

# Leistungsdiagnostik in der Schule

Herausgegeben von  
Kurt A. Heller

Mit Beiträgen von

Horst Bartnitzky, Duisburg  
Peter Birkel, Weingarten  
Peter Büscher, Heidelberg  
Reinhold Christiani, Düsseldorf  
Gerhard Eberle, Heidelberg  
Walter Fingerhut, Stuttgart  
Kurt A. Heller, München  
Ralf Horn, Weinheim  
Leo H. Klingen, Bonn  
Adam Kormann, Landshut

Reimer Kornmann, Heidelberg  
Andreas Krapp, München  
Hans-Peter Langfeldt, Bonn  
Erich Langhorst, Bonn  
Horst Nickel, Düsseldorf  
Bernhard Rosemann, Bochum  
Wilhelm Schreckenberg, Düsseldorf  
Michael Schümann, Hamburg  
Jürgen Wendeler, Gießen  
Wilhelm Wiczerkowski, Hamburg

4., völlig neubearbeitete Auflage

Verlag Hans Huber Bern Stuttgart Toronto

CIP-Kurztitelaufnahme der Deutschen Bibliothek

*Leistungsdiagnostik in der Schule* / hrsg. von Kurt A. Heller. Mit Beitr. von Horst Bartnitzky . . . - 4., völlig Neubearb. Aufl. - Bern; Stuttgart; Toronto: Huber, 1984.

Bis 3. Aufl. im Quelle-und-Meyer-Verl., Heidelberg

Bis 3. Aufl. u. d. T.: Leistungsbeurteilung in der Schule

ISBN 3-456-81345-7

NE: Heller, Kurt [Hrsg.]; Bartnitzky, Horst [Mitverf.]

22;11



# Inhaltsverzeichnis

<b>1.</b>	<b>Einleitung und Übersichtsreferat</b>	<b>15</b>
1.1	Zum Leistungsprinzip in der Schule	15
1.2	Aufgaben und Funktionen schulischer Leistungsdiagnostik	18
1.3	Das Gegenstandsproblem	20
1.4	Theoretische Grundlagen und Methoden der Schulleistungsdiagnostik	24
1.4.1	Testtheoretische Modelle und Methodenkonzepte	24
1.4.2	Objektive Verfahren der Leistungsmessung	31
1.4.3	Subjektive Verfahren der Leistungsbeurteilung	33
1.5	Schülerbeurteilung und Bildungsberatung	37
<b>2.</b>	<b>Schulleistung als pädagogisch-psychologisches Problem</b>	<b>39</b>
2.1	<i>Empirische Ansätze zur Aufklärung des Konstruktes «Schulleistung» (H.-P. LANGFELDT &amp; W. FINGERHUT)</i>	40
2.2	<i>Forschungsergebnisse zur Bedingungsstruktur der Schulleistung (A. KRAPP)</i>	46
2.2.1	Einleitung	46
2.2.2	Klärung einiger Begriffe	46
2.2.2.1	Schulleistung, Schulerfolg, Schulversagen	46
2.2.2.2	Bedingungsfaktoren, Determinanten, Ursachen	47
2.2.3	Ein Klassifikations- und Funktionsmodell möglicher Bedingungsfaktoren der Schulleistung	47
2.2.3.1	Intention und Grobstruktur des Modells	48
2.2.3.2	Differenzierung des Modells auf der Grundlage entwicklungs- theoretischer Überlegungen	49
2.2.4	Forschungsansätze zur Analyse der Bedingungen der Schulleistung	50
2.2.5	Forschungsergebnisse	54
2.2.5.1	Persönlichkeit des Schülers	54
2.2.5.2	Bedingungsfaktoren der familialen Umgebung	58
2.2.5.3	Bedingungsfaktoren der schulischen Umgebung	59
2.2.6	Ausblick	61

<b>3.</b>	<b>Theoretische Modelle der Schulleistungsdiagnostik</b>	<b>63</b>
3.1	<i>Die klassische Testtheorie als Grundlage normorientierter (standardisierter) Schulleistungstests (H.-P. LANGFELDT)</i> . . . . .	65
3.1.1	Messen und Testen . . . . .	65
3.1.1.1	Definition von Messen . . . . .	65
3.1.1.2	Skalen für Meßwerte . . . . .	66
3.1.1.3	Gütekriterien von Messungen . . . . .	68
3.1.2	Grundlagen der klassischen Testtheorie . . . . .	68
3.1.2.1	Fragestellung . . . . .	68
3.1.2.2	Axiome der klassischen Testtheorie . . . . .	69
3.1.2.3	Das Reliabilitätskonzept . . . . .	70
3.1.2.4	Validitätskonzepte . . . . .	72
3.1.2.5	Objektivitätsarten . . . . .	73
3.1.2.6	Zusammenhänge zwischen den Gütekriterien . . . . .	75
3.1.2.7	Zusammenfassung . . . . .	75
3.1.3	Reliabilität von Schultests . . . . .	75
3.1.3.1	Retest-Reliabilität . . . . .	76
3.1.3.2	Paralleltest-Reliabilität . . . . .	76
3.1.3.3	Halbierungs-Reliabilität . . . . .	77
3.1.3.4	Konsistenz-Reliabilität . . . . .	77
3.1.3.5	Das Konzept des Standardmeßfehlers . . . . .	78
3.1.4	Validität von Schultests . . . . .	79
3.1.4.1	Curriculare Validität . . . . .	79
3.1.4.2	Kriteriumsbezogene Validität . . . . .	80
3.1.4.3	Konstruktvalidität . . . . .	81
3.1.5	Objektivität von Schultests . . . . .	82
3.1.5.1	Durchführungs- und Auswertungsobjektivität . . . . .	82
3.1.5.2	Interpretationsobjektivität . . . . .	83
3.1.5.3	Normierung von Testergebnissen . . . . .	84
3.1.6	Formaler Aufbau standardisierter Schulleistungstests . . . . .	86
3.1.6.1	Aufgaben und Aufgabenanalyse . . . . .	86
3.1.6.2	Untertests und Testbatterien . . . . .	88
3.1.6.3	Ein praktisches Beispiel: Der AST 4 . . . . .	88
3.1.7	Diskussion . . . . .	92
3.1.7.1	Kritik an der klassischen Testtheorie . . . . .	92
3.1.7.2	Kritik an der Anwendung formeller Schulleistungstests . . . . .	94
3.1.8	Anhang: Zehn «Prüfsteine» zur Beurteilung von Schulleistungstests . . . . .	97
3.2	<i>Theoretische Ansätze der kriteriumsorientierten Leistungsmessung (P. BÜSCHER)</i> . . . . .	99
3.2.1	Terminologische Probleme . . . . .	99
3.2.2	Testtheoretische Probleme . . . . .	102

3.2.2.1	Unzulänglichkeit der klassischen Testtheorie bei kriteriumsbezogenen Tests .....	103
3.2.2.2	Itemanalyse bei kriteriumsbezogenen Tests .....	104
3.2.2.3	Reliabilitätsbestimmung bei kriteriumsbezogenen Tests .....	111
3.2.2.4	Validitätsbestimmung bei kriteriumsbezogenen Tests .....	113
3.2.3	Zusammenfassung .....	115
3.3	<i>Konzepte der Veränderungs- und Lernmessung (A. KORMANN)</i> .....	117
3.3.1	Definitions- und Abgrenzungsversuche .....	117
3.3.2	Meß- und testtheoretische Probleme .....	117
3.3.3	Methoden der Veränderungs- und Lernmessung .....	118
3.3.3.1	Graphische Darstellungen .....	118
3.3.3.2	«Klassische» Methoden .....	120
3.3.3.3	Beispiele für alternative Methoden der Veränderungs- und Lernmessung .....	121
3.4	<i>Anforderungsorientierte Leistungsdiagnostik (G. EBERLE &amp; R. KORNMANN)</i> .....	125
3.4.1	Zur Bestimmung individuell angemessener Lernanforderungen .....	126
3.4.2	Konzepte zur Strukturierung von Anforderungen .....	127
3.4.2.1	«Task Analysis» – Anforderungsanalyse .....	127
3.4.2.2	Die Analyse kognitiver Komponenten .....	128
3.4.2.3	Die Analyse der objektiven Anforderungsstruktur und der subjektiven Leistungsstruktur .....	129
3.4.3	Methodische Umsetzung .....	130
3.4.3.1	Die Fehleranalyse .....	130
3.4.3.2	Strukturorientierte Diagnostik .....	132
3.4.4	Abschließende Kritik .....	133
3.5	<i>Theoretische Grundlagen der Unterrichtsbeurteilung (W. SCHRECKENBERG)</i> .....	134
3.5.1	Der handlungstheoretische Aspekt .....	134
3.5.2	Der sozialwissenschaftliche Aspekt .....	137
3.5.3	Der pädagogisch-didaktische Aspekt .....	138
3.5.4	Das Beurteilungsmodell .....	138
3.5.5	Das Gewichtungssystem .....	140
4.	<b>Objektive Verfahren der Leistungsmessung</b> .....	143
4.1	<i>Zum Problem der Lernzieldefinition (R. HORN)</i> .....	145
4.1.1	Voraussetzungen der Leistungsmessung .....	145
4.1.2	Operationalisierte Lernziele .....	146
4.1.3	Erstellung von Prüfungsaufgaben .....	150

4.2	<i>Formelle (standardisierte) Schulleistungstests (R. HORN)</i> .....	154
4.2.1	Lesetests .....	154
4.2.2	Rechtschreibtests .....	155
4.2.3	Mathematiktests .....	157
4.2.4	Allgemeine Schulleistungstests .....	157
4.2.5	Fremdsprachentests .....	158
4.2.6	Tests für verschiedene Fächer .....	159
4.3	<i>Konstruktion und Auswertung informeller Schulleistungstests (Lernkontrolltests) (B. ROSEMANN)</i> .....	162
4.3.1	Einleitung .....	162
4.3.2	Die Formulierung der Items .....	162
4.3.2.1	Die Spezifikationstabelle .....	162
4.3.2.2	Itemtypen und ihre Konstruktion .....	165
4.3.3	Entwicklung der Test-«Vorform» .....	180
4.3.3.1	Gruppierung der Items und Erstellung des Testheftes .....	180
4.3.3.2	Aufgabenbewertung und Ermittlung der Gesamtleistung .....	182
4.3.4	Die Itemanalyse .....	183
4.3.4.1	Berechnung des Schwierigkeitsgrades einer Aufgabe .....	183
4.3.4.2	Berechnung der Trennschärfe einer Aufgabe .....	184
4.3.4.3	Distraktorenanalyse .....	186
4.3.5	Itemselektion und -revision .....	186
4.3.6	Die Reliabilität des Tests .....	189
4.3.7	Die Validität des Tests .....	191
4.3.8	Normorientierte Testauswertung .....	191
4.4	<i>Schulleistungsspezifische Lerntests (A. KORMANN)</i> .....	198
4.4.1	Zur theoretischen Legitimation der Lerntestmethodik .....	198
4.4.2	Voraussetzungen und Erwartungen an das Lerntestkonzept .....	199
4.4.3	Beispiele für schulleistungsspezifische Lerntests .....	200
4.4.3.1	Der Mengenfolgen-Test für Schulanfänger (MFT) .....	200
4.4.3.2	Der lerndiagnostische Rechtschreibtest von SCHNOTZ .....	201
4.4.4	Abschließende Bemerkung und Perspektiven .....	203
<b>5.</b>	<b>Subjektive Verfahren der Leistungsbeurteilung</b> .....	<b>205</b>
5.1	<i>Beobachtung und Beurteilung des Schülerverhaltens im Unterricht (E. LANGHORST)</i> .....	208
5.1.1	Fehlertendenzen bei der Beobachtung und Beurteilung menschlichen Verhaltens .....	208
5.1.2	Beobachtungsverfahren in der Schule .....	211
5.1.2.1	Beobachtung des Schülerverhaltens .....	211
5.1.2.2	Beschreibung des Schülerverhaltens .....	215

5.1.2.3	Beurteilung des Schülerverhaltens .....	220
5.1.3	Schlußbemerkungen .....	225
5.2	<i>Beurteilung mündlicher Prüfungsleistungen (P. BIRKEL)</i> .....	229
5.2.1	Zur Kritik an der mündlichen Leistungsprüfung .....	230
5.2.1.1	Sozialpsychologische Kritik .....	230
5.2.1.2	Psychoanalytische Kritik .....	231
5.2.1.3	Meßmethodische Kritik .....	231
5.2.2	Beurteilung mündlicher Leistungen im Unterricht .....	233
5.2.3	Vor- und Nachteile mündlicher vs. schriftlicher Prüfungen ....	235
5.3	<i>Schriftliche Prüfungen durch Klassenarbeiten (P. BIRKEL)</i> .....	237
5.3.1	Der klasseninterne Maßstab .....	237
5.3.2	Konsistenz und Konstanz von Klassenarbeitsnoten .....	240
5.3.3	Praxis der Leistungsfeststellung durch Klassenarbeiten .....	341
5.4	<i>Methodologische Probleme bei der Beurteilung von Schülerauf-</i> <i>sätzen (M. SCHÜMANN &amp; W. WIECZERKOWSKI)</i> .....	245
5.4.1	Grundprobleme in der Aufsatzbeurteilung .....	245
5.4.1.1	Allgemeines Modell der Beurteilung .....	246
5.4.1.2	Aufsätze als Ergebnis eines schriftsprachlichen Gestaltungsvor-	
	gangs .....	247
5.4.1.3	Aufsätze als sprachliches Handeln .....	248
5.4.2	Konstrukte der Aufsatzgüte als Basis für die Aufsatzbeurtei-	
	lung .....	249
5.4.2.1	Positiv und negativ formulierte Merkmale der Aufsatzgüte ....	250
5.4.2.2	Psycholinguistischer Ansatz der Aufgabenbewältigung .....	251
5.4.3	Kriterien der Aufsatzbeurteilung .....	253
5.4.3.1	Normative Kriterien der Aufsatzbeurteilung .....	253
5.4.3.2	Einige empirische Befunde .....	255
5.4.4	Aufsatzbeurteilung und Stellenwert von Aufsatzensuren .....	258
5.5	<i>Praktische Ansätze zur Beurteilung von Unterrichtsleistungen</i> <i>(W. SCHRECKENBERG)</i> .....	260
5.5.1	Die Unterschiedlichkeit der Beurteilungsmöglichkeiten im Be-	
	reich von Planung und Realisation des Unterrichts .....	260
5.5.2	Beurteilungsaspekte für die Unterrichtsplanung .....	261
5.5.2.1	Die Funktion der Unterrichtsplanung .....	261
5.5.2.2	Gütekriterien der Planung und ihre Indikatoren .....	262
5.5.3	Beurteilungsaspekte für die Realisation von Unterricht .....	264
5.5.3.1	Das Beobachtungssystem .....	264
5.5.3.2	Beobachtungssysteme und ihre Bedeutung für die Unterrichts-	
	beurteilung .....	265
5.5.3.3	Gütekriterien für die Realisation von Unterricht .....	266

5.5.3.4	Das Unterrichtsarrangement .....	267
5.5.3.5	Das Anspruchsniveau .....	269
5.5.4	Das Gewichtungsverhältnis der Beurteilungsaspekte für die Realisation des Unterrichts .....	270
5.5.5	Der «zentrale Beurteilungsaspekt» .....	270
<b>6.</b>	<b>Pädagogische Diagnostik und Schülerberatung .....</b>	<b>273</b>
6.1	<i>Schulreife und Schuleingangsdiagnostik</i> (H. NICKEL) .....	275
6.1.1	Wechselnde theoretische Konzepte und ihre praktischen Kon- sequenzen .....	275
6.1.1.1	Das entwicklungspsychologische Reifungskonzept .....	275
6.1.1.2	Das eigenschaftstheoretische Fähigkeitskonzept und seine Aus- wirkungen auf die Schuleingangsdiagnostik .....	275
6.1.2	Der interaktionistische ökopyschologische Ansatz .....	277
6.1.2.1	Schulreife als ökopyschologisches Konstrukt .....	277
6.1.2.2	Ein neues Schulreifemodell .....	278
6.1.3	Revision und Neugestaltung der Einschulungsdiagnostik .....	280
6.1.3.1	Einschulungsdiagnostik als Grundlage einer ökopyschologisch orientierten Interventionsstrategie .....	280
6.1.3.2	Schuleingangsdiagnostik als multidimensionaler Prozeß .....	281
6.2	<i>Förderungsdiagnostik im Primarbereich</i> (J. WENDELER) .....	283
6.2.1	Auslese- und Förderungsfunktion der Diagnostik .....	283
6.2.2	Fähigkeits- und Lernzieldiagnostik .....	284
6.2.3	Fähigkeitsdiagnostik in bezug auf das Lesen und Rechtschrei- ben .....	285
6.2.4	Lernzieldiagnostik der Rechtschreibfertigkeit .....	287
6.2.5	Die Diagnose von Lesefertigkeit und Leseverständnis .....	288
6.2.6	Diagnostik im Mathematikunterricht .....	289
6.3	<i>Neue Formen der Zeugnisschreibung in der Grundschule</i> (R. CHRISTIANI & H. BARTNITZKY) .....	292
6.3.1	Die Zeugnisregelungen .....	292
6.3.2	Der grundschulpädagogische Leistungsbegriff .....	292
6.3.3	Die Beurteilungspraxis ohne Zensuren .....	294
6.3.4	Das Verfassen zensurenfreier Zeugnisse .....	295
6.3.5	Zeugnisse im Beratungszusammenhang .....	298
6.4	<i>Schülerbeurteilung und Schullaufbahnberatung</i> (K. A. HELLER) .....	299
6.4.1	Modell der Schullaufbahnberatung .....	299
6.4.2	Der Diagnose-Prognose-Ansatz der Schuleignungsermittlung ..	300
6.4.2.1	Das Kriterienproblem .....	301

6.4.2.2	Auswahl der Prädiktoren .....	301
6.4.2.3	Verarbeitungs- und Entscheidungsstrategien .....	303
6.4.3	Paradigma der sukzessiven Entscheidungsfindung in der Schullaufbahnberatung .....	305
6.5	<i>Leistungsbewertung im Kurswahlssystem der Sekundarstufe (L. H. KLINGEN)</i> .....	308
6.5.1	Leistungsbeurteilung in einem durch sie begründeten Kurssystem .....	309
6.5.2	Leistungsbeurteilung in der Neigungsdifferenzierung .....	310
6.5.2.1	Grundsätzliche Bemerkungen zum Wahlverhalten .....	310
6.5.2.2	Darstellung des Systems der Grund- und Leistungskurse .....	310
6.5.2.3	Folgerungen für die Konstruktion valider Klausuren in beiden Kursarten .....	311
6.5.3	Grund- und Leistungskurse im Abiturbereich .....	312
6.5.4	Abschließende Bemerkungen .....	315
7.	<b>Literaturverzeichnis</b> .....	317
8.	<b>Autorenverzeichnis</b> .....	341
9.	<b>Personenregister</b> .....	343
10.	<b>Sachregister</b> .....	350



# 1. Einleitung und Übersichtsreferat

Die schulische Leistungsdiagnostik steckt voller Probleme, wie die Ausführungen in diesem Buch noch verdeutlichen werden. Dies gilt sowohl hinsichtlich der theoretischen Grundlegung als auch im Hinblick auf das Verfahrensangebot (Methoden). Sofern man Leistungsforderung und Leistungsbewertung als berechtigte Anliegen der Schule ansieht, womit jedoch keiner Verabsolutierung des Leistungsprinzips das Wort geredet wird, muß man sich auch den Problemen, die sich hierbei ergeben, stellen. Andernfalls praktiziert man *«Vogel-Strauß-Politik»* und huldigt bekannten Grundsätzen wie *«Wasch' mir den Pelz, aber mach' mich nicht naß!»* Aber selbst, wenn man Modeparolen folgend glaubt, dem Problem der Leistung in unserer Gesellschaft überhaupt aus dem Wege gehen zu können, begibt man sich meines Erachtens nur in neue Schwierigkeiten. Das Terrain, auf dem entsprechende Konflikte dann ausgetragen werden, mag zwar wechseln, etwa von der Leistungsebene auf das Feld ideologischer Auseinandersetzung (statt Leistungserweise werden dann *«Glaubens»-Bekennnisse* die Kriterien abgeben), der Problematik selbst wird man damit kaum begegnen (vgl. auch v. HENTIG 1970, S. 140ff.).

Somit wäre unsere Position abgesteckt: Wir erblicken in der schulischen Forderung nach Lernleistungen eine legitime Forderung. Diese stellt freilich nicht das einzige und vermutlich nicht einmal das wichtigste, gleichwohl ein unentbehrliches Bildungsziel dar. Erkennt man diese Aufgabenfunktion – neben anderen – an, so muß man auch die bei der Verwirklichung dieses Auftrags in Erscheinung tretenden Probleme artikulieren und wissenschaftlich begründete Wege (Methoden) zu ihrer Lösung aufzeigen. Zweifellos gilt ein solches Postulat zunächst für die pädagogische Funktion der Aktivierung von Schülerleistungen, erst danach erhält es seine volle Berechtigung im Hinblick auf die verschiedenen Ziele der Leistungsbeurteilung (z. B. im Rahmen unterrichtlicher Differenzierungsansätze, schulischer Effizienzkontrollen, der Schullaufbahnberatung usw.), die ja wiederum im Kontext schulischer Bildungsbemühungen zu sehen sind.

## 1.1 Zum Leistungsprinzip in der Schule

Mit dem Begriff der *Schulleistung* sei zunächst formelhaft *das gesamte Leistungsverhalten im Kontext schulischer Bildungsbemühungen* bezeichnet. Dabei verdienen der dynamische Aspekt (Lernprozeß) und der statische Aspekt (Lernprodukt) gleichermaßen Beachtung – entgegen manchen Begriffsvorstellungen, die nur das Ergebnis der Schülerleistung und nicht auch deren Bedingungsgefüge im Auge haben.

Diese doppelte Bedeutung des Leistungsbegriffs läßt sich auch etymologisch belegen, wie ZIEGENSPECK (1976, S.13f.) nachgewiesen hat. So steht das Wort *Leistung* für «Pflichthaftes» und bedeutet «etwas befolgen» oder «tun, was als Schuldigkeit vorgeschrieben ist». Das Substantiv *Leisten* kommt vom germanischen «laisti» gleich «Fußspur», während dem Verb *leisten* die Bedeutung «können, schaffen» oder auch «einer Spur nachgehen, nachspüren» zukommt. Diese Begriffe weisen auf den *dynamischen* Bedeutungsaspekt, das Prozeßhafte der Leistung hin.

Dagegen tritt bei den romanischen Wörtern, die auf die Wurzel «sta» zurückgehen, der *statische* Bedeutungsaspekt hervor. Dem lateinischen «praestare» entspricht im Italienischen heute noch «prestazione» (Leistung).

Analog zu dieser Etymologie unterscheidet BECKMANN (1978, S.45) einen prozeß- und einen produktorientierten Leistungsbegriff, wobei ersterer die psycho-physischen Prozesse miteinschließt, die beim Zustandekommen des Leistungsergebnisses beteiligt sind.

HECKHAUSEN (1974a, S.11ff.) bezeichnet *Leistung* als das «Ergebnis einer Untergruppe menschlicher Handlungen» und nennt fünf Bedingungen, die eine leistungsthematische Situation kennzeichnen: 1) Objektivierbarkeit und Aufgabencharakter des Handlungsergebnisses, 2) Bezug auf einen Schwierigkeitsmaßstab bzw. Kraftaufwand, 3) Möglichkeit des Gelingens oder Mißlingens einer Handlung, 4) Notwendigkeit subjektiv verbindlicher Gütemaßstäbe für die Bewertung des Handlungsergebnisses, 5) Urheberchaft des Handlungsergebnisses, d.h. dessen Abhängigkeit vom Handelnden selbst. Je nachdem, «welche Ursache man für das Zustandekommen des Handlungsergebnisses für ausschlaggebend hält, die Fähigkeit des Handelnden oder seine Anstrengung», unterscheidet HECKHAUSEN (1974a, S.19) zwischen einem fähigkeitszentrierten und einem anstrengungszentrierten Leistungsbegriff (vgl. noch SANG 1977, MIETZEL 1982).

Diese Merkmale des allgemeinen Leistungsbegriffs lassen sich unschwer reichsspezifisch auch auf Schulleistungen übertragen. Dabei wäre genauer zwischen dem *hypothetischen* Konstrukt «Schulleistung» und *empirischen* Begriffsvarianten (der Beobachtungssprache) zu unterscheiden (vgl. Kap. 2.1 und 2.2). Bezüglich der Operationalisierungsproblematik ist vor allem die Differenzierung in quantitative und qualitative Merkmale relevant: So liefern Zensuren oder Schulleistungstests *quantitative* Meßwerte (auf dem Niveau der Ordinal- und Intervallskala), während globale Urteile über «Schulerfolg» vs. «Schulversagen» u.ä. *qualitative* Beschreibungskategorien (auf dem Niveau der Nominal- bzw. Klassifikationskala) repräsentieren. Bei den quantitativen Variablen oder Indikatoren der Schulleistung handelt es sich somit um Merkmale mit kontinuierlich abgestuften Ausprägungsgrad, bei den qualitativen Kategorien hingegen um diskrete Daten; erreicht oder überschreitet der Schüler mit seiner Leistung bestimmte (Mindest-)Anforderungskriterien, gehört er zur Gruppe (Klasse) der «Erfolgreichen», unterschreitet er diese Norm, fällt er unter die Versagerquote (Repetenten bzw. Drop-outs). Zur ausführlicheren Information vgl. 2.2.2 und 3.1.1 in diesem Buch.

Die eingangs aufgestellte Definitionsformel schließt prinzipiell das Schüler- (bzw. Eltern-) und Lehrerverhalten ein, wenngleich die Vokabel «Schullei-

stung» häufig nur die Leistung des Schülers indiziert. Ein solchermaßen eingenger Leistungsbegriff ist jedoch nur akzentuierend, d.h. im Hinblick auf jeweils thematisierte Fragestellungen, angebracht. Die Gefahr, hierbei den Gesamt-komplex interdependenter Zusammenhänge aus dem Auge zu verlieren, ist freilich nicht von der Hand zu weisen. Nicht zufällig wurden deshalb die Beiträge von LANGFELDT & FINGERHUT zur Faktorenstruktur sowie von KRAPP zum Bedingungsgefüge der Schulleistung an den Anfang gestellt; möglichen Mißverständnissen sollte hiermit von vornherein begegnet werden. Wir müssen davon ausgehen, daß Schulleistungen mehrdimensional bedingt sind und von persönlichkeits- und sozialpsychologischen Faktoren abhängige Variablen(bündel) darstellen.

Umstrittener als diese Begriffsexplikation ist die Auffassung über das *Leistungsprinzip*, d.h. die Frage nach dem «Sinn und Unsinn des Leistungsprinzips in der Erziehungswissenschaft» (KLAFKI 1975). Publikationen hierzu sind fast Legion, und das Meinungsspektrum reicht von einseitigen Abqualifizierungen bis hin zu Interpretationsversuchen im Sinne von Selbstbestimmung bzw. Selbstverwirklichungsideal und Steigerung individueller Lebensqualität (vgl. LICHTENSTEIN-ROTHER 1964 und 1971, GEHLEN et al. 1975, SCHMÖLDERS 1975, BECKMANN 1978, NIPKOW 1978, SCHALLER 1979, PFISTNER 1981, SCHNITZER 1981, TEUSCH 1983 u.a.). Auch die Forderung nach Relativierung des Leistungsprinzips wird des öfteren erhoben (z.B. v. HENTIG 1970, KOZDON 1976, FLITNER 1980), wobei jeweils unterschiedliche Funktionsaspekte in den Vordergrund gerückt werden, etwa Probleme schulischer Leistungsdifferenzierung (GAUDE & TESCHNER 1970, HAUSSER 1981) oder soziale Gesichtspunkte der Chancengerechtigkeit (HECKHAUSEN 1974a, GEHLEN 1975, KLAFKI 1975, SCHALLER 1979, FLITNER 1980).

FURCK wies auf die doppelte Funktion der Schulleistung hin: die Persönlichkeitsbildung des Schülers und die Sicherung des volkswirtschaftlichen Leistungspotentials. «Das Problem der Leistung in der Schule spitzt sich so zu der Frage nach dem rechten Verhältnis von individueller Bildsamkeit und ihr angemessener Anforderung zu» (1975, S.118). Damit ist das *pädagogische Leistungsprinzip* artikuliert. Ähnlich betont der Strukturplan für das Bildungswesen (1970, S.35f.), daß die Forderung nach Leistung in der Schule unter dem pädagogischen Prinzip der *individuellen Förderung* stehen müsse. Dies schließt die Einübung des Wettbewerbs in altersgemäßen Formen ebenso ein, wie es andererseits nicht zur Sozialauslese führen darf. Eine solche Gefahr würde bestehen, «wenn die geforderten Leistungen nicht aufgrund der schulischen Lernprozesse, sondern nur oder vornehmlich aufgrund der Herkunft des Lernenden aus einem bestimmten sozialen Milieu erbracht oder nicht erbracht werden können». Sehr detailliert und ausführlich geht HECKHAUSEN (1974) im 10. und 11. Kapitel seines Buches über «Leistung und Chancengleichheit» auf die hier angesprochenen Probleme ein. Demnach ergeben sich zwischen dem Gleichheits- und dem Leistungsprinzip viele Widersprüche, die der Autor in zehn Dilemmata thesenartig zusammenfaßt. Seiner Meinung

nach kann das Bildungssystem nur einen Teil zur Chancengleichheit beitragen, wozu von ihm programmatisch verschiedene Ansätze vorgeschlagen werden.

## 1.2 Aufgaben und Funktionen schulischer Leistungsdiagnostik

Die Notwendigkeit von Schulleistungsdiagnosen ergibt sich aus übergeordneten Erziehungs- und Bildungszielen. Als Leitziel können dabei die *Persönlichkeitsbildung* (FURCK) und die *Erziehung zur Mündigkeit* (KLAFFKI), d.h. Selbst- und Mitbestimmungsfähigkeit des jungen Menschen, betrachtet werden. Im Zusammenhang mit diesem pädagogischen Auftrag stehen verschiedene Aufgaben pädagogisch-psychologischer Diagnostik, die hier als die didaktische Funktion, die evaluative Funktion und die Entscheidungsfunktion kurz dargestellt werden sollen.

### *(1) Didaktische Funktionen der Schulleistungsdiagnostik*

Diese erstrecken sich sowohl auf die curriculare Planung und Unterrichtsorganisation als auch auf die Lehr- und Lernsteuerung im Unterrichtsprozeß. STARK et al. (1973, S. 12f.) sehen die «didaktische Primärfunktion» schulischer Leistungsbeurteilung in der Überprüfung unterrichtlicher Maßnahmen «mit der Absicht, sie zu belassen, zu modifizieren oder zu korrigieren». Ähnlich äußert sich ZIELINSKI (1973, S. 58): «Nicht die Beurteilung und Benotung eines Schülers, die fast schon zum Selbstzweck geworden ist, sondern die Bewertung des Unterrichtserfolges und die Erfassung von Lernschwierigkeiten ist u. E. die vornehmste Aufgabe pädagogisch-psychologischer Prüfmethode, ausgerichtet auf das Ziel, rechtzeitiges und wirkungsvolles pädagogisches Handeln zu ermöglichen und zu initiieren.» Demnach dienen schulische Leistungskontrollen vorab zur Überprüfung aufgestellter Lernziele und zur Diagnose des Lehr-Lernprozesses. Die so gewonnenen Informationen bieten unverzichtbare Hilfen für die Curriculumentwicklung und -revision, die Lernzielbestimmung, die didaktische Planung und Analyse des Unterrichtsverlaufs usw. Die Leistungsbeurteilung ist somit notwendiger Bestandteil einer optimalen Unterrichtsgestaltung.

Unter dem Aspekt der (inneren) Bildungsreform wäre hier die Wechselwirkung von Curriculumentwicklung und Leistungsbeurteilung hervorzuheben (vgl. PETRI 1974). So erfordert die Überprüfung festgelegter Curricula, z. B. im Hinblick auf schulpädagogische und lernpsychologische Möglichkeiten, jeweils bestimmte Methoden ihrer Bewertung. Theoretisch bedeutet dies, daß *alle* kognitiven und nichtkognitiven Determinanten der konkret geforderten Schülerleistung in einem solchen Bewertungssystem Berücksichtigung finden müssen. Praktisch ist man dieser Idealforderung bislang allenfalls bezüglich

der Kontrolle der kognitiven Variablen nahegekommen, wohingegen nicht-kognitive Lernvoraussetzungen nur selten einbezogen wurden (vgl. Kap. 2.2 in diesem Buch). Andererseits bewirkt die Entwicklung bestimmter Beurteilungsmethoden oft auch Modifikationen und Ergänzungen schulischer Curricula, deren Beschränkung auf kognitive Inhalte keine Notwendigkeit darstellt. Daraus läßt sich die Forderung ableiten, die Diskussion über schulstestmethodische und curriculare Neuerungen nicht isoliert, sondern in wechselseitiger Abstimmung voranzutreiben. Nur so kann man der Gefahr unerwünschter Stabilisierungseffekte wirksam begegnen.

### *(2) Evaluative Funktionen der Schulleistungsdiagnostik*

Dieser Aspekt der Leistungsbeurteilung zielt auf die «Auto-Evaluation», die Orientierung des Schülers (bzw. seiner Eltern) über seine eigenen Schulleistungen, seine Lernfortschritte (intraindividueller Vergleich) und seine Position innerhalb der Klassen- oder Altersgruppe (interindividueller Vergleich). Dazu ist zweierlei notwendig: erstens die genaue Kenntnis der Lernanforderungen, was eindeutig definierte Lernziele erfordert, und zweitens die Transparenz des Beurteilungssystems bzw. der Bewertungsmaßstäbe. Ferner müssen dem Schüler Notwendigkeit, Zweck und Formen der Leistungsbeurteilung einsichtig gemacht werden.

Während im vorigen Abschnitt unterrichtsdidaktische und somit lehrerzentrierte bzw. gruppenimmanente Interessen (der Schulklasse bzw. Lerngruppe) im Vordergrund der Betrachtung standen, kommt unter dem Gesichtspunkt schülerzentrierter Aufgabenfunktionen der *Individualdiagnose* besondere Bedeutung zu. Hier sind diagnostizierte Begabungs-Leistungsdiskrepanzen (sog. Underachievement), vorübergehende oder andauernde Lern- und Leistungshemmungen (z.B. Konzentrationsstörungen), aber auch partielle vs. allgemeine Leistungsschwächen (z.B. Legasthenie, Rechenschwäche usw. vs. Lernbehinderung[en] im Sinne der Sonderschulbedürftigkeit) oft unerläßliche Voraussetzung für die Bestimmung angemessener (schul)pädagogischer Maßnahmen und therapeutischer Behandlungsansätze. Darüber hinaus will – und sollte – jeder Schüler, auch der «unauffällige», kontinuierlich über den persönlichen Stand und Fortschritt und erst recht über eventuelle Rückschritte im Lernprozeß informiert werden. Ob diese Information durch Tests, sog. Diagnosebögen oder Zensuren erfolgt, ist prinzipiell von zweitrangiger Bedeutung, sofern die verwendeten Methoden der Leistungsbeurteilung hinreichend objektiv, verlässlich und gültig sind. Daß diese Anforderungskriterien in der Praxis der Schülerbeurteilung – besonders bei den subjektiven Verfahren (z.B. der Notengebung) – häufig nicht erfüllt sind, betrifft eine Reihe von Problemen, die im folgenden noch ausführlich zur Sprache kommen.

### *(3) Entscheidungsfunktionen schulischer Leistungsanalysen*

Die Schulleistungsdiagnostik spielt auch eine wichtige Rolle bei der Schullaufbahnberatung bzw. – allgemein – beim Wechsel von Lerngruppen, d. h. überall

dort, wo Fragen der äußeren oder inneren Differenzierung im Schulsystem zum Problem werden. Frühere und/oder aktuelle Lernleistungen des Schülers bilden hier eine wichtige Grundlage für die Vorhersage des (künftigen) Schulerfolges. Dieser ist nach relevanten Lernbedingungen hin zu analysieren, was sowohl (kognitive und nichtkognitive) Schülermerkmale als auch Determinanten der sozialen (schulischen und familialen) Lernumwelt einschließt. Die Frage, unter welchen Bedingungen eine bestimmte Schulleistung erbracht wird oder zu erwarten ist, läßt sich umso zuverlässiger und gültiger beantworten, je umfassender und differenzierter diagnostische Informationen über den Bedingungskomplex des Schulerfolgs im Einzelfall zur Verfügung stehen (vgl. Kap. 6.4). Eine entsprechende Schulleistungsbeurteilung darf also nicht nur die Lernresultate (das Leistungsprodukt) berücksichtigen, sondern muß auch die persönlichen und sozialen Begleitumstände schulisch geforderter Leistungsprozesse in Betracht ziehen. Hierfür stehen unterschiedliche Methodenansätze der sog. Status- vs. Prozeßdiagnostik (z.B. Lerntests) zur Verfügung (vgl. ausführlicher KRAPP & PRELL 1978, SÜLLWOLD 1983).

Die einzelnen Aufgaben der Schulleistungsdiagnostik stehen nicht beziehungslos nebeneinander. Vorrangiges Ziel sollte die Unterstützung der Arbeit in der Schule sein, wozu die skizzierten Beurteilungsfunktionen in unterschiedlichem Maße beitragen.

### 1.3 Das Gegenstandsproblem

Gegenstand schulischer Leistungsdiagnosen sind sowohl Inhaltsdimensionen als auch Bedingungsfaktoren des Schulerfolgs bzw. einzelner Lernprobleme. Zunächst wird auf die *Hauptdimensionen* der Schulleistung eingegangen, um das Konstrukt «Schulleistung» zu erklären. Ausgehend von den üblichen Indikatoren der Schulleistung versuchen LANGFELDT & FINGERHUT in ihrem Beitrag eine empirische Beschreibung dieses Phänomens. Die zu diesem Zweck faktorenanalysierten Schulzeugnisse teils eigener, teils fremder Provenienz deuten übereinstimmend an, «daß die Schulleistung – so wie sie durch Schulnoten erfaßt wird – bei weitem nicht so differenziert ist, wie es Zeugnisse mit zehn und mehr Einzelnoten nahelegen». Für die Schülerleistungen auf der Sekundarstufe konnten lediglich drei Faktoren nachgewiesen werden: ein Fremdsprachenfaktor, ein mathematisch-naturwissenschaftlicher Faktor und ein sachkundlicher Faktor (nach DENIG & WEIS 1970). Für die Primarstufe (Grundschüler und lernbehinderte Schüler) ergaben sich sogar nur zwei Faktoren: ein allgemeiner Schulleistungsfaktor (vor allem auf die Fachzensuren in Deutsch und Rechnen bezogen) und ein Faktor der schulischen Disziplin (mit entsprechenden Ladungen auf den sog. Kopfnoten). Die Ergebnisse von FUNKE (1972) und ZIMMERMANN (1968) bestätigen – unabhängig voneinander –

diese gegenüber dem wesentlich differenzierteren Zeugnisbild stark reduzierte Faktorenstruktur der tatsächlichen Schülerleistung. Im Hinblick auf die Schulleistungsbeurteilung herkömmlicher Art (Zensurierung) ergäbe sich hieraus die Konsequenz, künftig nicht mehr als zwei bzw. drei unabhängige Zeugnisnoten zu vergeben. Zumindest unter dem Gesichtspunkt streng dimensionsbezogener Beurteilung wären mehr als drei Einzelnoten überflüssig. ROEDER & TREUMANN (1974) konnten jedoch nachweisen, daß diesen Hauptdimensionen der Schulleistung jeweils spezifische *Leistungsbereiche* entsprechen. Darunter sind Fachgebiete oder Teilbereiche eines Unterrichtsfaches, z. B. Vektorrechnung in der Mathematik oder Aussprache und Intonation im Englischen, zu verstehen. Unterschiedlichen Leistungsbereichen (wie Mathematik und Latein) kann somit ein und dieselbe Leistungsdimension (abstrakt-logisches Denken) zugrunde liegen. Die angeschnittene Frage nach der erforderlichen oder wünschenswerten Anzahl von Fachnoten wäre also auch unter der didaktischen bzw. schulpädagogischen Perspektive zu reflektieren.

Sofern man neben den Lehrerurteilen in Form von Zensuren noch weitere Daten, z. B. Schulleistungstests, Intelligenztests, biographische Informationen, in die Faktorenanalyse miteinbezieht, gewinnt man ein höher strukturiertes Modell der Schulleistung. Wie FINGERHUT & LANGFELDT (1974) unter Bezug auf ein entsprechendes Datenmaterial von 1756 Viertkläßkinder nachweisen konnten, ergeben sich dann folgende vier Faktoren: 1) «Schulische Leistungsfähigkeit» mit Ladungen auf den Zensuren der Kernfächer (Deutsch, Rechnen, Heimatkunde), sämtlichen Skalen des AST 4 (Allgemeiner Schulleistungstest von FIPPINGER) und den Untertests 3 + 4 des LPS (Leistungsprüfsystem von W. HORN); 2) «Genereller Notenfaktor» mit Ladungen auf den Fachleistungs- und Kopfnoten sowie den Variablen Alter und Geschlecht; 3) «Test-Intelligenz» mit Ladungen auf sämtlichen Skalen des LPS; 4) «Rechenleistungen» mit Ladungen auf der Rechennote, den Rechenskalen des AST 4 und den LPS-Subtests 3 + 4, 14 und 15 (Arbeitsprobe).

Diese Ergebnisse lassen einen gewissen Zusammenhang zwischen einigen Intelligenzfaktoren (Reasoning, Accuracy, Number) und der (Grund-)Schulleistung, besonders hinsichtlich des allgemeinen Leistungsstandards und der Rechenleistung, erkennen. Die gerade in der jüngeren Literatur häufig vertretene Auffassung, wonach zwischen Intelligenz und Schulleistung keine oder nur schwache Zusammenhänge bestehen, muß demnach korrigiert bzw. folgendermaßen präzisiert werden. «Es kann . . . nur gesagt werden, daß Schulnoten und Intelligenz relativ unabhängig voneinander sind», nicht aber Schulleistung und Intelligenz. Da Schulnoten allein nach den Befunden von LANGFELDT & FINGERHUT kein adäquates Abbild der Schulleistung geben, können diese «nur in mäßigem Umfange etwas über die *tatsächlichen* Bedingungen der Schulleistung aussagen». Dieses Ergebnis erhellt indirekt die Bedeutung kognitiver Faktoren im Hinblick auf die Konstituierung von Schulleistungen, ohne hiermit den Einfluß nichtkognitiver Determinanten als gering einzuschätzen.

Eine modifizierte Faktorenanalyse der Daten von FINGERHUT & LANGFELDT ergab schließlich fünf abgrenzbare Cluster: 1) Intelligenztestleistungen (LPS), 2) Schultestleistungen (AST 4), 3) Fachleistungsnoten (Deutsch, Rechnen, Heimatkunde), 4) Kopfnoten (Betragen, Fleiß, Aufmerksamkeit, Ordnung), 5) Merkmale biographischer Art (Alter, Geschlecht, soziale Herkunft). Aus der Distanz der Cluster zueinander postulieren die Autoren jetzt ein zweidimensionales Konstrukt «Schulleistung», demzufolge sich Schulleistungen im gegenwärtigen Bildungssystem als Funktion von «tatsächlicher Leistung aufgrund von Begabung» (Faktor I) und «Anpassung an dieses System» (Faktor II) erweisen. Siehe dazu Abb. 1 in Kap. 2.1.

Der folgende Beitrag von KRAPP (Kap. 2.2) beschäftigt sich mit der *Bedingungsstruktur* von Schulleistungen. Nach einschlägigen Begriffsklärungen wird ein Klassifikations- und Funktionsmodell der Schulleistung mit den drei Merkmalskomplexen Schülerpersönlichkeit, Familie und Schule entworfen. Mit dem Schulleistungskriterium stehen nur die Variablen der Schülerpersönlichkeit unmittelbar in Beziehung, während die Bedingungsmerkmale auf seiten der Familie und Schule indirekt – über die Beeinflussung der Schülerpersönlichkeit – auf das Kriterium wirken. In einem erweiterten Ansatz werden aus entwicklungstheoretischer Sicht *frühere* Entwicklungsbedingungen von *aktuellen* Einflußgrößen unterschieden. Diesen entspricht – in Anlehnung an motivationspsychologische Konzepte – das Begriffspaar «*habituelle* Lernvoraussetzungen» vs. «*aktualisierte* Leistungsbereitschaften».

Die Erfassung früherer Entwicklungsbedingungen erfordert zeitliche Längsschnittstudien, während aktuelle Bedingungsfaktoren (sog. Dispositionen) zeitlichen Querschnittsanalysen zugänglich sind, wobei die Unterscheidung von *Strukturmerkmalen* (Dispositionen) und *Prozeßmerkmalen* relevant ist. Typische Forschungsansätze der pädagogischen Psychologie sind im aufgezeigten Problemfeld Gruppenvergleiche (z. B. erfolgreicher und nichterfolgreicher Schüler) bzw. (einfache) Korrelationsstudien, sog. Ex-post-facto-Experimente, Untersuchungen zum Konzept partieller Schulleistungsschwächen (Underachievement, Legasthenie, Arithmasthenie oder Dyskalkulie), multiple Regressionsanalysen, Interaktionsanalysen (ATI-Forschung) und Moderatoranalysen sowie Pfadanalysen auf der Basis von Kausalmodellen.

Empirische Forschungsbefunde liegen vor allem zu *habituellen* Lern- und Leistungsvoraussetzungen vor. Als Einzelprädiktor erklärt demnach die Intelligenz den größten Anteil, nämlich 30 – 50% der Kriteriumsvarianz (Schulleistung). Dabei variiert die Höhe des Korrelationskoeffizienten in Abhängigkeit von anderen Faktoren. So sinkt nach einer Untersuchung von SIMONS et al. (1975) die Korrelation zwischen Intelligenz und Mathematikleistung (in der Grundschule) im Laufe der Zeit, während gleichzeitig die Bedeutung unterrichtsspezifischer *Vorkenntnisse* (sensu GAGNÉ) zunimmt, was sich in ansteigenden Korrelationskoeffizienten ausdrückt. Ebenso besteht eine Beziehung zur Unterrichtsqualität: «In «schlecht geführten» Klassen hat die (allgemeine) Intelligenz ein vergleichsweise stärkeres Gewicht als in «gut geführten» Klas-

sen.» Solche Befunde deuten auf die Existenz «*kognitiver Fähigkeitsschwellen*» hin, deren Unterschreiten zu Schulversagen führen kann, während oberhalb der hypothetischen Fähigkeitsschwelle die Auswirkung der Intelligenz auf das schulische Leistungsverhalten in stärkerem Maße von anderen Faktoren (konstellationen) abhängt, d. h. «hier ist eine gegenseitige Kompensation von Bedingungsvariablen möglich» (KRUG & RHEINBERG 1980, S. 58). Mit der gleichen Hypothese läßt sich der Tatbestand interpretieren, daß der korrelative Zusammenhang zwischen Intelligenz und Schulleistung im allgemeinen bei älteren Schülern weniger eng ausfällt als bei jüngeren. In eine ähnliche Richtung weisen auch Beobachtungen, wonach Intelligenz(test)leistungen und Schulnoten auf niedrigerem Bildungsniveau enger korreliert sind als auf höherem, z. B. Gymnasium (vgl. GAEDIKE 1974, S. 47ff.). Und daß dieser Zusammenhang bei Mädchen durchweg enger ist als bei Jungen, sei nur am Rande erwähnt. KEMMLER (1967, 1976) interpretiert diesen Befund mit dem Hinweis auf geschlechtsspezifische Differenzen, etwa eine bessere Aktivierung des Intelligenzpotentials aufgrund günstigerer Arbeitshaltung bei Mädchen.

Neben der Intelligenz wurde auch die *Kreativität* als Einflußfaktor der Schulleistung untersucht. Im Vergleich zur Intelligenz bzw. «konvergenter Denkproduktion» (sensu GUILFORD) korreliert Kreativität («divergentes Denken») deutlich niedriger mit dem Schulleistungskriterium, was durch schulische Präferenzen bezüglich konvergenter Denkaufgaben erklärt werden mag. Die im Laufe der Schulzeit zunehmende Bedeutung fachspezifischer *Vorkenntnisse* als Bedingungsvariable der Schulleistung geht auch aus der relativ hohen Korrelation (.65 bis .85) zwischen zeitlich aufeinanderfolgenden Zeugnisnoten der wichtigsten Unterrichtsfächer hervor.

Sieht man von somatischen und konstitutionellen Dispositionen ab, dann waren vor allem motivationale und affektive Merkmale als sog. *nichtkognitive Persönlichkeitsfaktoren* Gegenstand pädagogisch-psychologischer Untersuchungen. *Leistungsmotivation* und *Schulangst* sind hier besonders einflußreiche Bedingungen für den Schulerfolg bzw. das Schulversagen. Während – als gesichertes Ergebnis – «die Motivationsstärke von der subjektiv erlebten Aufgabenschwierigkeit abhängt, wobei für Erfolgs- bzw. Mißerfolgsmotivierte jeweils andere Schwierigkeitsgrade motivierend sind, . . . ist die Beziehung zwischen Motivationsstärke und Güte und Qualität einer Leistung» nach KRAPP noch weithin unklar. Wahrscheinlich ist hierfür eine kurvilineare Beziehung charakteristisch. Versteht man Schulleistung als kumulativen Lernprozeß, dann wäre die Motivationsstärke nach ATKINSON (1974) für die Aktualisierung kognitiver Leistungen und den arbeitszeitlichen Aufwand verantwortlich zu machen.

Zu dem Bedingungskomplex des *sozialen Lernumfeldes* zählen einmal *familiäre* Sozialisationsvariablen. Diese «Umweltkräfte» (MARJORIBANKS) oder ökologischen Faktoren lassen sich in *Strukturmerkmale* (Familiengröße, Einkommensverhältnisse, Wohnsituation usw.) und in *Prozeßmerkmale* (z. B. elterlicher Erziehungsstil, Sprachkode oder «Zeitperspektive» sensu LEWIN)

aufgliedern. In der neueren Forschung werden vor allem kausale Strukturanalysen (Pfadmodelle) auf der Grundlage familialer Sozialisationstheorien durchgeführt.

Zum andern wurde – analog zur Familienpsychologie – der Einfluß der *Schule* zur Erklärung individueller Leistungsunterschiede bei Schülern untersucht. Während – entgegen bekannter Alltagstheorien – *schulorganisatorische* Rahmenbedingungen sich als relativ bedeutungslos erwiesen haben, gibt es zahlreiche empirische Belege für den Einfluß der *Lehrinhalte* und *schulischen Lernbedingungen*. Dabei muß man sich die Beziehung von Lehrerverhalten (Unterrichtsqualität, Verhaltensstile) und Schülerleistung als *differentielle «Effektmuster»* (ATI-Forschung) vorstellen. Aber auch solche Effekte sind wohl eher indirekt hervorgerufen. KRAPP resümiert abschließend: «Lehren, Lernen und Leistungsverhalten sind nach neuem Verständnis eingebettet in dynamische, sich wechselseitig beeinflussende Ereignisketten.» Entsprechend diesem Paradigmawechsel ist eine Unterscheidung von «unabhängigen» und «abhängigen» Variablen hier praktisch kaum mehr möglich, wengleich die Erforschung spezifischer Einflußgrößen im erörterten Problemkontext nach wie vor ein wichtiges Ziel der pädagogischen Psychologie darstellt.

## 1.4 Theoretische Grundlagen und Methoden der Schulleistungsdiagnostik

### 1.4.1 Testtheoretische Modelle und Methodenkonzepte

Die *klassische Testtheorie*, auch *Meßfehlertheorie* genannt, bildet nach wie vor die Hauptgrundlage standardisierter (formeller und informeller) Schulleistungstests. Jeder, der sich solcher Meßverfahren im Rahmen der Schülerbeurteilung bedient, sollte deshalb – schon um unkritischer Anwendung und Fehleinschätzungen vorzubeugen – die theoretischen Voraussetzungen testdiagnostischen Vorgehens in etwa kennen. Da die klassische Testtheorie unabhängig vom Inhalt der einzelnen Verfahren (z. B. Intelligenztest, Leistungstest usw.) gilt, maß man diesem Konzept in der psychologischen Diagnostik lange Zeit praktisch uneingeschränkte Bedeutung bei. Obwohl heute die Relativierung dieses Anspruchs unbestritten ist, finden sich daneben noch Tendenzen, die Theorie als solche überhaupt abzulehnen. Ein solcher Schritt wäre jedoch übereilt, wie die Ausführungen von LANGFELDT – und später von BÜSCHER – belegen.

Das Wort «Test» bedeutet ursprünglich «Prüfung», «Stichprobe» oder «Zeugnis», wobei sich der Begriff in der Testdiagnostik vor allem auf standardisierte Verhaltensstichproben bezieht, d. h. auf Prüfungen, die unter genau festgelegten Bedingungen durchgeführt werden. Solche Prüfverfahren oder *Tests* messen bestimmte Merkmalsausprägungen (z. B. Schulleistung, Intelligenz, Ängst-

lichkeit usw.), indem sie *interindividuelle* Unterschiede (Unterschiede zwischen den einzelnen Individuen) oder *intraindividuelle* Unterschiede (Unterschiede in bezug auf dieselbe Person, etwa im Verlauf der Ontogenese) erfassen. Bei den sog. standardisierten Schulleistungstests steht der erste Aspekt im Vordergrund, d. h. der Vergleich eines individuellen Meßwertes mit der Gruppennorm (Durchschnittsleistung der Bezugsgruppe, z. B. Schulklasse oder Jahrgangsstufe).

In dem testtheoretischen Beitrag von LANGFELDT (Kap. 3.1) werden zunächst die Begriffe «Messen» (Testen) und «Meßskala» sowie die wichtigsten Anforderungskriterien in bezug auf Messungen, auch «Gütekriterien» genannt, erläutert. *Messen* (Testen) ist ein Vorgang des Vergleichens anhand eines Maßstabs (Testnormen), wobei die Meßwerte durch unterschiedliche Skalenniveaus repräsentiert sein können. Die wichtigsten Skalentypen sind: 1) die *Nominal-* oder *Klassifikationsskala* (unterste Ebene des Messens), die nur dem Kriterium der Äquivalenz genügt; 2) die *Ordinal-* oder *Rangskala*, die zwei Kriterien, dem der Äquivalenz und dem der rangmäßigen Ordnung entspricht; 3) die *Intervallskala* mit den Kriterien der Äquivalenz, der Rangbeziehung und der Intervallkonstanz; 4) die *Rational-* oder *Verhältnisskala* (oberste Ebene des Messens), die zusätzlich zu den aufgeführten Kriterien einen absoluten Nullpunkt aufweist. Schulleistungstests messen in der Regel auf Rangskalenniveau (z. B. in Form von Prozenträngen = PR), selten auf Intervallskalenniveau (z. B. T-Werte).

Unabhängig vom jeweiligen Skalenniveau wird der Interpretationsspielraum noch durch die sog. *Testgütekriterien* (Objektivität, Reliabilität, Validität) beeinflusst. *Objektivität* meint hier die Unabhängigkeit der Testergebnisse von der Person des Testleiters sowohl in bezug auf die Testanweisung (Instruktion) als auch in bezug auf die Testauswertung und Testinterpretation. Die Durchführungsbestimmungen eines Tests müssen so präzise festgelegt sein, daß intersubjektivität gewährleistet ist. Diese Forderung ist in der Praxis der Testdurchführung nicht immer leicht zu erfüllen. Die *Reliabilität* oder Zuverlässigkeit bezieht sich auf die Meßgenauigkeit. Ein Schulleistungstest ist beispielsweise dann reliabel (zuverlässig), wenn die Ergebnisse unabhängig vom Zeitpunkt der Messung zustandekommen, d. h. wenn die wiederholte Anwendung des betr. Tests bei derselben Klasse oder einzelnen Individuen (in etwa) zum gleichen Resultat führt. Es gibt verschiedene Aspekte der Reliabilität und dementsprechend unterschiedliche Methoden ihrer Kontrolle, die alle von LANGFELDT besprochen werden. Die *Validität* oder Gültigkeit bezieht sich auf die Genauigkeit, mit der ein Test das mißt, was er messen soll. So sagt die Validität eines Rechentests etwas darüber aus, wie genau tatsächlich Rechenkenntnisse (und nicht etwa Konzentrationsfähigkeit oder andere Merkmale) erfaßt werden. Während sich also die Reliabilität auf die Zuverlässigkeit des Meßinstrumentes (Tests) selbst und damit auf die formale Meßgenauigkeit bezieht, informiert die Validität darüber, «welche psychodiagnostischen Schlußfolgerungen die numerischen Resultate eines Tests zulassen und welchen Grad an

Sicherheit solche Schlußfolgerungen aufweisen» (MICHEL 1964, S.47). Damit ist die diagnostische Valenz oder Treffsicherheit eines Testverfahrens angesprochen. Was endlich den Zusammenhang zwischen Reliabilität und Validität betrifft, so gilt: Die Reliabilität ist die notwendige, aber keine hinreichende Voraussetzung für die Validität eines Tests. Sehr oft erweist sich die Validierung eines Schultests als das schwierigste Unterfangen überhaupt, wobei man sich nicht selten mit der Bestimmung der curricularen Gültigkeit begnügt. Seltener werden hier die Methoden der Kriterienvalidierung (z. B. zur Ermittlung der Übereinstimmungsvalidität von Testleistung und Schulnoten als «Außenkriterien») oder der Konstruktvalidierung eingesetzt. Bei der Konstruktion eines Schulleistungstests kommt man den Anforderungskriterien der Reliabilität und Validität dadurch entgegen, daß man schon frühzeitig – in der sog. Aufgaben- oder Itemanalyse – *Schwierigkeit* und *Trennschärfe* jeder einzelnen Testaufgabe ermittelt, um von vornherein die unbrauchbaren Aufgaben (Items) auszusondern.

Damit nun die Ergebnisse eines Schulleistungstests zuverlässig interpretiert werden können, sind zwei weitere Voraussetzungen unerläßlich: die Normierung des Tests und die Ermittlung des sog. (Standard-)Meßfehlers. In den *Testnormen* liegen hierfür relativierte Testwerte, d. h. auf die jeweilige Alters- oder Klassengruppe bezogene Vergleichsmaßstäbe vor. So besagt eine Schülerleistung auf dem 75. PR, daß der betreffende Schüler in dem untersuchten Schulfach bessere Leistungen erzielt als 75% seiner Bezugsgruppe (und schlechtere Leistungen als 25%), während eine Leistung auf dem 50. PR durchschnittliche Leistungsfähigkeit indiziert. Nun wäre es naiv anzunehmen, daß die ermittelte (beobachtete) Testleistung von PR = 75 exakt die tatsächliche (wahre) Leistungsfähigkeit des Schülers wiedergibt; vielmehr muß davon ausgegangen werden, daß in diesem Testergebnis – wie in jedem Meßwert – eine zunächst unbekannte Fehlerquote steckt, die zu Lasten der Irreliabilität bzw. mangelnden Objektivität einer Untersuchung geht. Diese Annahme trifft im Kern das *Meßfehlerkonzept*, das wichtigste Axiom der klassischen Testtheorie. Danach setzt sich jeder *beobachtete Wert* (Meßwert) aus dem «wahren Wert» (hypothetisch einem zeitlich konstanten Parameter) und einem «Fehlerwert» (Meßfehler) zusammen. Beide werden als unkorrelierte Größen postuliert, wobei der Meßfehler als Zufallsvariable auftreten soll. Mit dessen Hilfe werden dann – auf der Basis wahrscheinlichkeitstheoretischer Überlegungen – die sog. *Vertrauensbereiche* oder *Konfidenzintervalle* ermittelt, d. h. jene Bereiche, in denen mit einer bestimmten Wahrscheinlichkeit der wahre Wert erwartet wird. Auf diese Weise läßt sich die Interpretation der durch Tests erfaßten Schülerleistungen auf eine (meßtheoretisch) gesicherte Grundlage stellen – ein Vorteil gegenüber anderen Formen der Schülerbeurteilung (z. B. Zensurierung), der sowohl dem interindividuellen Vergleich als auch intraindividuellen Untersuchungszielen (z. B. der Erfassung von Lernleistungsfortschritten oder sog. Profilanalysen) zugute kommt.

Die *Kritik der klassischen Testtheorie* setzt an der Axiomatik an, d. h. an

Widersprüchen zwischen einigen testtheoretischen Voraussetzungen und empirischen Befunden dazu. Die wichtigsten Annahmen der klassischen Testtheorie lassen sich in folgenden drei Axiomen zusammenfassen. 1) *Existenzaxiom*: Zu jedem beobachteten (gemessenen) Wert existiert ein «wahrer» Wert im Sinne einer bestimmten individuellen Merkmalsausprägung (z. B. Höhe der Schulleistung). Diese wird als Konstante – wenigstens über einen gewissen Zeitraum hinweg – angenommen. 2) *Fehleraxiom*: Der Meßfehler einer Messung ist eine Zufallsvariable. Für diese gilt, daß die Summe bzw. das arithmetische Mittel der Fehlerwerte den Wert Null ergibt. 3) *Verknüpfungaxiom*: Der beobachtete Wert (Meßwert) setzt sich additiv aus wahren Wert und Fehlerwert zusammen. Daraus kann eine Reihe von Hilfssätzen abgeleitet werden, worauf wir nicht mehr eingehen wollen.

Die auf der Grundlage der klassischen Meßfehlertheorie konstruierten Schulleistungstests lassen vor allem folgende Phänomene ungeklärt: erstens die Beobachtung, daß – entgegen obiger Annahme – die Fehlerwerte nicht unabhängig von den wahren Werten auftreten, d. h. «daß unterschiedliche wahre Werte auch unterschiedliche Fehlerwerte bedingen» können und diese nicht allein von der (mangelnden) Reliabilität des Tests, sondern auch von der jeweiligen Untersuchungspopulation beeinflußt werden; zweitens die (prinzipielle) Variabilität der wahren Werte, also die Tatsache, daß Merkmalschwankungen (z. B. Intelligenz- oder Schulleistungszuwachs durch gezielte Fördermaßnahmen) auftreten können; drittens die damit verbundene Schwierigkeit, meßtheoretischen Anforderungen (vor allem der Reliabilität und Validität) und berechtigten schulpädagogischen Anliegen (Lernleistungsförderung) gleichermaßen zu genügen. Inwieweit diese Probleme durch neuere testtheoretische Ansätze einer Lösung zugeführt werden können, wird in den folgenden Beiträgen behandelt.

Nach terminologischen Vorklärungen befaßt sich BÜSCHER (Kap. 3.2) mit der durch das RASCH-Modell eingeleiteten neuen Entwicklung einer *psychologischen* Meßtheorie, die vor allem am Problem der Populationsabhängigkeit klassischer Leistungsmessung ansetzt. «Für gewöhnlich werden die Eigenschaften eines psychologischen Tests durch individuelle Unterschiede innerhalb einer bestimmten Population definiert, und die Beurteilung jeder einzelnen Person ist mit der Referenzpopulation verknüpft. In vielen Fällen ist man jedoch daran interessiert, Individuen *per se* zu vergleichen, ohne sich dabei auf eine Population beziehen zu müssen. Die Frage etwa, ob sich ein Kind während eines oder mehrerer Jahre verbessert hat, kann unmöglich mit Hilfe verschiedener Tests beantwortet werden, die an Populationen unterschiedlich alter Kinder geeicht wurden. Im Rahmen eines neuen testtheoretischen Ansatzes versuchte RASCH (1960) derartige Probleme zu lösen, um das Studium einzelner Personen oder einzelner Testitems (Testaufgaben) unabhängig von Referenzpopulationen zu ermöglichen» (STENE 1968, S. 229).

Am Beispiel *kriteriumsbezogener* Leistungstests behandelt nun BÜSCHER einschlägige Probleme. Ausführlich werden hierbei neue Verfahrensansätze

zur Test- bzw. Itemanalyse erörtert. Deren praktische Relevanz im Hinblick auf die Schülerbeurteilung ist jedoch – vorerst noch – ungeklärt. Symptomatisch hierfür ist die abschließende Bewertung des RASCH-Modells durch BÜSCHER. «Die praktische Anwendung des RASCH-Modells auf lehrzielorientierte Tests ist problematisch, da zum einen eine große Anzahl von Probanden erforderlich ist, zum andern der Konstruktions- und Rechenaufwand enorm ist und schließlich das Modell nicht mehr angewandt werden kann, wenn alle Probanden das Lehrziel erreicht (bzw. nicht erreicht) haben. Das bedeutet, wenn überhaupt, dann ist das RASCH-Modell nur für normbezogene Messung sinnvoll.»

Damit, so könnte es scheinen, sind wir wieder an den Anfang unserer Problem Diskussion verwiesen. Dem ist nicht so. Vielmehr stehen wir mitten in einer Umorientierungsphase, in der neue Denkmodelle entwickelt und die klassische Testtheorie relativiert (nicht verworfen) werden. Soviel steht jetzt schon fest, daß kriteriumsbezogene Messungen genauso wie normbezogene «ihre eigene bedeutsame pädagogische Funktion» haben. «Keineswegs wird das eine Meßverfahren das andere verdrängen.»

Die einzelnen Methoden der Schulleistungsdiagnostik werden gewöhnlich – so auch hier – in *objektive* oder i. e. S. *messende Verfahren* (Tests) und in *subjektive Verfahren* (z. B. Lehrerurteile) klassifiziert. Diese Einteilung bedeutet freilich nicht, daß Schulleistungstests immer ausreichend objektive Ergebnisse lieferten und Lehrerurteile völlig inobjektiv wären. Vielmehr erwartet man von den «*objektiven*» Verfahren, daß ihre instrumentellen Eigenschaften empirisch überprüft wurden und die betr. Kennwerte den Testgütekriterien (Objektivität, Reliabilität, Validität) hinreichend entsprechen. Ferner sind hierzu genaue Auswertungsrichtlinien (Normen oder Bestehenskriterium) erforderlich. Demgegenüber ist hinsichtlich der Meßeigenschaften sog. subjektiver Prüfverfahren in concreto gewöhnlich nichts bekannt; soweit Verfahrenskontrollen (im Rahmen der Forschung) durchgeführt wurden, fielen die Gütekriterien im allgemeinen deutlich niedriger aus. Daraus darf man jedoch keine voreiligen Schlüsse ziehen, etwa auf die sog. subjektiven Beurteilungsmethoden zu verzichten. Zur Wahrnehmung der eingangs skizzierten Aufgaben der Schulleistungsdiagnostik sind beide Methodenansätze unverzichtbar und können sich sinnvoll ergänzen.

So unterschiedlich wie die testtheoretischen Ansätze ist die in der Literatur verwendete Terminologie im Bereich der Schulleistungsmessung. Während sich die konventionellen Bezeichnungen vorwiegend an meßtheoretischen Kriterien orientieren, begründete ROSEMAN (1974, S. 163 ff.) ein Klassifikationskonzept auf den *Funktionen* schulischer Leistungsbeurteilung. Dabei ist die Unterscheidung von «Leistungsfeststellung» und «Leistungsbewertung» von Bedeutung. «Im ersten Falle verschafft man sich lediglich Information darüber, was die Schüler im Verlaufe des Unterrichtsgeschehens gelernt bzw. nicht gelernt haben. Diese Informationen per se kann der Lehrer in vielfältiger

Weise verwenden. Im zweiten Falle geht man einen Schritt weiter, man will die festgestellten Leistungen der Schüler bewerten, wobei verschiedene Bezugspunkte für die Bewertung in Betracht kommen können.» Für Testverfahren, die im Dienste der erstgenannten Funktionseinheit stehen, schlägt der Autor die Sammelbezeichnung «Lernsteuerungstests» vor, für Verfahren der zweiten Kategorie die Bezeichnung «Lernkontrolltests». Diesem Begriffspaar entspricht die auf SCRIVEN (1967) zurückgehende Einteilung in *formative* Evaluation (Lernsteuerungstests) und *summative* Evaluation (Lernkontrolltests). Ordnungskriterium hierfür ist somit die Verwendungsart von Leistungsdiagnosen (vgl. auch BLOOM et al. 1971).

Was den Akt der Leistungsdiagnose selbst angeht, so gliedert sich dieser in zwei mehr oder weniger abhängige Phasen: Der erste Schritt ist die *Leistungsfeststellung*, z. B. die Ermittlung der richtigen vs. falschen Lösungen von Prüfungsaufgaben. Der zweite Schritt ist die *Leistungsbewertung*, die eigentliche «Beurteilung», die entsprechende Vergleichsmaßstäbe erfordert. Dieser Abschnitt ist zugleich der problematischere Teil, worauf bereits KANT (1781) hingewiesen hat – und «eine Unterweisung (Übung) der Urtheilskraft» (1877, S. 139f.) forderte! Prinzipiell lassen sich in diesem Zusammenhang drei *Beurteilungsmaßstäbe* unterscheiden:

- (1) *Individueller oder personbezogener Maßstab*. Einschlägig sind hier *intra-individuelle* oder *ipsative* Vergleiche, z. B. der jetzigen Leistung eines Schülers mit seiner (eigenen) früheren.
- (2) *Sozialer oder gruppenbezogener Normmaßstab*. Einschlägig ist hier der *interindividuelle* Leistungsvergleich, d. h. der individuelle Vergleich mit der Durchschnittsleistung (sog. Norm) einer Bezugsgruppe (z. B. Schulklasse, Leistungskurs oder Altersjahrgang). Dieses Bezugssystem spielt bei allen Formen *normorientierter* Leistungsprüfung eine Rolle.
- (3) *Sachbezogener oder kriterialer Maßstab*. Dieses Bezugssystem wird zur Auswertung lehr- oder lernzielbezogener (kriteriumsorientierter) Leistungsprüfungen, etwa bei informellen *kriteriumsorientierten* Tests, herangezogen. Damit läßt sich feststellen, ob ein Schüler das Lernziel (Kriterium) erreicht hat oder wieviel Prozent der Schüler einer Klasse das Unterrichtsziel erreicht vs. verfehlt haben («Pass-fail»-Kriterium).

Während die bisherigen Methodenkonzepte überwiegend *statusdiagnostische* Ansätze (sensu PAWLIK) darstellen, werden in den beiden folgenden Beiträgen neuere Modelle zur *Prozeßdiagnostik* behandelt. Im Sinne unserer Begriffsexplikation (Kap. 1.2) sind der statische wie der dynamische Aspekt in der Leistungsdiagnostik gleichermaßen von Bedeutung, wenngleich bezüglich prozeßdiagnostischer Ansätze noch eklatante Defizite bestehen.

Zunächst geht KORMANN (Kap. 3.3) auf allgemeine testtheoretische und methodische Probleme der *Veränderungsmessung* ein, die er herkömmlichen Schulleistungstests gegenüberstellt. Sodann werden alternative Methoden der Lernmessung (z. B. wiederholtes Testen mit Kontingenztafeln auf der Basis der

Konfigurationsfrequenzanalyse, clusteranalytische Verfahren, Latent-trait-Modelle) behandelt und praktische Konsequenzen für die Leistungsdiagnostik aufgezeigt. In Anlehnung an WYGOTSKIS Konzept zur «Pädagogisierung der Psychodiagnostik» wurden bereits Anfang der 70er Jahre am Leipziger Psychologischen Institut von GUTHKE und Mitarbeitern erste *Lerntestmethoden* (nach dem Design «Vortest – Pädagogisierungsphase – Nachtest») entwickelt. Inzwischen sind die Arbeiten hierzu auf internationaler Ebene aufgegriffen und verschiedene Lerntests publiziert worden.

Zwei Verfahren dieses neuen Testtyps werden von KORMANN in Kap. 4.4 vorgestellt: der *Mengenfolgentest für Schulanfänger* von GUTHKE (1983) und der *lerndiagnostische Rechtschreibtest* von SCHNOTZ (1979, 1983). Sicherlich befindet sich die Entwicklung schulleistungsspezifischer Lerntests sowohl in theoretischer als auch in methodischer Hinsicht noch im Anfangs- oder (besser) Aufbaustadium, nachdem erst seit wenigen Jahren empirische Forschungsbefunde hierzu vorliegen. Gleichzeitig sind jedoch damit neue, z. T. unerwartete Probleme aufgetaucht, deren Klärung vermutlich noch viele Jahre beanspruchen wird. Entsprechend konstatiert KORMANN: «Beim gegenwärtigen Forschungsstand von einer «Renaissance der pädagogischen Diagnostik» sprechen zu wollen, dürfte . . . verfrüht sein.»

Eine weitere Alternative zur traditionellen Schulleistungsmessung stellt die sog. strukturorientierte Diagnostik oder Anforderungsanalyse (task analysis) dar (Kap. 3.4). Dieser «task analysis approach» wurde im angloamerikanischen Raum, vorab im Rehabilitations- bzw. sonderpädagogischen Bereich, entwickelt. Danach werden Lernschwierigkeiten «mit fehlenden, d. h. noch nicht erworbenen Teilqualifikationen, die zum Bewältigen bestimmter (Lern-)Anforderungen notwendig sind», erklärt. Dieser Ansatz berücksichtigt bekannte Forderungen der neueren Lernpsychologie (z. B. GAGNÉ 1969/1973, WYGOTSKI 1964/1977); vgl. noch GUTHKE 1977 und 1982).

Bei der *Anforderungsanalyse* wird zwischen der «objektiven» und der «subjektiven» Leistungsstruktur unterschieden; die methodische Umsetzung demonstrieren EBERLE & KORNMAN am Beispiel der *Fehleranalyse* (vgl. auch Kap. 6.2). Damit erhofft man sich wichtige Aufschlüsse über die Ursachen individueller Lernschwierigkeiten sowie konkrete Hinweise auf notwendige pädagogische Förderungsmaßnahmen. Theoretische Grundlage hierfür ist die Annahme, daß schulische Lernprozesse in hierarchischer Lernschrittfolge ablaufen, denen entsprechende Diagnoseeinheiten zugeordnet werden. So entwickelte SCHNOTZ (1979, 1983) eine strukturorientierte diagnostische Konzeption in bezug auf Rechtschreib- und kognitive Verstehensleistungen, während PROBST (1981) sowie KUTZER & PROBST (1983) Anforderungen bezüglich der Zahl- und Oberbegriffsbildung, von Rechenoperationen und Leseleistungen strukturierten und technologische Hilfen dazu ausarbeiteten. Von der Kenntnis darüber, «in welcher Abfolge welche Teilqualifikationen erworben werden und wie diese aufeinander zu beziehen sind», erhofft man sich direktere (und schnellere) Hilfe, als dies durch die Analyse kognitiver Lernfähigkeiten mög-

lich ist. Inwieweit diese Erwartungen tatsächlich erfüllt werden können, bleibt künftigen Erfahrungen vorbehalten. Immerhin bietet die anforderungsorientierte Konzeption der Schulleistungsdiagnostik schon jetzt eine Reihe interessanter Anknüpfungspunkte zur aktuellen Lehr-Lernforschung (vgl. TREIBER & WEINERT 1982).

Ausgehend von der Annahme, daß Schulleistungen nicht ausschließlich Resultat des Schülerengagements sind, sondern immer auch von Bedingungsfaktoren der familialen und schulischen Lernumwelt – besonders der Lehrerpersönlichkeit bzw. -verhaltensstile – beeinflusst werden, fällt der *Unterrichtsbeurteilung* eine wichtige (im Kontext der Schulleistungsdiagnostik häufig vernachlässigte) Funktion zu. SCHRECKENBERG (Kap. 3.5) reflektiert zunächst den Unterrichtsbegriff unter der handlungstheoretischen, der sozialwissenschaftlichen und der pädagogisch-didaktischen Perspektive. Für die Unterrichtsbeurteilung ergibt sich daraus die Forderung, Unterrichtsleistungen hinsichtlich folgender «Gütekriterien» zu erfassen und zu bewerten: dem diagnostischen, dem Planungs- und dem Realisationsaspekt. Das vom Autor konzipierte Modell der Unterrichtsbeurteilung enthält die Indikatoren Zieldimension, Anspruchsniveau, methodische Entscheidungen, Strukturierung des Unterrichts und Unterrichtsarrangement. *Unterricht* wird dabei als «ein im wesentlichen durch Lehrerhandeln bestimmter Prozeß» verstanden, der auf den Schüler ausgerichtet ist, aber auch von ihm beeinflusst wird. Für die Gewichtung der einzelnen Beurteilungskriterien kann es nach SCHRECKENBERG keine starre Regel oder eindeutige Priorität geben. Vielmehr muß der Gesamtzusammenhang von Unterricht als komplexe Einheit gesehen werden, wofür das «sinnstiftende Zentrum» des Unterrichts ein Leitprinzip sein könnte. Merkmalszuordnung und -gewichtung sind somit bei der Beurteilung von Unterrichtsleistungen flexibel zu handhaben.

#### 1.4.2 *Objektive Verfahren der Leistungsmessung*

Eine unerläßliche Vorbedingung für objektive Leistungsdiagnosen sind eindeutige *Lernzieldefinitionen*. Diese müssen mit den Unterrichts- oder Lehrzielen übereinstimmen und sind möglichst operational zu definieren. *Operationalisierung von Lernzielen* meint hier (nach HORN in Kap. 4.1) die genaue Festlegung der vom Schüler am Ende einer Unterrichtseinheit geforderten Verhaltensweisen (Operationen). «Diese Verhaltensweisen müssen direkt beobachtbar sein. Daher findet man bei Lernzielen häufig Formulierungen wie «Der Schüler soll . . . nennen (aufschreiben, lösen usw.) können».»

Bei der Festlegung von Lernzielen orientiert man sich heute vielfach an der BLOOMschen Taxonomie, die 6 verschiedene, hierarchisch geordnete Komplexitätsstufen enthält: Wissen, Verstehen, Anwendung, Analyse, Synthese, Evaluation. Die Kategorien 2 bis 6 betreffen dabei mehr «intellektuelle Fähigkeiten» und repräsentieren Formen zunehmender Komplexität des kognitiven

Verhaltens, die auch als «Strategien des Problemlösens» aufgefaßt werden können. Für die praktische Anwendung des BLOOMschen Klassifikationsmodells bei der Unterrichtsplanung gibt HORN detaillierte Hinweise und erläutert diese an einem Beispiel aus dem Naturkundeunterricht, wobei die Darstellung seines eigenen Lehrplan-Analyseschemas besondere Beachtung verdient. Ausführlich wird dann auf die Konstruktion von Prüfungsaufgaben eingegangen. Auch hierzu bietet der Autor eine Reihe praktischer Beispiele. Insgesamt wird deutlich, daß die Messung schulischer Lernleistungen vor allen Methodenfragen zunächst von exakten Lernzieldefinitionen abhängt.

Im nächsten Beitrag (Kap. 4.2) gibt HORN einen Überblick über die wichtigsten zur Zeit verfügbaren standardisierten Schulleistungstests und deren Anwendungsmöglichkeiten im Rahmen der Schülerbeurteilung. Unter *inhaltlichem* Gesichtspunkt lassen sich diese folgendermaßen gruppieren: 1) *fachübergreifende* Tests oder sog. Omnibusverfahren (z.B. Allgemeiner Schulleistungstest für 2. Klassen AST 2 bzw. für 3. Klassen AST 3); 2) *fachspezifische* Tests (z.B. Fremdsprachentests für Englisch und Französisch); 3) *lernbereichsorientierte* Tests, wie Lesetests (z.B. Lesetest LT 2), Rechtschreibtests (z.B. Diagnostischer Rechtschreibtest DRT 3), Mathematiktests, z.B. Strukturorientierter Rechentest von PROBST & KUTZER (1984). Im Hinblick auf den *Anwendungsbereich* bzw. die Zielgruppe oder Untersuchungs*population* könnte man schließlich Tests für verschiedene Schulstufen und Schularten unterscheiden: 1) Tests für den *Primar-* vs. *Sekundarstufenbereich* bzw. die einzelnen *Schulklassen*; 2) Tests für die *Schultypen* des Gymnasiums, der Realschule und Hauptschule vs. Grundschule – seltener Gesamtschulen, die fast ausschließlich informelle Tests zur Lernleistungskontrolle verwenden (vgl. GAUDE & TESCHNER 1970).

Hauptanliegen des folgenden Beitrags von ROSEMANN (Kap. 4.3) ist die Information über *Lernkontrolltests*, aufgezeigt am Beispiel informeller Leistungsmessung. Dabei soll der Leser mit den Konstruktionstechniken und einigen Problemen informeller Tests soweit vertraut gemacht werden, daß er in der Lage ist, «den einen oder anderen Test selbst zu entwickeln bzw. bestehende Tests kritisch zu beurteilen».

Ein den Ausführungen vorangestelltes Ablaufdiagramm vermittelt einen schnellen Überblick über die einzelnen Arbeitsschritte. Nach der *Operationalisierung der Lernziele* (siehe oben) und der Erstellung einer *Spezifikations-tabelle* (in der nicht nur die Anzahl der Testaufgaben oder Items festgelegt, sondern auch eine Entscheidung darüber getroffen wird, welche Lernziele überprüft und somit in den Test aufgenommen werden sollen) erfolgt die *eigentliche Aufgabenkonstruktion*. Hierzu werden vom Autor alle relevanten Item- und Antworttypen besprochen und an Hand einschlägiger Beispiele das Vorgehen veranschaulicht. Ausführlich erörtert ROSEMANN dann die einzelnen Schritte zur Entwicklung der sog. *Testvorform* (Aufgabengruppierung, Erstellung des Testaufgabenheftes einschließlich der Instruktion, Festlegung des Bewertungsschlüssels usw.) sowie die Vorbereitung und Durchführung der

*Itemanalyse* (Schwierigkeits- und Trennschärfenanalyse jeder einzelnen Testaufgabe sowie Distraktorenanalyse, Aufgabenselektion bzw. -revision). Schließlich werden die Methoden zur Überprüfung der *Reliabilität* und *Validität* sowie Verfahren zur *Normierung* des Tests dargestellt, wobei eine Reihe von Illustrationsbeispielen wiederum die praktische Arbeit der Testkonstruktion bzw. -auswertung und das Verständnis testtheoretischer Grundlagen – nunmehr am konkreten Objekt – wesentlich erleichtern dürfte. Der Weg von der Lernzieldefinition bis hin zur *Testendform* ist somit lückenlos beschrieben.

### 1.4.3 *Subjektive Verfahren der Leistungsbeurteilung*

Zu den «*subjektiven*» *Verfahren* rechnen wir alle nicht-messenden Methoden der Schüler- und Lehrerbeurteilung, also Verfahrensansätze, die im Sinne der Meß- bzw. Testgütekriterien im allgemeinen weniger objektiv sind und einen geringeren Grad an Zuverlässigkeit und Gültigkeit aufweisen als Testmethoden. Im einzelnen fallen hierunter sämtliche Formen konventioneller Notengebung, mündliche Prüfung, schriftliche Prüfungen wie Klassenarbeiten oder Schulaufgaben, Aufsatzbeurteilung u.ä. Trotz der schlechteren instrumentellen Eigenschaften sog. subjektiver Beurteilungsverfahren weisen diese jedoch auch (andere) Vorzüge gegenüber sog. objektiven Verfahren auf, weshalb eine umfassende und differenzierte Schulleistungsdiagnostik darauf nicht verzichten kann. Beiden Methodenansätzen kommt vielmehr eine wichtige Ergänzungsfunktion in der Schülerbeurteilung zu.

Auch die in der Schulpraxis viel zu wenig beachtete Methode der (wissenschaftlichen) *Verhaltensbeobachtung* und eine Reihe von Beurteilungs- oder *Schätzverfahren* (*Rating*) sind der Kategorie der subjektiven Diagnoseinstrumente zuzuordnen. Damit befaßt sich der Beitrag von LANGHORST (Kap. 5.1).

Über die Notwendigkeit der Verhaltensbeobachtung bzw. einzelner Ratingtechniken im Unterricht sollte es eigentlich keine Diskussion mehr geben. Ohne den Einsatz dieser Methode(n) müßten wir sehr oft auf wichtige Informationen verzichten, z.B. über das Schüler- und Lehrerverhalten im Interaktionsspiel des Unterrichtsablaufs, damit zugleich auf nicht via Tests erfaßbare Daten bezüglich der Unterrichtsplanung, Curriculumentwicklung, Lernzielbestimmung oder in der Schullaufbahnberatung, bei Lern- und Leistungsschwierigkeiten wie auch bei Verhaltensstörungen und sozialen Konflikten.

Bei der Beobachtung und Beurteilung menschlichen Verhaltens ist allerdings mit spezifischen *Fehlerquellen* zu rechnen, deren Kenntnis *eine* Voraussetzung zu ihrer Beherrschung darstellt. Hiervon wiederum verspricht man sich objektivere (und gerechtere?) Schüler- und Lehrerbeurteilungen. Nach einer modellhaften Beschreibung der Eindrucksbildung bzw. Personwahrnehmung geht LANGHORST deshalb auf die wichtigsten Urteilsfehler ein. Sodann wird ein Überblick über systematische und nichtsystematische Beobachtungs-

formen geboten, wobei das Hauptaugenmerk der *Verhaltensbeschreibung* (Aufzeichnungsmethoden) gilt. Wissenschaftlich fundierte Verhaltensbeobachtungen sind durch drei – unterscheidbare – Phasen charakterisiert: 1) die planmäßige, absichtliche bzw. selektierende und gegenstandszentrierte *Wahrnehmung* (Beobachtung i. e. S.), 2) die *Beschreibung* (Protokollierung) des beobachteten Verhaltens, 3) die *Interpretation oder Bewertung* (Beurteilung i. e. S.) der Beobachtungsdaten im Hinblick auf eine systematische Ordnung bzw. Hypothesenentscheidung und gegebenenfalls Interventionsplanung.

Zum Hauptproblem wird hierbei vor allem die zweite Stufe im Prozeß der Verhaltensbeobachtung, die Verhaltensbeschreibung. Dabei sollten nur *Beschreibungskategorien* – auf dem verbalen oder adverbialen Niveau – verwendet werden, während Protokollsätze unter Verwendung adjektivischer oder gar substantivischer Wörter («faul», «fleißig», «unkonzentriert» oder «Faulpelz», «Störer» usw.) bereits Deutungen bzw. Wertungen – keine beschreibenden Kategorien mehr – darstellen. Tatsächlich enthalten viele der sog. Schülerbeobachtungsbögen solche Bewertungen, die jedoch Bestandteil der Interpretationsphase sein sollten. Hingegen wird bei *Schätzskalen oder Ratingverfahren* die Beschreibung des Verhaltens ausgespart zugunsten sofortiger Bewertung des beobachteten Leistungsverhaltens (Beispiel: Notenskala). Durch zahlreiche Beispiele illustriert LANGHORST die vielfältigen Möglichkeiten der Verhaltensbeobachtung und -beurteilung im Unterricht.

Die Möglichkeiten zur objektiven Leistungsbeurteilung sind trotz des wachsenden Methodenangebots begrenzt, und diese Grenzen können mitunter in der Spezifität des Gegenstandes selbst liegen. Als Paradebeispiel hierfür sei die Aufsatzbeurteilung genannt. Andererseits wird m.E. zu Recht darauf hingewiesen, daß man – aus den verschiedensten Gründen – auf die Übung des Aufsatzes (und seine Bewertung) nicht verzichten sollte. Ähnliche Argumente sprechen für Klassenarbeiten und die prinzipielle Beibehaltung mündlicher Prüfungen. In Abwägung der jeweiligen Positiva sowie jeweils spezifischer Methodenschwächen kommt den einzelnen Prüfverfahren in der Schulleistungsdiagnostik unter Berücksichtigung bestimmter Funktionsziele unterschiedliche Bedeutung zu. Die Auswahl der diagnostischen Informationsquellen darf somit nicht allein nach den Testgütekriterien erfolgen, sondern muß immer auch inhaltliche und pädagogische bzw. pädagogisch-psychologische Gesichtspunkte berücksichtigen.

Ursprünglich dienten *mündliche Prüfungen* (vgl. Kap. 5.2) an den mittelalterlichen Hochschulen zur Erfassung und Bewertung studentischer Leistungen in der Rhetorik. Erst später erfolgte eine Ausdehnung dieser Prüfungsform auf andere Unterrichtsfächer, wobei dann nicht mehr so sehr die Redegewandtheit, die Argumentationsfähigkeit bzw. die Fertigkeit zur wissenschaftlichen Disputation Gegenstand mündlicher Prüfung waren, sondern mehr oder weniger nichtsprachbezogene Fachleistungen (z. B. in den Naturwissenschaften). Diese Ausweitung ist jedoch nicht unproblematisch, da hierbei sehr leicht der Zusammenhang von Lehr-/Lernziel und angemessener Prüfung ver-

loren geht. Andererseits bietet die mündliche Prüfungsform gegenüber schriftlichen Prüfungen prinzipiell eine Reihe von Vorzügen, so etwa die Möglichkeit, ein breiteres Themenspektrum anzusprechen und/oder die höheren Anforderungsebenen in der BLOOMschen Taxonomie kognitiver Lernziele (Verstehen, Anwendung u.ä.) stärker zu berücksichtigen. Weitere Vorteile wären die relativ kurze Prüfungszeit (pro Kandidat), die individuelle Abstimmung der Aufgabenschwierigkeit und die schnelle Rückmeldung des Prüfungsergebnisses. Schließlich sind manche Leistungsaspekte in sprachlichen Fächern (vorab im Fremdsprachenunterricht), etwa Artikulations- und Intonationsfähigkeit, nur mündlich zu erfassen.

Demgegenüber werden in der Literatur zahlreiche Kritikpunkte genannt, von denen – im Vergleich zur schriftlichen Prüfungsform – die durchweg schlechteren Meßeigenschaften (d.h. geringere Objektivität, Reliabilität und Validität) besonders hervorstechen. BIRKEL referiert kritische Einwände jedoch nicht nur in meßtheoretischer Hinsicht, sondern auch aus sozial- und tiefenpsychologischer Perspektive. Darüber hinaus werden unterrichtsspezifische Probleme angesprochen, deren Merkmalsvielfalt ein negatives Pauschalurteil über die mündlichen Prüfungen kaum aufrechterhalten läßt. Der abschließende Methodenvergleich ermöglicht nicht nur eine rasche Orientierung über Vorzüge und Schwachstellen mündlicher vs. schriftlicher Prüfungen, sondern wird auch für den konkreten Anwendungsfall die Wahl der jeweils optimalen Informationsquelle zur Leistungskontrolle erleichtern.

In seinem zweiten Beitrag (Kap. 5.3) thematisiert BIRKEL die traditionelle Form schriftlicher Leistungsprüfungen: *Klassenarbeiten* oder *Schulaufgaben*. Dabei werden – trotz insgesamt besserer instrumenteller Eigenschaften der schriftlichen Prüfungen – zahlreiche Berührungspunkte bzw. Überschneidungen mit der mündlichen Prüfungsform sichtbar. Der Autor plädiert deshalb für eine differenzierte Beantwortung der Methodenfrage, d.h. eine lehrzielbezogene Auswahl des Diagnoseinstrumentes, wobei prinzipiell schriftliche und mündliche Prüfungsformen ihren Platz haben. Dies gilt auch für entsprechende Methodenvarianten wie Zettelarbeit, Referat, Hausarbeit u.ä. Darüber hinaus läßt sich die Qualität der Leistungsdiagnose durch den Einsatz informeller und/oder formeller Tests (vgl. Kap.4 sowie 3.3 und 3.4) in vielen Fällen steigern.

Der nächste Beitrag von SCHÜMANN & WIECZERKOWSKI (Kap. 5.4) ist dem Problem der *Aufsatzbeurteilung* gewidmet. Nirgendwo sonst zeigt sich die Methodenproblematik (auch) der schriftlichen Prüfung so deutlich wie bei der Erfassung und Bewertung von Aufsätzen und aufsatzähnlichen Schülerarbeiten. In den bisherigen Bemühungen um größere Vergleichbarkeit der Bewertungsstandards werden zwei – sich ergänzende – Ansätze erkennbar: «1) die Prozesse bei der Beurteilung von Aufsätzen, insbesondere aber die Gewichte verschiedener in die Bewertung einfließender Parameter, zu analysieren und 2) die Benotung schriftlicher Produkte durch die Entwicklung von Normen und Auswertungsrichtlinien auf eine objektivere Basis zu stellen.» Die Auto-

ren sehen hierin einen nützlichen Beitrag, aber noch keine hinreichenden Voraussetzungen für «gerechtere Notengebung». Vor allem bleibt fraglich, ob sich dadurch die Komplexität des Gegenstandes einfangen läßt. Berücksichtigung finden müßten neben Ziel und Funktion des Schüleraufsatzes auf verschiedenen Altersstufen noch «die jeweiligen Schüler-Objekt-Beziehungen in den Aufgabenstellungen und die Objekt-Beurteiler-Relationen, von denen einerseits die Art der Aufgabenerledigung abhängt und andererseits wesentliche Einflüsse auf die wertende Einordnung eines Produkts in den Erwartungshorizont des Beurteilers ausgehen».

SCHÜMANN & WIECZERKOWSKI entwerfen deshalb ein allgemeines Modell der Aufsatzbeurteilung, das nicht ausschließlich an den Kriterien der Meß- bzw. Fehlertheorie orientiert ist. Demnach könnte man Aufsätze als «sprachliches Handeln» und entsprechende Niederschriften als Produkte eines «schriftsprachlichen Gestaltungsvorgangs» betrachten. Die Güte des Aufsatzprodukts ist nach dem Funktionsgefüge-Modell von BECK (1974) von vier Hauptkomponenten und einer größeren Variablenzahl abhängig, die miteinander interagieren und die Komplexität von Aufsätzen bestimmen. Daraus lassen sich wiederum negative vs. positive Merkmale der Aufsatzgüte gewinnen, die als Beurteilungskriterien dienen können. Zur ersten Gruppe gehören z.B. Fehlleistungen im lexikalisch-semantischen Bereich oder im morphologisch-syntaktischen Bereich, zur zweiten Gruppe sprachliche Gestaltung, Inhalt, Gesamtaufbau usw. Die Autoren informieren nicht nur über die in der Literatur vorgeschlagenen Kriterien(kataloge) und Normen zur Aufsatzbeurteilung, sondern referieren auch wichtige empirische Forschungsbefunde zur behandelten Thematik. Abschließend wird der Stellenwert von Aufsatznoten im Kontext der Schülerbeurteilung und Schulleistungsanalyse diskutiert.

Neben der Schülerbeurteilung muß in einem umfassenden Schulleistungskonzept die *Lehrerbeurteilung* – unter dem Aspekt der Unterrichtsleistung – Berücksichtigung finden. SCHRECKENBERG stellt hierzu (in Kap. 5.5) praktische Ansätze vor. Zunächst werden aus der Funktionsbestimmung der Unterrichtsplanung Gütekriterien abgeleitet bzw. Indikatoren für die Beurteilung des Planungsbereichs gewonnen. Hierunter fallen z.B. die Diagnose der Ausgangslage (d.h. des sachstrukturellen Entwicklungsstandes sensu HECKHAUSEN), Intentionen und Ziele des Unterrichts (fachwissenschaftlicher und didaktischer Aspekt), methodische Entscheidungen und ihre Begründungen (Lehrstrategien), Strukturierung des Unterrichts (hinsichtlich lernpsychologischer Angemessenheit, Lehrziel und Zeitdimension). Nach Meinung des Autors sind hierzu nicht nur objektive Aussagen möglich, sondern ist auch das Gewichtungproblem lösbar: «Priorität hat das Verhältnis der in der Diagnose dargestellten Ausgangslage zur Zielsetzung.»

Sodann werden Beobachtungssysteme und Gütekriterien für die Erfassung und Bewertung der *Realisation* von Unterricht dargestellt. Wichtige Beurteilungsdimensionen sind dabei das Lehrer- (und Schüler-)Verhalten, Unterrichtsarrangement und «Anspruchsniveau» (d.h. Leistungsforderung des Un-

terrichts an den Schüler). Das Gewichtungproblem möchte SCHRECKENBERG hier – wenigstens für manche Unterrichtsstunden – von einem «zentralen Beurteilungsaspekt» her lösen. Dabei muß dieser jedoch «sehr sorgfältig im Hinblick auf seine Relevanz für das Gesamtgefüge der Unterrichtsbeurteilung geprüft und analysiert werden» (vgl. auch Kap. 3.5).

## 1.5 Schülerbeurteilung und Bildungsberatung

Schließlich werden anhand ausgewählter Praxisfelder Anwendungsprobleme der pädagogisch-psychologischen Diagnostik im Kontext der Schule behandelt. Diagnostische Untersuchungen dienen ja keinem Selbstzweck, sondern übernehmen auxiliäre Funktionen. Diese sollen hier am Beispiel der Schulberatung aufgezeigt werden. Zugleich ist damit wohl der wichtigste – wenngleich nicht der häufigste – Anlaß für Schulleistungsdiagnosen angesprochen.

Vor dem Hintergrund eines Paradigmawechsels entwicklungspsychologischer Auffassungen über die «Schulreife» versteht NICKEL (Kap. 6.1) diesen Begriff als ökopsychologisches Konstrukt. Sein Schulreifemodell basiert auf einem interaktionistischen ökopsychologischen Ansatz und beinhaltet drei Faktoren (Schule, Schüler, Ökologie) mit jeweils einzelnen Teilkomponenten. Entsprechend fordert NICKEL eine revidierte *Einschulungsdiagnostik* als «ökopsychologisch orientierte Interventionsstrategie», die «Schulreife» oder Schulfähigkeit als «dynamisches Produkt aktiver Wechselbeziehungen zwischen Individuum, familiärem System, Schule und Gesellschaft» begreift.

WENDELER (vgl. Kap. 6.2) erläutert zunächst kurz die Hauptfunktionen pädagogisch-psychologischer Diagnostik: die Auslese- und die Förderungsfunktion. Seiner Meinung nach sind beide notwendig, wenngleich er nur die Förderungsdagnostik ausführt. In Anlehnung an MERCER (1979) wird zwischen Fähigkeits- und Lernzieldiagnostik unterschieden. Während das *Fähigkeitsmodell* basale Dispositionen (als Ursachen für Schulerfolg vs. Schulversagen bzw. bestimmte Schulleistungsschwächen) unterstellt, geht man im Sinne des *Lernzielmodells* davon aus, daß nicht hypothetische Konstrukte (basale Fähigkeiten), sondern schwach entwickelte Fertigkeiten bzw. (Vor-)Kenntnislücken die aktuellen Schulleistungsschwierigkeiten bedingen. WENDELER meint, daß im Rahmen förderungsdagnostischer Ziele die *Lernzieldiagnostik* das adäquatere Modell darstellt und demonstriert dessen Vorzüge anhand verschiedener Lernbereiche sowie einzelner Unterrichtsfächer (Lesen und Rechtschreiben, konkreter mathematischer Operationen u.ä.); hierbei wird der *Fehleranalyse* besondere Bedeutung zugemessen.

CHRISTIANI & BARTNITZKY (Kap. 6.3) stellen die *Individualisierung* des Unterrichts und die *soziale Koedukation* als Kernaufgaben der Grundschule heraus. Leistungsbeurteilungen hätten demnach nur dann einen Sinn, wenn die individuelle Bezugsnorm oder die anforderungsbezogene (kriteriale) Norm zur Anwendung gelangen (vgl. auch 1.4.1 oben). Konsequenterweise fordern

die Autoren zur Diagnose individueller Lernfortschritte eine Beurteilungspraxis *ohne* Zensuren (die ja gewöhnlich normorientiert, also mit Hilfe des sozialen Bezugssystems, erstellt werden). Für die Lernerfolgskontrollen eignen sich tägliche Verhaltensbeobachtungen und Ratings (vgl. Kap. 5.1), Schülerbeobachtungsbögen u.ä. Schließlich geben CHRISTIANI & BARTNITZKY konkrete Arbeitsanweisungen zur Abfassung zensurenfreier Zeugnisse (in der Grundschule).

Während in den bisherigen Beiträgen förderungsdiagnostische Ziele im Vordergrund individueller Schulleistungsanalysen standen, rückt bei Fragestellungen der *Schullaufbahnberatung* die *Auslese-* oder (besser) die *Klassifikationsfunktion* in den Mittelpunkt des Interesses. Zunächst wird (in Kap. 6.4) ein Strukturmodell der Schullaufbahnberatung beschrieben, wobei neben der Bildungsweginformation die Schulleistungsdiagnose und die Begabungsdiagnose das Fundament für die Eignungsermittlung (Schulerfolgsprognose) bilden. Dieser liegt der *Diagnose-Prognose-Ansatz* zugrunde. Ausführlicher werden deshalb die Messung des Kriteriums (Schulerfolg), die Auswahl der Prädiktoren (Bedingungsvariablen der Schulleistung bzw. des Schulerfolgs) sowie relevante Entscheidungsstrategien und Prognosemodelle dargestellt. In diesem Zusammenhang interessiert auch ein Vergleich klinischer vs. statistischer Vorhersagen, wobei nach WIGGINS' (1973) sowie eigenen Untersuchungsergebnissen (ALLINGER & HELLER 1975, HELLER et al. 1978a/b) die Prognose durch Kombination objektiver und subjektiver Informationsquellen (statistischer und klinischer Prädiktoren) optimiert werden kann. Schließlich wird ein Modell zur sukzessiven Entscheidungsfindung in der Schullaufbahnberatung vorgeschlagen, dem u.a. Schulleistungsdaten zugrundeliegen und das in ein interaktives Beratungsgespräch einmündet.

Im letzten Beitrag (Kap. 6.5) behandelt KLINGEN Fragen und Methodenprobleme der Leistungsbeurteilung im *Kurswahlsystem* der Sekundarstufe. Leistungsdiagnosen und Neigungsdifferenzierung kommt dabei schulform- bzw. schulstufenabhängig unterschiedliche Bedeutung zu. Eingehender werden nach einer kurzen Erläuterung des Systems der Grund- und Leistungskurse die verschiedenen Methoden und Verfahrensansätze zur Schulleistungsdiagnose im Kurswahlsystem beschrieben und den Problemen des Abiturbereichs besondere Aufmerksamkeit geschenkt. Diese Hinweise sind umso nützlicher, als zu dieser Thematik relativ wenige Publikationen vorliegen.

Die Ausführungen in diesem Buch sollten deutlich machen, daß zur Bewältigung der einzelnen Aufgaben und Funktionen schulischer Leistungsdiagnostik eine breite, differenzierte Palette von Informationsquellen erforderlich ist. Dabei wird man sowohl objektive (Tests) als auch subjektive («klinische») Verfahren berücksichtigen müssen, um einigermaßen zuverlässige und gültige, andererseits jedoch pädagogisch relevante Daten über den Schulleistungsstand einzelner oder auch von Lerngruppen zu gewinnen. Konkrete Anwendungsbeispiele dürften dabei die praktische Umsetzung diagnostischer Informationen im Kontext der Schule unterstützen.

## 7. Literaturverzeichnis

- AEBLI, H.: Die geistige Entwicklung als Funktion von Anlage, Reifung, Umwelt- und Erziehungsbedingungen. In: ROTH, H. (Hrsg.): *Begabung und Lernen*. Klett, Stuttgart 1969, 151 – 191.
- AEBLI, H.: *Grundformen des Lehrens*. Klett-Cotta, Stuttgart 1976.
- AHRENS, G.: *Wie zensieren wir? Lebendige Schule* 1964, 19, 368 – 381.
- ALLINGER, U., HELLER, K.: Automatische Klassifikation von psychologischen Untersuchungsbefunden. In: Kultusministerium Baden-Württemberg (Hrsg.): *Bildungsberatung in der Praxis*. Neckarverlag, Villingen 1975, 142 – 169.
- ANASTASI, A.: *Psychological Testing*. McMillan, London 1968<sup>3</sup>.
- ANGER, H., BARGMANN, R., VOIGT, M.: *Verständiges Lesen (VL 5 – 6)*. Beltz, Weinheim 1965.
- ANGERMAIER, M.: Psycholinguistischer Entwicklungstest (Deutsche Bearbeitung des «Illinois Test of Psycholinguistic-Abilities» von S. A. KIRK, J. H. MCCARTHY & W. D. KIRK). Beltz, Weinheim 1974.
- American Psychological Association (APA): *Standards for Educational and Psychological Tests*. APA, Washington, D. C. 1974.
- ARBINGER, R., SALTERN, M. v.: Schulische Umwelt und soziales Klima in Schulklassen. *Psychologie in Erziehung und Unterricht* 1984, 31, 81 – 100.
- ASCH, S. E.: Forming impressions of personality. *Journal of Abnormal and Social Psychology* 1946, 41, 258 – 290.
- ASCHERSLEBEN, K.: Untersuchungen zur Reliabilität von Schulnoten. *Schule und Psychologie* 1971, 18, 147 – 154.
- ASCHERSLEBEN, K.: Transformation und Normierung in Pädagogik und Psychologie. *Psychologie in Erziehung und Unterricht* 1973, 20, 77 – 88.
- ATKINSON, J. W.: *An introduction to motivation*. Van Nostrand Reinhold, New York 1964.
- ATKINSON, J. W.: Motivational determinants of intellectual performance and cumulative achievement. In: ATKINSON, J. W., RAYNOR, J. D. (Hrsg.): *Motivation and achievement*. Winston, Washington, D. C. 1974, 389 – 410 (a).
- ATKINSON, J. W.: Strength of motivation and efficiency of performance. In: ATKINSON, J. W., RAYNOR, J. D. (Hrsg.): *Motivation and achievement*. Winston, Washington, D. C. 1974, 193 – 218 (b).
- ATKINSON, J. W., LENS, W.: Fähigkeit und Motivation als Determinanten momentaner und kumulativer Leistung. In: HECKHAUSEN, H. (Hrsg.): *Fähigkeit und Motivation in erwartungswidriger Schulleistung*. Hogrefe, Göttingen 1980, 129 – 188.
- AUCKENTHALER, A.: Versuch einer psychologischen Analyse der mündlichen Prüfung. *Zeitschrift für experimentelle und angewandte Psychologie* 1975, 22, 391 – 408.
- AUGST, G.: Die linguistischen Grundlagen der Rechtschreibung. In: AUGST, G. (Hrsg.): *Deutsche Rechtschreibung mangelhaft? Quelle & Meyer, Heidelberg* 1974, 9 – 47.
- AURIN, K. et al.: *Gleiche Chancen im Bildungsgang (= Bd. 9 der Schriftenreihe des Kultusministeriums Baden-Württemberg zur Bildungsforschung, Bildungsplanung, Bildungspolitik)*. Neckarverlag, Villingen 1968.
- AMMER, C., BUGGLE, F., WETZEL, H., WILHELM, M.: *Veränderung von Schülerverhalten*. Urban & Schwarzenberg, München 1976.

- BALTES, P.B., NESSELROADE, J.: Die entwicklungspsychologische Analyse von individuellen Unterschieden in mehreren Meßgrößen. In: BALTES, P.B., ECKENBERGER, L. (Hrsg.): Entwicklungspsychologie der Lebensspanne. Klett, Stuttgart 1979, 145 – 178.
- BARTNITZKY, H.: Zensurenfreie Zeugnisse: Ja – aber wie? Die Grundschule 1981, 13, 222 – 223.
- BARTNITZKY, H., CHRISTIANI, R.: Zeugnisschreiben in der Grundschule. Hilfen zur Schülerbeurteilung ohne und mit Zensuren. Dieck, Heinsberg 1979.
- BAUR, R.: Elternhaus und Bildungschancen. Beltz, Weinheim 1972.
- BECK, O.: Kriterien zur Aufsatzbeurteilung. Hase & Köhler, Mainz 1974.
- BECK, O., HOFEN, N.: Sprach-Leistungstest für 4. Klassen. Beltz, Weinheim 1981.
- BECK, O., PAYRHUBER, F.J. (Hrsg.): Aufsatzbeurteilung heute. Herder, Freiburg 1978<sup>4</sup>.
- BECKMANN, H.-K. (Hrsg.): Leistung in der Schule. Westermann, Braunschweig 1978.
- BENNER, D., RAMSEGER, J.: Fallstudie zur Überprüfung der Neuregelung der Grundschulzeugnisse im ersten und zweiten Schuljahr im Hinblick auf die Resonanz bei Lehrern und Eltern unter besonderer Berücksichtigung der Region Münster. Der Kultusminister des Landes Nordrhein-Westfalen, Düsseldorf 1981.
- BEREITER, C.: Some Persisting Dilemmas in the Measurement of Change. In: HARRIS, C.W. (Hrsg.): Problems in measuring change. Madison Wisc. 1963, 3 – 20.
- BERK, R.A. (Hrsg.): Criterion-referenced Measurement: The State of the Art. Johns Hopkins, Baltimore 1980.
- BERLINER, D.C., CAHEN, L.S.: Trait-treatment-interaction and learning. In: KERLINGER, F.N. (Hrsg.): Review of Research in Education. Peacock, Itasca 1973, 58 – 94.
- BETHÄUSER, H., REICHENBECHER, H.: Approximative Expertenabstimmung in der Bildungsberatung. In: HELLER, K. (Hrsg.): Handbuch der Bildungsberatung, Bd.3. Klett, Stuttgart 1976, 861 – 877.
- BETZ, D.: Rhythmische Schwankungen als Fehler in der Notengebung bei mündlichen Prüfungen. Psychologie in Erziehung und Unterricht 1974, 2, 1 – 14.
- BIÄSCH, H.: Psychologische Hintergründe des Examensverhaltens. In: Jahrbuch der Psychologie, Psychotherapie und Medizinische Anthropologie 1968, 16, 115 – 129.
- BIBL, W., LÜHMANN, R.: Einige neue Entwicklungen in der Theorie kriteriumsorientierter Tests. Lernzielorientierter Unterricht 1981, 2, 28 – 39.
- BIGLMAIER, F.: Lesetests-Serie. Reinhardt, München 1969<sup>2</sup>.
- BIRKEL, P.: Mündliche Prüfungen. Kamp, Bochum 1978.
- BIRKEL, P.: Leistungsmessung durch Tests. Westermanns Pädagogische Beiträge 1978, 30, 258 – 263 (a).
- BIRKEL, P.: Schriftliche Prüfungen und Klassenarbeiten in Geographie und Biologie. In: KLAUER, K.J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd.3. Schwann, Düsseldorf 1978 (b), 683 – 688.
- BIRKEL, P., FRITZ, V.: Sprechflüssigkeit und Vorinformationen als validitätsmindernde Faktoren bei mündlichen Prüfungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1980, 12, 282 – 289.
- BLOCK, J.H.: Criterion-referenced Measurements: Potential. School Review 1971, 289 – 298.
- BLOOM, B.S.: Taxonomy of Educational Objectives. Handbook 1: Cognitive Domain. David McKay Company, New York 1956. – Dt. Taxonomie von Lernzielen im kognitiven Bereich. Beltz, Weinheim 1972, 1976<sup>5</sup>.
- BLOOM, B.S., HASTINGS, J. TH., MADUS, G.F.: Handbook on formative and summative evaluation of student learning. McGraw-Hill, New York 1971.
- BOCK, R.D.: Univariate and Multivariate Analysis of Variance of Time-Structured Data. In: NESSELROADE, J.R., BALTES, P.B. (Hrsg.): Longitudinal Research in the Study of Behavior and Development. Academic Press, New York 1979.

- BOLSCHO, D., SCHWARZER, C. (Hrsg.): Beurteilen in der Grundschule. Urban & Schwarzenberg, München 1979.
- Bonner Vereinbarung zur Neugestaltung der gymnasialen Oberstufe in der Sekundarstufe II. Luchterhand, Neuwied 1972.
- BORMUTH, J.R.: On the theory of achievement tests items. The University of Chicago Press, Chicago 1970.
- BRANDTSTÄDTER, J.: Soziale Schicht, Umwelt und Intelligenz: Eine Pfadanalyse der Korrelationsbefunde von MARJORIBANKS. Psychologische Beiträge 1976, 18, 35 – 53.
- BREUER, H., WEUFFEN, M.: Gut vorbereitet auf das Lesen- und Schreibenlernen? Deutscher Verlag der Wissenschaften, Berlin 1975, 1978.
- BRICKENKAMP, R. (Hrsg.): Handbuch psychologischer und pädagogischer Tests. Hogrefe, Göttingen 1975, 1983<sup>2</sup>.
- BROMBACH, E. et al.: Differenzierung nach Leistung? (Forschungsbericht Nr. 7) Max-Traeger-Stiftung, Berlin 1971.
- BRONFENBRENNER, U.: Ökologische Sozialisationsforschung. Klett, Stuttgart 1976.
- BRONFENBRENNER, U.: Die Ökologie der menschlichen Entwicklung. Klett, Stuttgart 1981.
- BROPHY, J.E., GOOD, T.L.: Die Lehrer-Schüler-Interaktion. Urban & Schwarzenberg, München 1976.
- BROWN, A.L., FRENCH, L.A.: The Zone of Potential Development: Implications for Intelligence Testing in the Year 2000. In: STERNBERGER, R.J., DETTERMAN, D.K. (Hrsg.): Human Intelligence. Allex Publishing Corp., Norwood 1979.
- BROWN, J.S., BURTON, R.R.: Diagnose models for procedural bugs in basic mathematical Skills. Cognitive Science 1978, 2, 155 – 192.
- BULL, G.M.: An Examination of the Final Examination in Medicine. The Lancet 1956, 368 – 372.
- CAMPBELL, D.T., STANLEY, J.C.: Experimental Designs for Research on Teaching. In: GAGE, N.L. (Hrsg.): Handbook on Teaching. Rand McNally, Chicago 1963.
- CARTER, R.S.: Non-intellectual variables involved in teacher's marks. Journal of Educational Research 1953, 44, 81 – 95.
- CARVER, R.P.: Special problems in measuring change with psychometric devices. Evaluative Research: Strategies and Methods. American Institutes for Research, Pittsburgh 1970.
- CERKES JULKOWSKI, M.: Fehleranalyse als eine Methode der Förderungsdiagnostik. In: KORNMAN, R., MEISTER, H., SCHLEE, J. (Hrsg.): Förderungsdiagnostik. Schindele, Heidelberg 1983, 78 – 83.
- CHIOUT, H., STEFFENS, W.: Unterrichtsvorbereitung und Unterrichtsbeurteilung. Diesterweg, Frankfurt/M. 1970, 1978<sup>4</sup>.
- CLAUSS, G., EBNER, H.: Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen. Deutscher, Frankfurt 1970.
- COHEN, R., REY, R.E., SIXTL, F.: Die Kombination von Häufigkeit und Intensität im diagnostischen Urteil. Psychologische Forschung 1969, 33, 9 – 20.
- COHEN, R., SCHÜMER, R.: Kontexteffekte bei der Eindrucksbildung. Zeitschrift für Sozialpsychologie 1972, 3, 313 – 328.
- COLEMAN, J.S. et al.: Equality of educational opportunity. U.S. Government Printing Office, Washington, D.C. 1966.
- COMBE, A.: Kritik der Lehrerrolle. List, München 1971.
- COMBER, L.C., KEEVES, J.P.: Science education in nineteen countries: An empirical study. Wiley, New York 1973.
- COOK, T.D., CAMPBELL, D.T.: Quasi-Experiments: Nonequivalent Control Groups Designs. In: COOK, T.D., CAMPBELL, D.T. (Hrsg.): Quasi-Experimentation: Design and Analysis Issues for Field Settings. Rand McNally, Chicago 1979, 95 – 146.

- COPEI, F.: Der fruchtbare Moment im Bildungsprozeß. Quelle & Meyer, 1959, Heidelberg 1963<sup>7</sup>.
- COX, R., GRAHAM, G. T.: The development of a sequentially scaled achievement test. *Journal of Educational Measurement* 1966, 3, 147 – 150.
- COX, R., VARGAS, J.S.: A comparison of item selection techniques for norm-referenced and criterion-referenced tests. Paper read at the Annual Meeting of the National Council on Measurement in Education, Chicago, Ill. 1966.
- CRAMBERT, A. C.: Estimation of Validity for Criterion-Referenced Tests. Paper presented at the Annual Meeting of the American Educational Research Association, New York 1977.
- CRANACH, V. M., KALBERMATTEN, V., INDERMÜHLE, K., GUGLER, B.: *Zielgerichtetes Handeln*. Huber, Bern 1980.
- CRANO, W. D., KENNY, D. A., CAMPBELL, D. T.: Does intelligence cause achievement?: A cross-lagged Panel analysis. *Journal of Educational Psychology* 1972, 63, 258 – 275.
- CRONBACH, L. J.: *Essentials of Psychological Testing*. Harper & Row, London 1970<sup>3</sup>.
- CRONBACH, L. J., MEEHL, P. E.: Construct Validity in Psychological Tests. *Psychological Bulletin* 1955, 52, 281 – 302.
- CRONBACH, L. J., RAJARATNAM, N., GLEESER, G. C.: Theory of generalizability: Aliberalization of reliability theory. *British Journal of Statistical Psychology* 1963, 16, 137 – 163.
- CRONBACH, L. J., GLEESER, G. C.: *Psychological Tests and Personnel Decisions*. University of Illinois Press, Urbana 1965<sup>2</sup>.
- CRONBACH, L. J., FURBY, L.: How we should measure «Change» – or should we? *Psychological Bulletin* 1970, 74, 68 – 80.
- CRONBACH, L. J., SNOW, R. E.: *Aptitudes and instructional methods*. Irvington, New York 1977.
- DAHL, T.: *Toward an Evaluative Methodology for Criterion-Referenced Measures: Objective-Item-Congruence*. CSE Report No. 15, Los Angeles 1971.
- DANN, H.-D. et al. (Hrsg.): *Analyse und Modifikation subjektiver Theorien von Lehrern. Ergebnisse und Perspektiven eines Kolloquiums. (Forschungsbericht 43 des Zentrums I Bildungsforschung an der Universität Konstanz)*. Konstanz 1982.
- DAWYDOW, W. W., LOMPSCHER, J., MARKOWA, A. K.: *Ausbildung der Lerntätigkeit bei Schülern*. Volk und Wissen, Berlin 1982.
- DENIG, F., WEIS, V.: Zur Faktorenstruktur der Fachnoten an deutschen Sekundarschulen. *Zeitschrift für Empirische Forschung* 1970, 4, 210 – 232.
- Deutscher Bildungsrat (Empfehlungen der Bildungskommission): *Strukturplan für das Bildungswesen*. Bundesdruckerei, Bonn 1970/Klett, Stuttgart 1972.
- DICKER, H.: *Untersuchung zur Beurteilung von Mathematikaufgaben*. Unveröffentl. Diplomarbeit an der EWH Rheinland-Pfalz, Abt. Landau 1973.
- DIEDERICH, P. B., FRENCH, J. W., CARLTON, S. T.: Factors in Judgement of Writing Ability. *Research Bulletin, Series 61-15*. Educational Testing Service, Princeton N.J. 1961.
- DÖRNER, D.: *Problemlösen als Informationsverarbeitung*. Klett-Cotta, Stuttgart 1976.
- DONAT, H.: *Persönlichkeitsbeurteilung*. Ehrenwirth, München 1970<sup>2</sup>.
- DRENTH, P. J. D.: *Der Psychologische Test*. Barth, München 1969.
- DRESSMANN, H.: *Unterrichtsklima*. Urban & Schwarzenberg, München 1982.
- DUMKE, D.: Die Auswirkungen von Lehrererwartungen auf Intelligenz und Schulleistungen. *Psychologie in Erziehung und Unterricht* 1977, 24, 93 – 108.
- DUMKE, D.: *Förderung lernschwacher Schüler*. Reinhardt, München 1980.
- DYER, J. L., MILLER, L. B.: Note on CRANO, KENNY and CAMPBELL's «Does intelligence cause achievement?» *Journal of Educational Psychology* 1974, 66, 49 – 51.
- DYSON, A. P.: *Oral Examining in French*. Modern Language Association, London 1972.

- EBEL, R.L.: Criterion-referenced measurements: Limitations. *School Review* 1971, 282 – 288.
- EBEL, R.L.: *Essentials of educational measurement*. Englewood Cliffs, Prentice-Hall 1979.
- ECHTERHOFF, W.: Lern- und Veränderungsmessung. In: KLAUER, K.J. (Hrsg.): *Handbuch der Pädagogischen Diagnostik*, Bd. 1. Schwann, Düsseldorf 1978, 157 – 175.
- EELLS, W.C.: Die Zuverlässigkeit wiederholter Benotung von aufsatzähnlichen Prüfungsarbeiten. In: INGENKAMP, K. (Hrsg.): *Die Fragwürdigkeit der Zensurengebung*. Beltz, Weinheim 1971, 117 – 122.
- EGGERT, D.: *Hannover-Wechsler Intelligenztest für das Vorschulalter (HAWIVA)*. Deutsche Bearbeitung der Wechsler Preschool and Primary Scale of Intelligence. Huber, Bern 1975.
- EICHLER, W., THURNER, F.: Deutsch. In: KLAUER, K.J. (Hrsg.): *Handbuch der Pädagogischen Diagnostik*, Bd. 3. Schwann, Düsseldorf 1978, 761 – 770.
- EISEBITT, R.: *Untersuchungen über den Zusammenhang von Intelligenztests und Schulleistungen*. Unveröffentl. Staatsexamensarbeit, Universität Köln 1983.
- ENGEMANN, A.: *Prüfen als Handlung. Eine handlungspsychologische Analyse des Urteils in mündlichen Prüfungen*. Minerva, München 1983.
- FASSNACHT, G.: *Systematische Verhaltensbeobachtung*. Reinhardt, München 1979.
- FELLER, G., HUGOW, K.: *Mathematik-Test für 2. Klassen*. Beltz, Weinheim 1981.
- FEND, H.: *Konformität und Selbstbestimmung*. Beltz, Weinheim 1971.
- FEND, H.: *Überblick über Ergebnisse wissenschaftlicher Begleituntersuchungen zu Schulversuchen im Bereich der Sekundarstufe I in der Bundesrepublik Deutschland*. Gesamtschulinformationen 1976, 9, 113 – 141.
- FEUERSTEIN, R.: *The dynamic assessment of retarded performers*. University Park Press, Baltimore 1979.
- FIEDLER, F.E.: The contingency model: A theory of leadership effectiveness. In: PROHANSKY, H., SEIDENBERG, B. (Hrsg.): *Basic studies of social psychology*. Holt, London 1965, 538 – 551.
- FINGERHUT, W., LANGFELDT, H.: Erfahrungen mit dem Allgemeinen Schulleistungstest für 4. Klassen (AST 4). *Psychologie in Erziehung und Unterricht* 1973, 20, 249 – 257.
- FINGERHUT, W., LANGFELDT, H.-P.: Leistungsbeurteilung durch Notengebung. In: HELLER, K. (Hrsg.): *Leistungsbeurteilung in der Schule*. Quelle & Meyer, Heidelberg 1974, 253 – 270.
- FIPPINGER, F.: *Allgemeiner Schulleistungstest für vierte Klassen (AST 4)*. Beltz, Weinheim 1966, 1977<sup>2</sup>, Neuauflage 1978.
- FISCHER, G.H. (Hrsg.): *Psychologische Testtheorie*. Huber, Bern 1968 (a).
- FISCHER, G.H.: Kritik der klassischen Testtheorie. In: FISCHER, G.H. (Hrsg.): *Psychologische Testtheorie*. Huber, Bern 1968, 54 – 77 (b).
- FISCHER, G.H.: Stochastische Testmodelle. In: FISCHER, G.H. (Hrsg.): *Psychologische Testtheorie*. Huber, Bern 1968, 78 – 132 (c).
- FISCHER, G.H.: *Einführung in die Theorie psychologischer Tests*. Huber, Bern 1974.
- FISCHER, G.H.: Linear logistic models for the attitudinal and behavioral changes under the influence of man communication. In: KEMPF, W.H., REPP, B.H. (Hrsg.): *Some mathematical models for socialpsychology*. Huber, Bern 1977, 102 – 151.
- FISCHER, G.H., FORMANN, A.K.: Veränderungsmessung mittels linear-logistischer Modelle. *Zeitschrift für Differentielle und Diagnostische Psychologie* 1982, 3, 75 – 99.
- FLAMMER, A.: Die differentielle Psychologie in der Krise. In: PULVER, U., LANG, A., SCHMID, F.W. (Hrsg.): *Ist Psychodiagnostik verantwortbar?* Huber, Bern 1978, 76 – 81.
- FLAMMER, A., SCHMID, H.: *Lerntests: Konzept, Realisierung, Bewährung. Eine Übersicht*. Schweizerische Zeitschrift für Psychologie 1982, 41, 114 – 138.

- FLAMMER, A., PERRIG-CHIELLO, P., RÜEGG, T.: Zeugnisnoten vor dem Übertritt in die Sekundarstufe. Univ.-Verlag, Freiburg/Schweiz 1983.
- FLITNER, A.: Gerechtigkeit und Unrecht in unserem Schulwesen – Pädagogische Perspektiven. Bayerische Schule 1980, 33, 360 – 367.
- FLITNER, A.: Besinnung auf die Grundschule. Die Grundschule 1981, 13, 454 – 458.
- FLÜGGE, I.: Die Einschätzung der Schwierigkeit von Aufgaben durch den Lehrer. Mitteilungen und Nachrichten der Hochschule für Internationale Pädagogische Forschung, Frankfurt/M. 1954.
- FREUDENTHAL, H.: Schülerleistungen im internationalen Vergleich. Zeitschrift für Pädagogik 1975, 21, 889 – 910.
- FRICKE, R.: Testgütekriterien bei lehrzielorientierten Tests. Zeitschrift für erziehungswissenschaftliche Forschung 1972, 150 – 175 (a).
- FRICKE, R.: Über Meßmodelle in der Schulleistungsdiagnostik. Schwann, Düsseldorf 1972 (b).
- FRICKE, R.: Kriteriumsorientierte Leistungsmessung. Kohlhammer, Stuttgart 1974.
- FROSTIG, M., REINARTZ, A., REINARTZ, E.: Visuelle Wahrnehmungsförderung. Crüwell, Dortmund 1977.
- FUNKE, E.H.: Grundschulzeugnisse und Sonderschulbedürftigkeit. Marhold, Berlin 1972.
- FURCK, C.L.: Das pädagogische Problem der Leistung in der Schule. Beltz, Weinheim 1961, 1975<sup>5</sup>.
- GAEDIKE, A.-K.: Determinanten der Schulleistung. In: HELLER, K. (Hrsg.): Leistungsbeurteilung in der Schule. Quelle & Meyer, Heidelberg 1974, 46 – 93.
- GAGE, N.L., BERLINER, D.C.: Pädagogische Psychologie – Lehrerhandbuch. Erziehungswissenschaftliche Grundlagen für die Unterrichtspraxis. Urban & Schwarzenberg, München 1977.
- GAGNÉ, R.M.: Die Bedingungen des menschlichen Lernens. Schroedel, Hannover 1969, 1973<sup>3</sup>.
- GAGNÉ, R.M.: Der Erwerb von Wissen. In: HOFER, M., WEINERT, F.E. (Hrsg.): Pädagogische Psychologie. Grundlagentexte 2: Lernen und Instruktion. Fischer, Frankfurt/M. 1973.
- GARVIN, A.D.: The Applicability of Criterion-Referenced Measurement by Content Area and Level. In: PROPHAM, W.J. (Hrsg.): Criterion-Referenced Measurement. Educational Technology Publications, Englewood Cliffs, N.J. 1971.
- GAUDE, P., TESCHNER, W.P.: Objektivierte Leistungsmessung in der Schule. Diesterweg, Frankfurt/M. 1970.
- GEHLEN, A.: Arbeiten – Ausruhen – Ausnützen. Wesensmerkmale des Menschen. In: GEHLEN, A. et al.: Sinn und Unsinn des Leistungsprinzips. DTV, München 1975<sup>3</sup>, 7 – 19.
- GEHLEN, A. et al.: Sinn und Unsinn des Leistungsprinzips. DTV, München 1975<sup>3</sup>.
- GERSTEIN, H.: Erfolg und Versagen im Gymnasium. Beltz, Weinheim 1972.
- GERSTER, J.: Schülerfehler bei schriftlichen Rechenaufgaben. Herder, Freiburg 1982.
- GIBSON, E., LEVINE, H.: The Psychology of Reading. The MIT Press, Cambridge, Mass. und London 1976<sup>3</sup>.
- GLASER, R.: Instructional technology and the measurement of learning outcomes. American Psychologist 1963, 18, 519 – 521.
- GLASER, R.: A Criterion-Referenced Test. In: PROPHAM, W.J. (Hrsg.): Criterion-Referenced Measurement. Educational Technology Publications, Englewood Cliffs, N.J. 1971.
- GLÜCK, G.: Rechenleistungen und Rechenfehler. Dissertation, Universität Tübingen 1971.

- GOLDSTEIN, H.: The Design and Analysis of Longitudinal Studies. Academic Press, New York und San Francisco 1979.
- GOODMAN, L.A.: Analyzing qualitative/categorical data. Addison-Wesley, London 1978.
- GOODMAN, Y.M., BURKE, C.L.: Reading Miscue Inventory – Manual. New York 1972.
- GRAUMANN, C.F.: Interaktion und Kommunikation. In: GRAUMANN, C.F. (Hrsg.): Sozialpsychologie (= Handbuch der Psychologie, Bd. VII/2. Halbb.). Hogrefe, Göttingen 1972.
- GRAUMANN, C.F.: Grundzüge der Verhaltensbeobachtung. In: GRAUMANN, C.F., HECKHAUSEN, H. (Hrsg.): Pädagogische Psychologie, Bd. 1. Fischer, Frankfurt/M. 1973, 14 – 41.
- GRIMM, M. et al.: Regensburger Modell – Lesen lernen mit Habakuk. Sellier, Freising/Diesterweg, Frankfurt/M. 1978.
- GROEN, G.J., RESNICK, L.B.: Can preschool children invent addition algorithms? Journal of Educational Psychology 1977, 69, 645 – 652.
- GROFFMANN, K.J.: Die Entwicklung der Intelligenzmessung. In: GROFFMANN, K.J., MICHEL, L. (Hrsg.): Intelligenz- und Leistungsdiagnostik (= Enzyklopädie der Psychologie, Themenbereich B, Serie II, Bd. 2). Hogrefe, Göttingen 1983, 1 – 103.
- GRÖNLUND, N.E.: Measurement and evaluation in teaching. McMillan, New York 1968.
- GRZESIK, J.: Die Steuerung von Lernprozessen im Unterricht. Quelle & Meyer, Heidelberg 1976.
- GRZESIK, J.: Unterrichtsplanung. Quelle & Meyer, Heidelberg 1979.
- GUILFORD, J.P.: The structure of intellect. Psychological Bulletin 1956, 53, 267 – 293.
- GUILFORD, J.P., HOEPPNER, R.: The Analysis of Intelligence. McGraw-Hill, New York 1971. – Dt. Die Analyse der Intelligenz. Beltz, Weinheim 1976.
- GULLIKSEN, H.: Theory of Mental Tests. Wiley, New York 1950.
- GUTHKE, J.: Zur Diagnostik der intellektuellen Lernfähigkeit. Deutscher Verlag der Wissenschaften, Berlin 1972/Klett, Stuttgart 1977.
- GUTHKE, J.: The Learning Test Concept – An Alternative to the Traditional Static Intelligence Test. The German Journal of Psychology 1982, 6, 306 – 324.
- GUTHKE, J.: Mengenfolgen-Test (MFT). Psychodiagnostisches Zentrum, Sektion Psychologie der Humboldt-Universität zu Berlin. Testzentrale, Stuttgart 1983.
- GUTHKE, J., WOHLRAB, U. (Hrsg.): Neuere Ergebnisse der Lerntestforschung – Diagnostische Programme als Lerntestvariante. Forschungsbericht der Sektion Psychologie. Karl-Marx-Universität, Leipzig 1982.
- GUTTMAN, L.: The basis for scalogram analysis. In: STOURFER, S.A. et al. (Hrsg.): Studies in social psychology in World War II, Vol. IV. Princeton, N.J. 1950.
- GUTTMAN, L., SCHLESINGER, I.M.: Development of diagnostic analytical and mechanical ability tests through facet design and analysis. The Israel Institute of Applied Social Research, Jerusalem 1966.
- HAASE, H.: Tests im Bildungswesen. Hogrefe, Göttingen 1978.
- HACKER, W.: Allgemeine Arbeits- und Ingenieurpsychologie. Deutscher Verlag der Wissenschaften, Berlin 1973.
- HAENISCH, H., LUKESCH, H.: Ist die Gesamtschule besser? Urban & Schwarzenberg, München 1980.
- HAERTEL, G.D., WALBERG, H.J., WEINSTEIN, T.: Psychological Models of Educational Performance: A Theoretical Synthesis of Constructs. Review of Educational Research 1983, 53, 75 – 91.
- HAHN, H.: Flensburger Norm für die Aufsatzbeurteilung. Die Pädagogische Provinz 1969, 20, 309 – 310.
- HAMBLETON, R.K., NOVICK, M.R.: Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement 1973, 10, 159 – 170.

- HAMBLETON, R. K., SWAMINATHAN, H., ALGINA, J., COULSON, D. B.: Criterion-referenced Testing and Measurement: A Review of Technical Issues and Developments. *Review of Educational Research* 1978, 48, 1 – 47.
- HAMMAKER, M., SHAFTO, M., TRABASSO, T.: Judging creativity. A method for assessing how and by what criteria it is done. *Journal of Educational Psychology* 1975, 67, 478 – 483.
- HANEIS, E.: Über das falsche Interesse an einer richtigen Aufsatzbenotung. *Linguistische Berichte* 1976, 45, 83 – 89.
- HANKE, B., LOHMÖLLER, J. B., MANDL, H.: *Schülerbeurteilung in der Grundschule*. Oldenbourg, München 1980.
- HARDESTY, F. P., PRIESTER, H. J.: *Hamburg-Wechsler Intelligenztest für Kinder (HAWIK)*. Huber, Bern 1966<sup>3</sup> – Revisionsfassung (HAWIK-R) 1983.
- HARRIS, C. W.: An interpretation of Livingston's reliability coefficient for criterion-referenced tests. *Journal of Educational Measurement* 1972, 9, 27 – 29.
- HARTOG, P., RHODES, E. C.: Prüfungszensuren in Geschichte und Englisch. In: INGENKAMP, K. (Hrsg.): *Die Fragwürdigkeit der Zensurengebung*. Beltz, Weinheim 1971, 78 – 89.
- HARTUNG, K.: Schulversagen aus organischer Ursache. *Öffentliches Gesundheitswesen* 1973, 35, 87 – 92.
- HASEMANN, K.: Verhaltensbeobachtung. In: HEISS, R. (Hrsg.): *Handbuch der Psychologie*, Bd. 6: *Psychologische Diagnostik*. Hogrefe, Göttingen 1964, 807 – 836.
- HASEMANN, K.: Verhaltensbeobachtung und Ratingverfahren. In: GROFFMANN, K. J., MICHEL, L. (Hrsg.): *Verhaltensdiagnostik (= Enzyklopädie der Psychologie, Themenbereich B, Serie II, Bd. 4)*. Hogrefe, Göttingen 1983, 434 – 488.
- HAUSSER, K. (Hrsg.): *Modelle schulischer Differenzierung*. Urban & Schwarzenberg, München 1981.
- HECKHAUSEN, H.: Förderung der Lernmotivierung und der intellektuellen Tüchtigkeit. In: ROTH, H. (Hrsg.): *Begabung und Lernen*. Klett, Stuttgart 1969, 193 – 228.
- HECKHAUSEN, H.: Leistung und Chancengleichheit. Hogrefe, Göttingen 1974 (a).
- HECKHAUSEN, H.: Motive und ihre Entstehung. In: WEINERT, F. E. et al. (Hrsg.): *Funk-Kolleg Pädagogische Psychologie*, Bd. 1. Fischer, Frankfurt 1974, 133 – 168 (b).
- HECKHAUSEN, H. (Hrsg.): *Fähigkeit und Motivation in erwartungswidriger Schulleistung*. Hogrefe, Göttingen 1980 (a).
- HECKHAUSEN, H.: *Motivation und Handeln*. Springer, Berlin 1980 (b).
- HEIPCKE, K.: *Lehr- und Handlungsziele im Unterricht*. betrifft: *erziehung* 1974, 8, 15 – 19.
- HELLER, K.: *Aktivierung der Bildungsreserven*. Huber, Bern/Klett, Stuttgart 1970.
- HELLER, K.: *Intelligenzmessung*. Neckar-Verlag, Villingen 1973.
- HELLER, K. (Hrsg.): *Leistungsbeurteilung in der Schule*. Quelle & Meyer, Heidelberg 1974, 1978<sup>3</sup>.
- HELLER, K.: *Beratung beim Übergang von der Primarstufe zur Sekundarstufe I*. In: HELLER, K. (Hrsg.): *Handbuch der Bildungsberatung*, Bd. 2. Klett, Stuttgart 1975, 430 – 473.
- HELLER, K.: *Computerunterstützte Interpretation von Testbefunden in der Schullaufbahnberatung*. In: HELLER, K. (Hrsg.): *Handbuch der Bildungsberatung*, Bd. 3. Klett, Stuttgart 1976, 879 – 900.
- HELLER, K.: *Beurteilung und Beratung beim Übergang in die Sekundarstufe*. In: BOLSCO, D., SCHWARZER, C. (Hrsg.): *Beurteilen in der Grundschule*. Urban & Schwarzenberg, München 1979, 185 – 206.
- HELLER, K.: *Entscheidungsmodell zur Übertrittsberatung*. In: HELLER, K., NICKEL, H. (Hrsg.): *Modelle und Fallstudien zur Erziehungs- und Schulberatung*. Huber, Bern 1982, 111 – 118.

- HELLER, K., ROSEMANN, B.: Planung und Auswertung empirischer Untersuchungen. Klett-Cotta, Stuttgart 1974, 1981<sup>2</sup>.
- HELLER, K., GAEDIKE, A.-K., WEINLÄDER, H.: Kognitiver Fähigkeits-Test für 4. bis 13. Klassen (KFT 4 – 13). Beltz, Weinheim 1976, 1984<sup>2</sup>.
- HELLER, K., NICKEL, H., ROSEMANN, B.: Beurteilen und Beraten (= Bd.4 des Studienprogrammes «Psychologie in der Erziehungswissenschaft»). Klett-Cotta, Stuttgart 1978 (a).
- HELLER, K., ROSEMANN, B., STEFFENS, K.: Prognose des Schulerfolgs. Beltz, Weinheim 1978 (b).
- HELLER, K., NICKEL, H. (Hrsg.): Modelle und Fallstudien zur Erziehungs- und Schulberatung. Huber, Bern 1982.
- HELMREICH, R.: Strategien zur Auswertung von Längsschnittdaten. Klett, Stuttgart 1977.
- HENTIG, H. v.: Systemzwang und Selbstbestimmung. Klett, Stuttgart 1968, 1970<sup>3</sup>.
- HENTIG, H. v.: Das Beurteilungssystem der Bielefelder Laborschule. Neue Sammlung 1982, 22, 238 – 260.
- HERBIG, M.: Verfahren zur experimentellen Validierung lehrzielorientierter Tests. Zeitschrift für erziehungswissenschaftliche Forschung 1973, 7, 235 – 239.
- HESS, T., MÜLLER, A.: Schuleintrittsdiagnostik: Neue Perspektiven. Psychologie in Erziehung und Unterricht 1984, 31, 57 – 63.
- HETZER, H.: Die seelischen Veränderungen des Kindes bei dem ersten Gestaltwandel. Barth, Leipzig 1936.
- HETZER, H.: Zum Problem der Schulreife. Westermanns Pädagogische Beiträge 1953, 5, 6 – 15.
- HETZER, H.: Aktuelle Probleme beim Eintritt in die Schule in psychologisch-pädagogischer Sicht. Vortrag auf der Studententagung der Universität des Saarlandes, Saarbrücken 1969.
- HETZER, H., TENT, L.: Der Schulreifetest – Auslesemittel oder Erziehungshilfe? Dr. Piorowski, Lindau 1958/Beltz, Weinheim 1969<sup>3</sup>.
- HETZER, H., TENT, L.: Weilburger Testaufgaben für Schulanfänger (WTA). Beltz, Weinheim 1971.
- HILKE, R.: Grundlagen normorientierter und kriterienorientierter Tests. Huber, Bern 1980.
- HIVELY, W. et al.: A «univers defined» system of arithmetic achievement tests. Journal of Educational Measurement 1968, 5, 275 – 290.
- HÖGER, D.: Analyse der Intelligenzstruktur bei männlichen Gymnasiasten der Klassen 6 bis 9 (Untersekunda – Oberprima). Psychologische Forschung 1964, 27, 419 – 474.
- HOFER, A.: Die pädagogische, didaktische und psycholinguistische Problematik von Testverfahren zur Feststellung von Legasthenie. In: HOFER, A. (Hrsg.): Lesenlernen: Theorie und Unterricht. Schwann, Düsseldorf 1976, 321 – 348.
- HOFER, M.: Zur impliziten Persönlichkeitstheorie von Lehrern. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1970, 2, 197 – 209.
- HOFER, M. (Hrsg.): Informationsverarbeitung und Entscheidungsverhalten von Lehrern. Urban & Schwarzenberg, München 1981.
- HOFER, M., WEINERT, F.E., ZIELINSKI, W.: Kognitive Bedingungen individualisierendes Verhalten von Lehrern. DFG-Abschlußbericht, Braunschweig 1979.
- HOLTZ, K.L.: Diagnostik im Rahmen der Pädagogischen Verhaltensmodifikation. In: KLAUER, K.J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd.4. Schwann, Düsseldorf 1978, 873 – 889.
- HOLZMAN, T.G., GLASER, R., PELLEGRINO, J.W.: Process training derived from a computer simulation theory. Memory and Cognition 1976, 4, 349 – 356.

- HOPF, C., NEVERMANN, K., RICHTER, H.: Schulaufsicht und Schule. Eine empirische Analyse der administrativen Bedingungen schulischer Erziehung. Klett-Cotta, Stuttgart 1980.
- HOPF, D.: Übergangsauslese und Leistungsdifferenzierung. Diesterweg, Frankfurt/M. 1970.
- HOPKINS, K. D., STANLEY, J. C.: Educational and psychological measurement and evaluation. Prentice-Hall, Englewood Cliffs 1981.
- HORN, R.: Lernziele und Schülerleistung. Beltz, Weinheim 1972.
- HORN, R., INGENKAMP, K., JÄGER, R. S. (Hrsg.): Tests und Trends 83 (Jahrbuch der Pädagogischen Diagnostik). Beltz, Weinheim 1983.
- HORN, W.: Leistungsprüfsystem (LPS). Hogrefe, Göttingen 1962.
- HORN, W.: Prüfsystem für Schul- und Bildungsberatung (PSB). Hogrefe, Göttingen 1969, 1983<sup>2</sup>.
- HORST, P.: Messung und Vorhersage. Eine Einführung in die psychologische Testtheorie. Beltz, Weinheim 1971.
- HUBER, H. P.: Psychometrische Einzelfalldiagnostik. Beltz, Weinheim 1973.
- HUMMELL, H. J., ZIEGLER, R. (Hrsg.): Korrelation und Kausalität, 3 Bde. Enke, Stuttgart 1976.
- HUSÉN, T.: Policy implications of individual differences in learning ability: A comparative perspective. Scandinavian Journal of Educational Research 1978, 22, 173 – 191.
- HUYNH, H.: On the Reliability of Decisions in Domain-Referenced Testing. Journal of educational measurement 1976, 13, 253 – 264.
- HUYNH, H.: The Kappamax Reliability Index for Decisions in Domain-Referenced Testing. Paper Presented at the Annual Meeting of the American Educational Research Association, New York 1977.
- INGENDAHL, W.: Sprechen und Schreiben. Quelle & Meyer, Heidelberg 1975.
- INGENKAMP, K.: Die deutschen Schulleistungstests. Beltz, Weinheim 1962.
- INGENKAMP, K.: Sind Zensuren aus verschiedenen Klassen vergleichbar? Erziehung 1969, 2, 12 – 14 (a).
- INGENKAMP, K.: Zur Problematik der Jahrgangsklasse. Beltz, Weinheim 1969 (b).
- INGENKAMP, K. (Hrsg.): Die Fragwürdigkeit der Zensurenggebung. Beltz, Weinheim 1971, 1975<sup>2</sup> (a).
- INGENKAMP, K.: Probleme der schulischen Leistungsbeurteilung unter besonderer Berücksichtigung des Deutschunterrichts. Der Deutschunterricht 1971, 23, 54 – 76 (b).
- INGENKAMP, K.: Tests in der Schulpraxis. Beltz, Weinheim 1971 (c).
- JACKSON, R.: Developing Criterion-Referenced Tests. TM Reports Number 1, ERIC Clearinghouse on Tests, Measurement and Evaluation, Princeton, N.J. 1970.
- JACOBS, B., STRITTMATTER, P.: Der schulängstliche Schüler. Eine empirische Untersuchung über mögliche Ursachen und Konsequenzen der Schulangst. Urban & Schwarzenberg, München 1979.
- JÄGER, R.: Differentielle Diagnostizierbarkeit in der psychologischen Diagnostik. Hogrefe, Göttingen 1978.
- JÄGER, R., BEETZ, E., ERLER, R., WALTER, R.: Mannheimer Schuleingangsdiagnostikum (MSD). Beltz, Weinheim 1976.
- JÄGER, R. S.: Der Moderatoransatz in der psychologischen Diagnostik. In: Mitteilungen und Nachrichten des DIPF, Nr. 112/113, Frankfurt/M. 1983, 121 – 131.
- JAGODZINSKI, W., ZÄNGLE, M.: Über einige Probleme der Anwendung der Pfadanalyse – Bemerkungen zu einer Studie über soziokulturelle Determinanten der Fähigkeit zum Role-Taking. Zeitschrift für Soziologie 1977, 6, 49 – 61.
- JANTZEN, W.: Untersuchungen zur Faktorenstruktur von Intelligenz und Schulleistungen bei guten und schlechten Schülern dritter Grundschulklassen (und zur Einstel-

- lungsstruktur ihrer Lehrer). Zeitschrift für empirische Forschung 1971, 5, 44 – 62 und 92 – 106.
- JENCKS, C. et al.: Chancengleichheit. Rowohlt, Hamburg 1973.
- JENDROWIAK, H., KREUZER, K.: Lehrer beurteilen Lehrer. Methoden – Befunde – Alternativen. Schwann, Düsseldorf 1980.
- JUNG, U.: Legasthenie als linguistisches Defizit. Linguistische Berichte 1976, 41, 22 – 38.
- KALVERAM, K. T.: Über Faktorenanalyse, Kritik eines theoretischen Konzepts und seine mathematische Neuformulierung. Archiv für die gesamte Psychologie 1970, 118, 92 – 118.
- KANT, I.: Kritik der reinen Vernunft. Text der Ausgabe 1781 mit Beifügung sämtlicher Abweichungen der Ausgabe 1787, Hrsg. K. KEHRBACH. Reclam jun., Leipzig 1877<sup>2</sup>.
- KAUTTER, H., STORZ, L.: Schulleistungstestbatterie für Lernbehinderte (SBL I, SBL II). Beltz, Weinheim 1972.
- KAUTTER, H.-J., MUNZ, W.: Verfahren der Aufnahme und Überweisung in die Sonderschule. In: Deutscher Bildungsrat: Gutachten und Studien der Bildungskommission 34, Sonderpädagogik, Bd. 3. Klett, Stuttgart 1974, 235 – 285.
- KEEVES, J. P.: Educational environment and student achievement. Almqvist & Wiksell, Stockholm 1972.
- KEMMLER, L.: Erfolg und Versagen in der Grundschule. Hogrefe, Göttingen 1967, 1975<sup>3</sup>.
- KEMMLER, L.: Schulerfolg und Schulversagen. Hogrefe, Göttingen 1976.
- KERN, A.: Sitzenbleiberelend und Schulreife. Herder, Freiburg 1951, 1966<sup>5</sup>.
- KIFER, E., BRAMBLE, W.: The Calibration of a Criterion-Referenced Test. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago 1974.
- KLAFKI, W.: Sinn und Unsinn des Leistungsprinzips in der Erziehungswissenschaft. In: GEHLEN, A. et al. (Hrsg.): Sinn und Unsinn des Leistungsprinzips. DTV, München 1975<sup>3</sup>, 73 – 110.
- KLAFKI, W.: Die bildungstheoretische Didaktik. Westermanns Pädagogische Beiträge 1980, 1, 32 – 37.
- KLAHR, D.: Informationsverarbeitungsmodelle der Denkentwicklung. In: KLUWE, R. H., SPADA, H. (Hrsg.): Studien der Denkentwicklung. Huber, Bern 1981.
- KLAUER, K. J.: Neuere Entwicklungen im Bereich der Lehrstoffanalyse – Schwerpunkt Makroanalyse. Zeitschrift für Pädagogik 1976, 22, 387 – 398.
- KLAUER, K. J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, 4 Bde. Schwann, Düsseldorf 1978 (a).
- KLAUER, K. J.: Perspektiven Pädagogischer Diagnostik. In: KLAUER, K. J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd. 1. Schwann, Düsseldorf 1978, 3 – 14 (b).
- KLAUER, K. J.: Gedanken zu einer ipsativen Veränderungsmessung. In: SEIDEL, G. (Hrsg.): Orientierungen zum pädagogischen Handeln. Hogrefe, Göttingen 1982, 98 – 114.
- KLAUER, K., FRICKE, R., HERBIG, M., RUPPRECHT, H., SCHOTT, F.: Lehrzielorientierte Tests. Schwann, Düsseldorf 1972.
- KLEBER, E. W., MEISTER, H., SCHWARZER, C., SCHWARZER, R.: Beurteilung und Beurteilungsprobleme. Beltz, Weinheim 1976.
- KLEBER, E. W.: Probleme der diagnostischen Validität der Informationserhebung und Lösungsansätze für die Schuleingangsdiagnostik. In: MANDL, H., KRAPP, A. (Hrsg.): Schuleingangsdiagnose. Hogrefe, Göttingen 1978, 121 – 133 (a).
- KLEBER, E. W.: Probleme des Lehrerurteils. In: KLAUER, K. J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd. 3. Schwann, Düsseldorf 1978, 589 – 618 (b).
- KLEBER, E. W.: Tests in der Schule. Reinhardt, München 1979.
- KLEIN, S. P., KOSECOFF, J. P.: Procedures and Issues in the Validation of Criterion-Refer-

- enced Tests. Paper presented at the Annual Meeting of the National Council of Measurement in Education, Washington, D.C. 1975.
- KLEITER, E. F.: K-MICRO-Stat 1, BASIC-Statistik-Programm-System für Micro-Computer. Zeitschrift für Empirische Pädagogik 1982, 6, 1 – 21. Manual: NATIC, Wiesbaden 1982.
- KLEITER, E. F., PETERMANN, F.: Abbildung von Lernwegen. Oldenbourg, München 1977.
- KLEITER, E. F., TIMMERMANN, G.: Über einen Algorithmus zur hierarchischen Voraussetzungs-Struktur-Analyse. Psychologische Beiträge 1977, 19, 355 – 390.
- KLINGER, E.: Fantasy need achievement as a motivational construct. Psychological Bulletin 1966, 66, 291 – 308.
- KLUWE, R.: Wissen und Denken: Modelle, empirische Befunde und Perspektiven für den Unterricht. Kohlhammer, Stuttgart 1979.
- KLUWE, R. H., SPADA, H. (Hrsg.): Studien zur Denkentwicklung. Huber, Bern 1981.
- KÖTTER, L., GRAU, U.: Zur Bedingtheit der uneinheitlichen Benotung von Schüleraufsätzen. Zeitschrift für experimentelle und angewandte Psychologie 1965, 12, 278 – 301.
- KOOB, B.: Theorie, Methoden und Probleme der Schuleingangsdiagnostik. In: JÄGER, R. S. et al. (Hrsg.): Tests und Trends (Jahrbuch der Pädagogischen Diagnostik). Beltz, Weinheim 1981.
- KORMANN, A.: Veränderungsmessung. In: SCHIEFELE, H., KRAPP, A. (Hrsg.): Handlexikon zur Pädagogischen Psychologie. Ehrenwirth, München 1981, 391 – 394.
- KORMANN, A.: Möglichkeiten von Lerntests für Diagnose und Optimierung von Lernprozessen. In: INGENKAMP, K. et al. (Hrsg.): Tests und Trends (Jahrbuch der Pädagogischen Diagnostik). Beltz, Weinheim 1982, 97 – 117.
- KORMANN, A.: Zum gegenwärtigen Stand einer förderungsorientierten Diagnostik. Positionreferat am 12. Kongreß für Angewandte Psychologie in Düsseldorf 1983. Hogrefe, Göttingen 1984.
- KORMANN, A., SPORER, S. L.: Learning Tests – Concepts and Critical Evaluation. Studies in Educational Evaluation 1983, 9, 169 – 184.
- KORNMAN, R.: Minimalisieren Schulleistungstests die Zahl der Fehlentscheidungen? Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1972, 4, 282 – 286.
- KORNMAN, R., ENDRIGKEIT, F., SANDER, H.: Sind lernbehinderte Schüler in Gruppen-Intelligenztests benachteiligt? Diagnostica 1972, 18, 111 – 121.
- KORNMAN, R., MEISTER, H., SCHLEE, J. (Hrsg.): Förderungsdiagnostik. Schindele, Heidelberg 1983.
- KOZDON, B.: Das Leistungsprinzip in Gesellschaft und Schule. Klinkhardt, Bad Heilbrunn 1976.
- KRAPP, A.: Bedingungen des Schulerfolges. Oldenbourg, München 1973.
- KRAPP, A.: Bedingungsfaktoren der Schulleistung. Psychologie in Erziehung und Unterricht 1976, 23, 91 – 109.
- KRAPP, A.: Ursachen und Folgen des Schulversagens. In: KLAUER, K. J., REINARTZ, A. (Hrsg.): Handbuch der Sonderpädagogik, Bd. 9: Sonderpädagogik in allgemeinbildenden Schulen. Marhold, Berlin 1978, 31 – 46.
- KRAPP, A.: Prognose und Entscheidung. Beltz, Weinheim 1979.
- KRAPP, A.: Schulreife und Schulfähigkeit. In: ROST, D. H. (Hrsg.): Entwicklungspsychologie für die Grundschule. Klinkhardt, Bad Heilbrunn 1980.
- KRAPP, A., MANDL, H.: Zur Problematik der Schulleistungstests. Die Bayerische Schule 1973, 26, 209 – 212. – Abgedruckt auch in der Hamburger Lehrerzeitung 1973, 26, 365 – 371.
- KRAPP, A., MANDL, H.: Vorhersage und Erklärung der Schulleistung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1976, 8, 192 – 219.

- KRAPP, A., MANDL, H.: Einschulungsdiagnostik. Eine Einführung in Probleme und Methoden der pädagogisch-psychologischen Diagnostik. Beltz, Weinheim 1977.
- KRAPP, A., PRELL, S.: Klassifikation pädagogisch-diagnostischer Aufgaben. In: KLAUER, K. J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd.4. Schwann, Düsseldorf 1978, 831 – 854.
- KRATHWOHL, D. R., BLOOM, B. S., MASIA, B. B.: Taxonomie of Educational Objectives. Handbook 2: Affective Domain. David McKay, New York 1964. – Dt. Taxonomie von Lernzielen im affektiven Bereich. Beltz, Weinheim 1975, 1978<sup>2</sup>.
- KRATZMEIER, H.: Heidelberger Sprachfreier Schulreifeftest (HSST). Marhold, Berlin 1969 (a).
- KRATZMEIER, H.: Reutlinger Test für Schulanfänger (RTS). Beltz, Weinheim 1969 (b).
- KRAUTH, J.: The present state in the measurement of success and change. In: LÜER, G. (Hrsg.): Bericht über den 33. Kongreß der Deutschen Gesellschaft für Psychologie in Mainz 1982, Bd.2. Hogrefe, Göttingen 1983, 769 – 776 (a).
- KRAUTH, J.: Typenlehre. In: BREDEKAMP, J., FEGER, H. (Hrsg.): Strukturierung und Reduzierung von Daten (= Enzyklopädie der Psychologie, Themenbereich B, Serie I, Bd.4) 1983, 440 – 496 (b).
- KRAUTH, J., LIENERT, G. A.: Fundamentals and Modifications of configurational frequency analysis. Interdisciplinaria 1982, 2, 1 – 14.
- KREPPER, W., WINTHER, A.: Anwendung eines Selbstregulationsprogramms in Schulklassen. In: PERLWITZ, E. (Hrsg.): Verhaltensformung in der Schule. Westermann, Braunschweig 1978, 116 – 131.
- KROHNE, H. W.: Angst bei Schülern und Studenten. Hoffmann und Campe, Hamburg 1977.
- KROPP, R. P., STOKER, H. W.: The Construction and Validation of Tests of the Cognitive Processes as Described in the Taxonomy of Educational Objectives. Florida State University, 1966.
- KRUG, S., RHEINBERG, F.: Erwartungswidrige Schulleistung im Entwicklungsverlauf und ihre Ursachen: Ein überholtes Konstrukt in neuer Sicht. In: HECKHAUSEN, H. (Hrsg.): Fähigkeit und Motivation in erwartungswidriger Schulleistung. Hogrefe, Göttingen 1980, 53 – 102.
- KUBINGER, K.: Die Typisierung in Kinder mit hoher und geringer latenter Lernfähigkeit. Zeitschrift für Differentielle und Diagnostische Psychologie 1983, 4, 129 – 140.
- KUBINGER, K., NEUMAYER, R., SCHUBERT, M. T., WURST, E.: Apparative Konzentrationsförderung bei Kindern. Zeitschrift für Klinische Psychologie 1979, 8, 56 – 66.
- KÜFFNER, H.: Fehlerorientierte Tests: Konzept und Bewährungskontrolle. Beltz, Weinheim 1981.
- KÜFFNER, H., SIKORSKI, P. B.: Zur Diagnose des Lernerfolgs mit fehlerorientierten Tests. In: MICHAELIS, W. (Hrsg.): Bericht über den 32. Kongreß der Deutschen Gesellschaft für Psychologie in Zürich. Hogrefe, Göttingen 1981, 514 – 517.
- KÜHN, R.: Zur differentiellen Vorhersagbarkeit von Schulerfolg. Eine Erkundungsstudie zur Auffindung von Moderatoreffekten. DIPP, Frankfurt/M. 1977.
- KÜHN, R.: Bedingungen für Schulerfolg. Hogrefe, Göttingen 1983.
- Kultusminister NW: Ausbildungs- und Prüfungsordnung, Oberstufe des Gymnasiums. Greven, Köln 1979.
- Kultusminister NW: Richtlinien für die gymnasiale Oberstufe in Nordrhein-Westfalen. Greven, Köln 1981 (a).
- Kultusminister NW: Richtlinien und Lehrpläne für den Sport in den Schulen im Lande Nordrhein-Westfalen, Bd. V, Gymnasiale Oberstufe. Greven, Köln 1981 (b).
- Kultusministerkonferenz (KMK): Einheitliche Prüfungsanforderungen in der Abiturprüfung (Beschlüsse der KMK). Luchterhand, Neuwied 1975, 1979<sup>2</sup>.

- KUTZER, R.: Mengen- und Figuren-Struktur-Niveau-orientiertes Unterrichtswerk. Diesterweg, Frankfurt/M. 1979.
- KUTZER, R., PROBST, H.: Strukturbezogene Aufgaben zur Prüfung mathematischer Einsichten. Beltz, Weinheim 1983.
- KVALE, S.: Prüfung und Herrschaft. Beltz, Weinheim 1972.
- LANDA, L.N.: Algorithmierung im Unterricht. Volk und Wissen, Berlin 1969.
- LANGE, B., KUFFNER, H., SCHWARZ, R.: Schulangst und Schulverdrossenheit (= Forschungsberichte des Landes NW, Nr. 3153). Westdeutscher Verlag, Opladen 1983.
- LANGEHEINE, R.: Kausalanalyse qualitativer Daten mit manifesten und latenten Variablen. Zeitschrift für Sozialpsychologie 1982, 13, 163 – 176.
- LANGFELDT, H.-P.: Determinanten von Bildungsempfehlungen. Dissertation, Universität Bonn 1977.
- LANGFELDT, H.-P.: Möglichkeiten und Grenzen konventioneller Leistungsprüfungen im Unterricht. In: TWELLMANN, W. (Hrsg.): Handbuch Schule und Unterricht, Bd. 4/2. Schwann, Düsseldorf 1981, 705 – 721.
- LAUTMANN, R.: Gesellschaftliche Mechanismen im Examen. In: ECKSTEIN, B. (Hrsg.): Hochschulprüfungen: Rückmeldung oder Repression? Blickpunkt Hochschuldidaktik 1971, 13, 35 – 41.
- LAVIN, D.E.: The predication of performance. Russel Sage Foundation, New York 1965.
- LAZARSFELD, O., HENRY, N.W.: Latent Structure Analysis. Houghton Mifflin, Boston (Mass.) 1968.
- LENK, H.: Erklärung, Prognose, Planung. Rombach, Freiburg 1972.
- LICHTENSTEIN-ROTHER, I.: Das Problem der Leistung in der Schule. Die Deutsche Schule 1964, 54, 469 – 485.
- LICHTENSTEIN-ROTHER, I. (Hrsg.): Schulleistung und Leistungsschule. Klinkhardt, Bad Heilbrunn 1971.
- LIENERT, G.A.: Testaufbau und Testanalyse. Beltz, Weinheim 1961, 1967<sup>2</sup>, 1969<sup>3</sup>.
- LIVINGSTON, S.A.: Criterion-referenced applications of classical test theory. Journal of Educational Measurement 1972, 9, 13 – 25.
- LOCKOWANDT, O.: FROSTIGS Entwicklungstest der visuellen Wahrnehmung (FEW). Beltz, Weinheim 1974, 1976<sup>2</sup>.
- LÖSCHENKOHL, E.: Über den prognostischen Wert von Schulreifetests. Beltz, Weinheim 1975.
- LÖSCHENKOHL, E.: Multivariate Methoden zur Aufklärung von Schulerfolg und Lernerfolg. In: MANDL, H., KRAPP, A. (Hrsg.): Schuleingangsdiagnose. Hogrefe, Göttingen 1978, 103 – 112.
- LOMPSCHE, J.: Theoretische und experimentelle Untersuchungen zur Entwicklung geistiger Fähigkeiten. Volk und Wissen, Berlin 1972.
- LOTTER, H.: Geeichte Testdiktate für das 4. – 8. Schuljahr. Prögel, Sachsen 1967, 1978<sup>4</sup>.
- LUEG, C.H.: Kriterien für Beurteilung von Schüleraufsätzen. Mitteilung des Deutschen Germanistenverbandes 1972, 1, 8 – 9.
- LUKESCH, H.: Erziehungsstile, Pädagogische und psychologische Konzepte. Kohlhammer, Stuttgart 1975.
- LUKESCH, H., PERREZ, M., SCHNEEWIND, K.A.: Familiäre Sozialisation und Intervention. Huber, Bern 1980.
- MAGER, R.F.: Lernziele und Unterricht. Beltz, Weinheim 1965, 1977.
- MAGNUSSON, D.: Testtheorie. Deuticke, Wien 1969.
- MALGADI, R.G., BARCHER, P.R.: Some Information-Processing Models of Creative Writing. Journal of Educational Psychology 1977, 69, 512 – 518.
- MALGADI, R.G., BARCHER, P.R.: Psychological scaling of essay creativity: Effects of productivity and novelty. Journal of Educational Psychology 1979, 71, 717 – 725.

- MANDL, H.: Kognitive Entwicklungsläufe von Grundschulern. Empirische Befunde zum Schereneffekt. Oldenbourg, München 1975.
- MANDL, H.: Zur Psychologie der Textverarbeitung. Ansätze, Befunde, Probleme. Urban & Schwarzenberg, München 1981.
- MANDL, H. et al.: Lernen mit Texten. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1980, 12, 44 – 74.
- MARJORIBANKS, K.: Environment, social class and mental abilities. Journal of Educational Psychology 1972, 63, 103 – 109.
- MARSHALL, J. L., HAERTEL, E. H.: A single-administration reliability index for criterion-referenced tests: The mean split-half coefficient of agreement. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C. 1975.
- MASENDORF, F.: Empirische Lehrzielhierarchien als förderungsdiagnostisches Prinzip. In: KORNMANN, R., MEISTER, H., SCHLEE, J. (Hrsg.): Förderungsdiagnostik. Schindele, Heidelberg 1983, 107 – 112.
- MASENDORF, F., ROEDER, B., KRETSCHMANN, R.: Zur differentiellen Erfassung von Schulleistungsänderungen mit Hilfe der Konfigurationsfrequenzanalyse. Psychologie in Erziehung und Unterricht 1976, 23, 377 – 380.
- MAYER, R. E.: The Promise of Cognitive Psychology. Freeman, San Francisco 1981.
- MCKEACHIE, W. J.: Lipsticks as a determiner of first impression of personality: an experiment for the general psychology course. Journal of sociological Psychology 1952, 36, 214.
- MCNEIL, J. D., POPHAM, W. J.: The Assessment of Teacher Competence. In: TRAVERS, R. (Hrsg.): Second Handbook of Research on Teaching. Rand McNally, Chicago 1973.
- MEDLEY, D. M., MITZEL, H. E.: Measuring Classroom Behavior by Systematic Observation. In: GAGE, N. L. (Hrsg.): Handbook of Research on Teaching. Rand McNally, Chicago 1965.
- MEEHL, P. E.: Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. University of Minnesota Pr., Minneapolis 1954.
- MEES, U.: Einführung in die systematische Verhaltensbeobachtung. In: MEES, U., SELG, H. (Hrsg.): Verhaltensbeobachtung und Verhaltensmodifikation. Klett, Stuttgart 1977.
- MEIER, F.: Studentisches Arbeiten als Prozeß. Eine Anwendung der Zeitreihenanalyse in der Feldforschung. Hain, Meisenheim 1981.
- MEIS, R.: Diagnostischer Rechtschreibtest (DRT 4 – 5). Beltz, Weinheim 1969.
- MEIS, R.: Duisburger Vorschul- und Einschulungstest (DVET). Beltz, Weinheim 1973.
- MERCER, C. D.: Children and Adolescents with Learning Disabilities. Merrill & Howell, Columbus 1979.
- MERKELBACH, V.: Kritik des Aufsatzunterrichts. Diesterweg, Frankfurt/M. 1972.
- MERKENS, H.: Zum Problem der Konstruktion von lehrzielorientierten Tests. Schule und Psychologie 1972, 19, 139 – 140.
- MICHAEL, W. B.: Prediction. In: EBEL, R. L. (Hrsg.): Encyclopedia of educational research. MacMillan, London 1969, 982 – 993.
- MICHEL, L.: Allgemeine Grundlagen psychometrischer Tests. In: HEISS, R. (Hrsg.): Handbuch der Psychologie, Bd. 6: Psychologische Diagnostik. Hogrefe, Göttingen 1964, 19 – 70.
- MICHEL, L., MAI, N.: Entscheidungstheorie und Probleme der Diagnostik bei Cronbach & Gleser. Diagnostica 1968, 14, 99 – 121.
- MIETZEL, G.: Interpretation von Leistungen. Leske, Opladen 1982.
- MILLER, G. A., GALANTER, E., PRIBRAM, K. H.: Strategien des Handelns. Klett, Stuttgart 1973.

- MÖBUS, C., NAGL, W.: Messung, Analyse und Prognose von Veränderungen. In: BREDEKAMP, J., FEGER, H. (Hrsg.): Hypothesenprüfung (= Enzyklopädie der Psychologie, Themenbereich B, Serie I, Bd. 5). Hogrefe, Göttingen 1983, 239 – 470.
- MÖLLER, C.: Technik der Lernplanung. Beltz, Weinheim 1969.
- MOELLER, M.L.: Zur Psychodynamik des Prüfungswesens. Zeitschrift für Psychotherapie und Psychologie 1972, 22, 1 – 13.
- MOSER, H.: Didaktisches Planen und Handeln. Kösel, München 1978.
- MÜHLE, G., SCHELL, C. (Hrsg.): Kreativität und Schule. Piper, München 1969.
- MÜLLER, R.: Diagnostischer Rechtschreibtest für 2. Klassen (DRT 2). Beltz, Weinheim 1965, 1983<sup>2</sup>.
- MÜLLER, R.: Diagnostischer Rechtschreibtest für 3. Klassen (DRT 3). Beltz, Weinheim 1965, 1983<sup>2</sup>.
- NEUBAUER, W.F.: Selbstkonzept und Identität im Kindes- und Jugendalter. Reinhardt, München 1976.
- NEUMANN, J.: Die Erfassung von Denkleistungen mit Hilfe von Markoffketten. Zeitschrift für Psychologie 1965, 171, 421 – 430.
- NEWCOMB, T.: Sozialpsychologie. Hain, Meisenheim 1959.
- NICKEL, H.: Untersuchungen zur Bedeutung einer erhöhten Motivation für eine einheitliche Auffassung in der visuellen Wahrnehmung vierjähriger Kinder. Psychologische Rundschau 1968, 19, 9 – 17.
- NICKEL, H.: Entwicklungspsychologie des Kindes- und Jugendalters, 2 Bde. Huber, Bern 1972, 1982<sup>4</sup> (Bd. 1) und 1975, 1979<sup>3</sup> (Bd. 2).
- NICKEL, H.: Entwicklungsstand und Schulfähigkeit. Reinhardt, München 1976.
- NICKEL, H.: Schulfähigkeit, Schulreife. In: SCHIEFELE, H., KRAPP, A. (Hrsg.): Handlexikon der Pädagogischen Psychologie. Ehrenwirth, München 1981, 311 – 316 (a).
- NICKEL, H.: Schulreife und Schulversagen. Ein ökopyschologischer Erklärungsansatz und seine praktischen Konsequenzen. Psychologie in Erziehung und Unterricht 1981, 28, 19 – 37 (b).
- NICKEL, H.: Schuleingangsberatung auf der Grundlage eines ökopyschologischen Schulreifemodells. In: HELLER, K.A., NICKEL, H. (Hrsg.): Modelle und Fallstudien zur Schul- und Erziehungsberatung. Huber, Bern 1982, 81 – 88 (a).
- NICKEL, H.: Schulreife und Schulerfolg als ökologisches Problem. In: JANIG, H., LÖSCHENKOHL, E., SCHOFENEGGER, J., SÜSSENBACHER, G. (Hrsg.): Umweltpsychologie. Literas, Wien 1982, 32 – 37 (b).
- NICKEL, H., WIECZKOWSKI, W.: Einflüsse auf die Beurteilung von Schüleraufsätzen. Ergebnisse einer quasi-experimentellen Versuchsreihe. In: HELLER, K. (Hrsg.): Leistungsbeurteilung in der Schule. Quelle & Meyer, Heidelberg 1974, 271 – 308.
- NIEGEMANN, H.M., TREIBER, B.: Lehrstoffstrukturen, Kognitive Strukturen. In: TREIBER, B., WEINERT, F.E. (Hrsg.): Lehr-Lern-Forschung. Urban & Schwarzenberg, München 1982, 37 – 65.
- NIEMEYER, W.: Bremer Lautdiskriminationstest (BLDT). Bremer Hilfen für leserschreibschwache Kinder. Herbig, Bremen 1976.
- NIPKOW, K.E.: Leistungsprinzip und Lernverständnis. In: BECKMANN, H.K. (Hrsg.): Leistung in der Schule. Westermann, Braunschweig 1978, 7 – 32.
- NITKO, A.J.: Distinguishing the many varieties of criterion-referenced Tests. Review of Educational Research 1980, 50, 461 – 485.
- NUNNALLY, J.C., ATOR, N.A.: Educational measurement and evaluation. McGraw-Hill, New York 1964, 1972<sup>2</sup>.
- OLDENBÜRGER, H.-A.: Clusteranalyse. In: BREDEKAMP, J., FEGER, H. (Hrsg.): Strukturierung und Reduzierung von Daten (= Enzyklopädie der Psychologie, Themenbereich B, Serie I, Bd. 4). Hogrefe, Göttingen 1983, 390 – 439.

- OPP, K. D., SCHMIDT, P.: Einführung in die Mehrvariablenanalyse. Rowohlt, Reinbek 1976.
- OSBURN, H. G.: Item sampling for achievement testing. *Educational and Psychological Measurement* 1968, 28, 95 – 104.
- OSTERLAND, J.: Verhaltensbeurteilung von Schülern mit Hilfe von Ratingskalen. In: HELLER, K. (Hrsg.): *Handbuch der Bildungsberatung*, Bd. 3. Klett, Stuttgart 1976, 805 – 816.
- PANCHAUD, G., HENSLER, H., FALL, S.: La valeur objective des examens. *Etudes Pédagogiques, Annuaire de l'Instruction Publique en Suisse*. Payot, Lausanne 1969, 55 – 72.
- PAUSE, G.: Merkmale der Lehrerpersönlichkeit. In: INGENKAMP, K. (Hrsg.): *Handbuch der Unterrichtsforschung*, Teil II. Beltz, Weinheim 1970, 1353 – 1528.
- PAWLIK, K. (Hrsg.): *Diagnose der Diagnostik*. Klett, Stuttgart 1976.
- PAWLIK, K.: Modell- und Praxisdimensionen psychologischer Diagnostik. In: PAWLIK, K. (Hrsg.): *Diagnose der Diagnostik*. Klett, Stuttgart 1976, 19 – 43.
- PEDERSEN, D. M., SHINEDLING, M. M., JOHNSON, D. L.: Effects of sex of examiner and subject on children's quantitative test performance. *Journal of Personality and Social Psychology* 1968, 10, 251 – 254.
- PEISERT, H. G., DAHRENDORF, R.: Der vorzeitige Abgang vom Gymnasium. (Schriftenreihe des Kultusministeriums Baden-Württemberg «Bildung in neuer Sicht», Bd. 6). Neckar-Verlag, Stuttgart 1967.
- PELLEGRINO, J. W., GLASER, R.: Cognitive Correlates and Components in the Analysis of individual Differences. In: STERNBERG, R. J., DETTERMAN, D. K. (Hrsg.): *Human Intelligence*. Ablex Publish Company, Norwood 1979.
- PETERMANN, F.: *Veränderungsmessung*. Kohlhammer, Stuttgart 1978.
- PETERMANN, F., HEHL, F.-J. (Hrsg.): *Einzelfallanalyse*. Urban & Schwarzenberg, München 1979.
- PETRI, G.: Zur Problematik der Leistungsmessung: Analysen und Entwicklungsansätze (Nr. 2 der Arbeitsberichte der Reihe II des Zentrums für Schulversuche und Schulentwicklung). Graz 1974.
- PFISTNER, H.-J.: *Lehrer beurteilen Schüler. Beiträge zur Beurteilungspraxis und zur Theorie der Leistung*. Lang, Bern 1981.
- PIAGET, J., SZEMINSKA, A.: Die Entwicklung des Zahlbegriffs beim Kinde. Klett, Stuttgart 1959, 1969<sup>2</sup>.
- PIONTKOWSKI, U.: Interaktion und Kommunikation im Unterricht. In: TREIBER, B., WEINERT, F. E. (Hrsg.): *Lehr-Lern-Forschung*. Urban & Schwarzenberg, München 1982, 149 – 176.
- PLICKAT, H., WIECZERKOWSKI, W.: *Lernerfolg und Trainingsformen im Rechtschreibunterricht*. Klinkhardt, Bad Heilbrunn 1979.
- POKORNY, A. D., FRAZIER, S. H.: An Evaluation of Oral Examinations. *Journal of Medical Education* 1966, 41, 28 – 40.
- POPHAM, W. J.: Indices of adequacy for criterion-referenced test items. A Symposium Presentation at a Joint Session of the National Research Association, Minnesota 1970.
- POPHAM, W. J.: *Criterion-referenced measurement*. Prentice Hall, Englewood Cliffs, N.J. 1978.
- POPHAM, W. J.: *Modern educational Measurement*. Prentice Hall, Englewood Cliffs, N.J. 1981.
- POPHAM, W. J., HUSEK, T. R.: Implications of criterion-referenced Measurement. *Journal of Educational Measurement* 1969, 6, 1 – 9.
- PREISER, S.: Zur «Objektivität» mündlicher Prüfungen im Fach Psychologie. *Psychologische Rundschau* 1975, 26, 256 – 281.

- PREISER, S.: Kreativitätsforschung. Wissenschaftliche Buchgesellschaft, Darmstadt 1976.
- PROBST, H.: Strukturorientierte Diagnostik. Kritische Behindertenpädagogik in Theorie und Praxis. Jarick, Oberbiel 1979.
- PROBST, H.: Theorie der kognitiven Entwicklung in ihrer Bedeutung für eine Neuorientierung der Diagnostik. In: KASZTANTOW, U. (Hrsg.): Beiträge zur sonderpädagogischen Theorie und Praxis. Marhold, Berlin 1980, 35 – 47.
- PROBST, H.: Zur Diagnostik und Didaktik der Oberbegriffbildung. Jarick, Oberbiel 1981.
- PROBST, H.: Aufgaben und Verfahren qualitativer Diagnostik. In: Fachschaftsinitiative Sonderpädagogik Würzburg (Hrsg.): Diagnostik im Interesse der Betroffenen. Reichenberg-Lindflur 1982.
- PROBST, H.: Diagnostik? Wenn ja, welche? – Argumente und Beispiele für eine inhaltsbezogene pädagogische Diagnostik. In: KORNMANN, R., MEISTER, H., SCHLEE, J. (Hrsg.): Förderungsdiagnostik. Schindele, Heidelberg 1983, 113 – 124 (a).
- PROBST, H.: Tests zur Diagnostik spezifischer Lernvoraussetzungen. In: HORN, R. et al. (Hrsg.): Tests und Trends 3 (Jahrbuch der Pädagogischen Diagnostik). Beltz, Weinheim 1983, 77 – 105 (b).
- PROBST, H., KUTZER, R.: Strukturbezogene Aufgaben zur Prüfung mathematischer Einsichten. Beltz, Weinheim 1984.
- PUTZ-OSTERLOH, W.: Problemlöseprozesse und Intelligenzleistungen. Huber, Bern 1981.
- QUACK, L.: Zur Bedingungsanalyse der Schulleistung: Der Beitrag kognitiver und nicht-kognitiver Merkmale der Schülerpersönlichkeit. In: KLAUER, K.J., KORNDT, H.-J. (Hrsg.): Jahrbuch für empirische Erziehungswissenschaft 1979. Schwann, Düsseldorf 1979, 93 – 116.
- RADATZ, H.: Fehleranalysen im Mathematikunterricht. Vieweg, Braunschweig 1979.
- RASCH, G.: Probabilistic Models for Some Intelligence and Attainment Tests. Danmarks pædagogiske Institut, Copenhagen 1960.
- RENN, H.: Die Messung von Sozialisierungswirkungen. Oldenbourg, München 1973.
- RESNICK, L.B.: Hierarchies in Children's learning. A Symposium. Instructional Science 1973, 2, 311 – 362.
- RESNICK, L.B.: Task analysis in instructional design: Some cases from mathematics. In: KLAHR, D. (Hrsg.): Cognition and Instruction. Lawrence Erlbaum, Hillsdale, N.J. 1976.
- RESNICK, L.B.: Die Bedeutung von Erfindungen in der Entwicklung mathematischer Kompetenz. In: KLUWE, R.H., SPADA, H. (Hrsg.): Studien zur Denkentwicklung. Huber, Bern 1981, 361 – 411.
- REUCHLIN, M.: Testergebnisse und Zensuren der Klassenlehrer. In: INGENKAMP, K. (Hrsg.): Die Fragwürdigkeit der Zensurengebung. Beltz, Weinheim 1971, 164 – 167.
- REVENSTORF, D., KEESER, W.: Zeitreihenanalyse von Therapieverläufen. In: PETERMANN, F., HEHL, F.-J. (Hrsg.): Einzelfallanalyse. Urban & Schwarzenberg, München 1979, 183 – 228.
- RHEINBERG, F.: Situative Determinanten der Beziehung zwischen Leistungsmotiv und Schul- und Studienleistung. In: SCHMALT, H.-D., MEYER, W.-U. (Hrsg.): Leistungsmotivation und Verhalten. Klett, Stuttgart 1976, 249 – 264.
- RHEINBERG, F.: Leistungsbewertung und Lernmotivation. Hogrefe, Göttingen 1980.
- RIEDER, O.: Allgemeiner Schulleistungstest für 2. Klassen (AST 2). Beltz, Weinheim 1971.
- ROBECK, M.C., WILSON, J.A.R.: Psychology of Reading: Foundations of Instruction. Wiley & Sons, New York 1974.
- ROEDER, P.M.: Einleitung zum Themenschwerpunkt «Lehr-Lern-Forschung». Zeitschrift für Pädagogik 1982, 28, 355 – 364.

- ROEDER, P.M., TREUMANN, K.: Dimensionen der Schulleistung, 2 Bde. Klett, Stuttgart 1974.
- ROSEMANN, B.: Zur Problematik der Klassifikation von Schultests. In: HELLER, K. (Hrsg.): Leistungsbeurteilung in der Schule. Quelle & Meyer, Heidelberg 1974, 158 – 166.
- ROSEMANN, B.: Prognosemodelle in der Schullaufbahnberatung. Reinhardt, München 1978.
- ROSENSHINE, B., FURST, M.: The use of direct observation to study teaching. In: TRAVERS, R.M.W. (Hrsg.): Second handbook of research on Teaching. Rand McNally, Chicago 1973, 122 – 138.
- ROST, J.: Diagnostik des Lernzuwachses. IPN-Arbeitsberichte, Nr.26. Universität Kiel 1977.
- ROST, J., SPADA, H.: Probabilistische Testtheorie. In: KLAUER, K.J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd. 1. Schwann, Düsseldorf 1978, 59 – 97.
- ROTH, H. et al.: Frankfurter Schulreifeftest (FST). Beltz, Weinheim 1968<sup>5</sup>.
- ROUDABUSH, G.E.: Item Selection for Criterion-Referenced Tests. Paper presented at the American Educational Research Association, New Orleans 1973.
- RUBINSTEIN, S.L.: Sein und Bewußtsein. Akademie Verlag, Berlin 1962.
- RÜDIGER, D.: Der prozeßdiagnostische Ansatz mit einem Beispiel curriculärer Prozeßdiagnostik im Erstleseunterricht. In: BOLSCO, D., SCHWARZER, C. (Hrsg.): Beurteilen in der Grundschule. Urban & Schwarzenberg, München 1979, 162 – 184.
- RÜDIGER, D.: Prozeßdiagnostik. In: SCHIEFELE, H., KRAPP, A. (Hrsg.): Handlexikon zur Pädagogischen Psychologie. Ehrenwirth, München 1981, 289 – 293.
- RÜDIGER, D., KORMANN, A., PEETZ, H.: Schuleintritt und Schulfähigkeit. Reinhardt, München 1976.
- SADER, M., KEIL, W.: Bedingungskonstanz in der psychologischen Diagnostik. Archiv für die Gesamtpsychologie 1966, 118, 297 – 308.
- SAMSTAG, K., SANDER, A., SCHMIDT, R.: Diagnostischer Rechentest für 3. Klassen (DRE 3). Beltz, Weinheim 1971.
- SAMSTAG, K.: Informationen zum Lernen im Vorschulalter. Klinkhardt, Bad Heilbrunn 1971.
- SANDER, E.: Lernstörungen. Kohlhammer, Stuttgart 1981.
- SANG, F.: Elternreaktionen auf Schulleistung. Klett-Cotta, Stuttgart 1977.
- SCHÄFER, H.: Bildwortserie zur Lautagnosieprüfung und zur Schulung des phonematischen Gehörs. Eigenverlag, Gießen 1975.
- SCHALLBERGER, U.: Psychodiagnostik und psychologische Theorieentwicklung. In: PULVER, U., LANG, A., SCHMID, F.W. (Hrsg.): Ist Psychodiagnostik verantwortbar? Huber, Bern 1978, 81 – 86.
- SCHALLER, K.: Leistung und Selbstverwirklichung. Neue Deutsche Schule, Essen 1979.
- SCHEEER-NEUMANN, G.: Intervention bei Lese-Rechtschreibschwäche. Kamp, Bochum 1976, 1979<sup>2</sup>.
- SCHENK-DANZINGER, L.: Erscheinungsformen der Legasthenie. Handbuch der Legasthenie. Beltz, Weinheim 1968.
- SCHENK-DANZINGER, L.: Schuleintrittsalter, Schulfähigkeit und Lesereife. Deutscher Bildungsrat, Gutachten und Studien der Bildungskommission, Bd. 7. Klett, Stuttgart 1969.
- SCHMACK, E.: Nordrhein-Westfalens neues Grundschulzeugnis in der Bewährung – Ergebnisse einer Stichprobenauswertung. In: BOLSCO, D., BURK, K.H., HAARMANN, D. (Hrsg.): Grundschule ohne Noten. Arbeitskreis Grundschule, Frankfurt/M. 1979, 130 – 147.
- SCHMALOHR, E.: Zur akustischen Durchgliederungsfähigkeit als Voraussetzung des Lesenlernens bei 4- bis 6jährigen Kindern. Schule und Psychologie 1968, 15, 295 – 303.

- SCHMIDT, H.J.: Grundschatulzeugnisse in Niedersachsen. Bericht über eine Untersuchung. Neubauer, Lüneburg 1980.
- SCHMIDT, H.P.: Zur Problematik des Rechtschreibversagens bei Lernbehinderten. Zeitschrift für Heilpädagogik 1980, 31, 186 – 193.
- SCHMIDT-DENTER, U.: Sozialverhalten von Kindern. Universität Düsseldorf 1980.
- SCHMÖLDERS, G.: Leistungsprinzip und Qualität des Lebens. In: GEHLEN, A. et al.: Sinn und Unsinn des Leistungsprinzips. DTV, München 1975<sup>3</sup>, 20 – 30.
- SCHNEEWIND, K. A., LUKESCH, H. (Hrsg.): Familiäre Sozialisation. Klett, Stuttgart 1978.
- SCHNEIDER, K., KREUZ, A.: Die Effekte unterschiedlicher Anstrengung auf die Mengen- und Güteleistung bei einer einfachen und schweren Zahlensymbolaufgabe. Psychologie und Praxis 1979, 23, 34 – 42.
- SCHNEIDER, W.: Kausalmodelle zur Beschreibung und Erklärung von Rechtschreibleistungen in der Grundschule – Eine Sekundäranalyse. Zeitschrift für Empirische Pädagogik 1982, 6, 123 – 143.
- SCHNITZER, A. (Hrsg.): Schwerpunkt: Leistung in der Schule. Oldenbourg, München 1981.
- SCHNOTZ, W.: Schätzung von Aufgabenschwierigkeiten durch Lehrer. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1971, 3, 106 – 120.
- SCHNOTZ, W.: Lerniagnose als Handlungsanalyse. Beltz, Weinheim 1979.
- SCHNOTZ, W.: Zur Diagnostik von Verstehensleistungen. In: KORNMANN, R., MEISTER, H., SCHLEE, J. (Hrsg.): Förderungsdiagnostik. Schindele, Heidelberg 1983, 137 – 146.
- SCHORB, A.O., SCHMIDBAUER, M.: Aufstiegsschulen im sozialen Wettbewerb. Klett, Stuttgart 1969.
- SCHOTT, F., KRETSCHMER, I.: Konstruktion lehrzielvalider Testaufgaben aufgrund einer normierten Lehrstoffanalyse. In: KLAUER, K.J. et al. (Hrsg.): Lehrzielorientierte Leistungsmessung. Schwann, Düsseldorf 1977, 26 – 64.
- SCHOTT, F., NEEB, K.-E., WIEBERG, H.J.W.: Lehrstoffanalyse und Unterrichtsplanung. Vieweg, Braunschweig 1981.
- SCHRAND, H., PORTMANN, R., STARK, G.: Englisch-Einstufungstest 6 +. Beltz, Weinheim 1973.
- SCHRECKENBERG, W.: «Guter» Unterricht – «schlechter» Unterricht. Zur Theorie und Praxis der Unterrichtsbeurteilung. Schwann, Düsseldorf 1980, 1984<sup>2</sup>.
- SCHRECKENBERG, W.: Vom «guten» zum «besseren» Lehrer. Über Eignung und Leistung von Lehrern. Schwann, Düsseldorf 1982.
- SCHRÖTER, G.: Die ungerechte Aufsatzzensur. Kamp, Bochum 1971, 1974<sup>4</sup>.
- SCHRÖTER, G.: Von der Theorie zur Praxis der Aufsatzbewertung (Sekundarstufe I). Praxis Deutsch 1975, 2, 49 – 51.
- SCHRUMPF, B.: Die Analyse von Ausdrucksfehlern – Ein Weg zur Bestimmung von Übungsschwerpunkten bei der Arbeit am schriftlichen Ausdruck. Deutschunterricht 1975, 28, 267 – 289.
- SCHULZ, W., TESCHNER, W.P., VOIGT, J.: Verhalten im Unterricht. Seine Erfahrung durch Beobachtungsverfahren. In: INGENKAMP, K. (Hrsg.): Handbuch der Unterrichtsforschung, Teil I. Beltz, Weinheim 1970, 633 – 852.
- SCHWARTZ, E.: Eingangsbericht statt Zensuren-Zeugnis! Grundschule 1978, 20, 100 – 104.
- SCHWARZ, E.: Experimentelle und quasi-experimentelle Anordnungen in der Unterrichtsforschung. In: INGENKAMP, K., PAREY, E. (Hrsg.): Handbuch der Unterrichtsforschung, Teil I. Beltz, Weinheim 1970, 445 – 632.
- SCHWARZER, C., SCHWARZER, R.: Praxis der Schülerbeurteilung. Kösel, München 1977.
- SCHWARZER, C.: Gestörte Lernprozesse. Urban & Schwarzenberg, München 1980.
- SCHWARZER, R.: Sequentielle Prädiktion des Schulerfolgs. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1979, 11, 170 – 180.

- SCHWARZER, R.: Fallstudie zur Evaluation der Neuregelung von Zeugnissen in der Grundschule. Düsseldorf 1981 (a).
- SCHWARZER, R.: Streß, Angst und Hilflosigkeit. Die Bedeutung von Kognition und Emotionen bei der Regulation von Belastungssituationen. Kohlhammer, Stuttgart 1981 (b).
- SCHWARZER, R., STEINHAGEN, K. (Hrsg.): Adaptiver Unterricht. Kösel, München 1975.
- SCRIVEN, M.: The methodology of evaluation. AERA Monograph Series on Curriculum Evaluation 1967, 1, 39 – 83.
- SEITZ, W.: Über den Zusammenhang von Persönlichkeitseigenarten, Schulnoten und HAWIK-Leistungen bei Volksschülern. Psychologische Beiträge 1970, 13, 579 – 602.
- SEITZ, W.: Über die Beziehung von Persönlichkeitsmerkmalen zu Schul- und Intelligenztestleistungen bei Volksschülern. Zeitschrift für experimentelle und angewandte Psychologie 1971, 18, 307 – 336.
- SEITZ, W., LÖSER, G.: Über die Beziehung von Persönlichkeitsmerkmalen zu Schul- und Intelligenzleistungen bei Volksschülern. Zeitschrift für experimentelle und angewandte Psychologie 1969, 16, 671 – 679.
- SHAVELSON, R.J., BLOCK, J.H., RAVITCH, M.M.: Criterion-referenced testing: comments on reliability. Journal of Educational Measurement 1972, 9, 133 – 137.
- SHOEMAKER, D.M., OSBURN, H.G.: Computer-aided item sampling for achievement testing. Educational and Psychological Measurement 1969, 29, 165 – 172.
- SIGNER, M.: Hörtraining bei auditiv differenzierungsschwachen Kindern. Haupt, Bern 1979.
- SIMONS, H., WEINERT, F.E., AHRENS, H.J.: Untersuchungen zur differential-psychologischen Analyse von Rechenleistungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1975, 7, 153 – 169.
- SOMMERLA, G.: Zur Praxis der Leistungsfeststellung und -bewertung in der Schule. Westermanns Pädagogische Beiträge 1976, 28, 450 – 461.
- SOMMERLA, G.: Leistungsbeurteilung in der Primarstufe. Ehrenwirth Grundschulmagazin 1979, 6, 5 – 8.
- SPACHE, G.D.: Good Reading for Poor Reader. Garrard Publishing Company, Champaign, Ill. 1974.
- STARCH, D., ELLIOT, E.C.: Die Verlässlichkeit der Zensierung von Mathematikarbeiten. In: INGENKAMP, K. (Hrsg.): Die Fragwürdigkeit der Zensurenggebung. Beltz, Weinheim 1971, 69 – 77.
- STARK, G. et al.: Theorie und Praxis der Leistungsmessung in Gesamtschulen. DIPF, Frankfurt/M. 1973.
- STEGMÜLLER, W.: Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, Bd. I: Erklärung und Begründung. Springer, Berlin 1969.
- STENE, J.: Einführung in RASCHS Theorie der psychologischen Messung. In: FISCHER, G. (Hrsg.): Psychologische Testtheorie. Huber, Bern 1968, 229 – 268.
- STELZL, I.: Was bringt das Rasch-Modell in der Praxis? Psychologische Beiträge 1972, 15, 298 – 310.
- STELZL, I.: Fehler und Fallen der Statistik. Huber, Bern 1982.
- STRAKA, G.A.: Zielerreichendes Lernen. In: SCHIEFELE, H., KRAPP, A. (Hrsg.): Handlexikon zur Pädagogischen Psychologie. Ehrenwirth, München 1981, 421 – 426.
- STRASSER, S.: Pädagogische Gesamtheorie als praktische Wissenschaft. Zeitschrift für Pädagogik 1972, 18, 659 – 684.
- SUBKOWIAK, M.J.: Estimating Reliability from a Single Administration of a Criterion-Referenced Test. Journal of Educational Measurement 1976, 13, 265 – 276.
- SÜLLWOLD, F.: Schultests. In: HEISS, R. (Hrsg.): Psychologische Diagnostik (= Handbuch der Psychologie, Bd. 6). Hogrefe, Göttingen 1964, 352 – 384.

- SÜLLWOLD, F.: Pädagogische Diagnostik. In: GROFFMANN, K. J., MICHEL, L. (Hrsg.): Intelligenz- und Leistungsdiagnostik (= Enzyklopädie der Psychologie, Themenbereich B, Serie II, Bd.2). Hogrefe, Göttingen 1983, 307 – 386.
- SWAMINATHAN, H. et al.: Reliability of Criterion-Referenced Tests: A Decision-Theoretic Formulation. *Journal of Educational Measurement* 1974, 11, 263 – 267.
- TAUSCH, R., TAUSCH, A.: Erziehungspsychologie. Hogrefe, Göttingen 1977<sup>8</sup>.
- TENT, L.: Die Auslese von Schülern für weiterführende Schulen. Hogrefe, Göttingen 1969.
- TENT, L., FINGERHUT, W., LANGFELDT, H.P.: Quellen des Lehrerurteils. Beltz, Weinheim 1976.
- TEUSCH, M.: Das Leistungsprinzip in der Regelschule und den Alternativschulen. Unveröffentl. Staatsexamensarbeit, Universität Köln 1983.
- TEWES, U. (Hrsg.): Hamburg-Wechsler-Intelligenztest für Kinder, Revision (HAWIK-R). Huber, Bern 1983.
- THOMAE, H.: Beobachtung und Beurteilung von Kindern und Jugendlichen. Karger, Basel 1954, 1980<sup>13</sup>.
- THORNDIKE, E.L.: A constant error in psychological ratings. *Journal of applied Psychology* 1920, 4, 25 – 29.
- THORNDIKE, R.L.: Concepts of underachievement. Teachers College Press, Columbia University, New York 1963.
- THORNTON, G.R.: The effect of wearing glasses upon judgements of personality traits of persons seen briefly. *Journal of applied Psychology* 1944, 28, 203 – 207.
- TIEDEMANN, J.: Leistungsversagen in der Schule. Goldmann, München 1977.
- TILMANN, H.: Lehrerbeurteilung. Quelle & Meyer, Heidelberg 1981.
- TISMER, K.-G.: Verhaltensbeobachtung bei Kindern und Jugendlichen. In: HELLER, K. (Hrsg.): *Handbuch der Bildungsberatung*, Bd.3. Klett, Stuttgart 1976, 817 – 836.
- TREIBER, B.: Qualifizierung und Chancenausgleich in Schulklassen. Teil I: Theorien, Methoden, Ergebnisse. Lang, Frankfurt 1980.
- TREIBER, B., PETERMANN, F.: Probleme der Unterrichtsdifferenzierung aus der Sicht des ATI-Forschungsprogramms. *Zeitschrift für Pädagogik* 1976, 22, 526 – 546.
- TREIBER, B., WEINERT, F.E. (Hrsg.): *Lehr-Lern-Forschung*. Urban & Schwarzenberg, München 1982.
- TROST, G.: Vorhersage des Studienerfolgs. Westermann, Braunschweig 1975.
- TRUDEWIND, C.: Häusliche Umwelt und Motiventwicklung. Hogrefe, Göttingen 1974.
- TSCHERNER, K.: Schulleistungsentwicklungen in Abhängigkeit von Beurteilungsprozessen. Athenäum, Königstein/Ts. 1979.
- TUCKER, L.R.: Implications of Factor Analysis of Three-Way Matrices in Measuring Change. In: HARRIS, CH. (Hrsg.): *Problems in Measuring Change*. University of Wisconsin Press, Madison, Wisconsin 1963.
- TUCKER, L.R., DAMARIAN, F., MESSICK, S.: A base-free measure of change. *Psychometrika* 1966, 31, 457 – 473.
- TYLER, R.W.: *Basic principles of curriculum and instruction*. University of Chicago Press, Chicago 1950. – Dt. *Curriculum und Unterricht*. Schwann, Düsseldorf 1973.
- ÜBERLA, K.: *Faktorenanalyse*. Springer, Berlin 1971<sup>2</sup>.
- ULSHÖFER, R.: Wie beurteilen Sie diesen Reifeprüfungsaufsatz? *Der Deutschunterricht* 1948/1949, 1, 95 – 98.
- VALETTE, R.M.: *Modern Language Testing*. Harcourt, New York 1977<sup>2</sup>.
- VOLPERT, W.: *Handlungsstrukturanalyse als Beitrag zur Qualifikationsforschung*. Pahl-Rugenstein, Köln 1974.
- WAHL, D.: *Erwartungswidrige Schulleistungen*. Beltz, Weinheim 1975.
- WAHL, D.: Implizite Theorien. In: SCHIEFELE, H., KRAPP, A. (Hrsg.): *Handlexikon zur Pädagogischen Psychologie*. Ehrenwirth, München 1981, 172 – 176.

- WALTER, H.: Neue Wege zum optimalen Unterricht. Beobachtung und Beurteilung von Schüler- und Lehrerverhalten. Ehrenwirth, München 1973.
- WALTER, H.: Angst bei Schülern. Reinhardt, München 1977.
- WANG, M.C.: Approaches to the validation of learning hierarchies. Western Regional Conference on Testing Problems (Proceedings). Educational Testing Service, Princeton, N.J. 1969, 14 – 38.
- WARR, P.B., KNAPPER, C.: The perception of people and events. Wiley & Sons, London 1968.
- WASNA, M.: Motivation, Intelligenz und Lernerfolg. Kösel, München 1972.
- WEBER, A.: Aufsatzlehre und Aufsatzbeurteilung. In: BECK, O., PAYRHUBER, F.-J. (Hrsg.): Aufsatzunterricht heute. Herder, Freiburg 1977, 137 – 178.
- WECHSLER, D.: Die Messung der Intelligenz Erwachsener. Textband zum Hamburg-Wechsler-Intelligenztest für Erwachsene (HAWIE). Huber, Bern 1956.
- WEEDE, E.: Zur Pfadanalyse. Neuere Entwicklung, Verbesserungen, Ergänzungen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 1972, 24, 101 – 117.
- WEINER, B.: Die Wirkung von Erfolg und Mißerfolg auf die Leistung. Klett, Stuttgart 1975.
- WEINERT, F.E.: Schule und Beruf als Sozialisationsbedingung. In: GRAUMANN, C.F. (Hrsg.): Handbuch der Psychologie. Bd. 7/2: Sozialpsychologie. Hogrefe, Göttingen 1972, 825 – 885.
- WEINERT, F.E.: Fähigkeits- und Kenntnisunterschiede zwischen Schülern. In: WEINERT, F.E. et al. (Hrsg.): Funk-Kolleg Pädagogische Psychologie, Bd.2. Fischer, Frankfurt/M. 1974, 763 – 793 (a).
- WEINERT, F.E.: Instruktion als Optimierung von Lernprozessen, Teil I: Lehrmethoden. In: WEINERT, F.E. et al. (Hrsg.): Funk-Kolleg Pädagogische Psychologie, Bd.2. Fischer, Frankfurt/M. 1974, 797 – 826 (b).
- WEINERT, F.E., PETERMANN, F.: Erwartungswidrige Schülerleistungen oder unterschiedlich determinierte Schulleistungen. In: HECKHAUSEN, H. (Hrsg.): Fähigkeit und Motivation in erwartungswidriger Schulleistung. Hogrefe, Göttingen 1980, 19 – 52.
- WEINERT, F.E., SCHNEIDER, W., TREIBER, B.: Bedingungsanalyse von (erwartungswidrigen) Schulleistungen. In: HECKHAUSEN, H. (Hrsg.): Fähigkeit und Motivation in erwartungswidriger Schulleistung. Hogrefe, Göttingen 1980, 106 – 127.
- WEISS, R.: Über die Zuverlässigkeit der Ziffernbenotung bei Aufsätzen. Schule und Psychologie 1965, 12, 252 – 269 (a).
- WEISS, R.: Zensur und Zeugnis. Haslinger, Linz 1965 (b).
- WEISS, R.: Über die Zuverlässigkeit der Ziffernbenotung bei Rechenarbeiten. Schule und Psychologie 1966, 13, 144 – 151 und 198 – 207.
- WEISS, R.: Vor- und Nachteile der Leistungsbeurteilung durch Ziffernnoten. Schule und Psychologie 1969, 16, 198 – 207.
- WENDELER, J.: Eine Aufgabenanalyse anhand des Testmodells von RASCH. Archiv für die gesamte Psychologie 1968, 120, 218 – 230.
- WENDELER, J.: Standardarbeiten. Beltz, Weinheim 1969.
- WENDELER, J.: Bemühungen um die Vereinheitlichung der Aufsatzbeurteilung. In: HELLER, K. (Hrsg.): Leistungsbeurteilung in der Schule. Quelle & Meyer, Heidelberg 1974, 309 – 323 (a).
- WENDELER, J.: Satzlehre 6+. Beltz, Weinheim 1974 (b).
- WENDELER, J.: Wortlehre 4. Beltz, Weinheim 1978.
- WENDELER, J.: Förderungsdiagnostik bei Schulleistungsschwächen in der Grundschule. Psychologie in Erziehung und Unterricht 1981, 28, 293 – 305.
- WIECZERKOWSKI, W., NICKEL, H., ROSENBERG, L.: Einige Bedingungen der unterschiedlichen Bewertung von Schüleraufsätzen. Psychologische Rundschau 1968, 11, 280 – 295.

- WIECZKOWSKI, W., NICKEL, H.: Einflüsse auf die Beurteilung von Schüleraufsätzen. In: HELLER, K. (Hrsg.): Leistungsbeurteilung in der Schule. Quelle & Meyer, Heidelberg 1974, 271 – 308.
- WIECZKOWSKI, W., OEVESTE, H. ZUR: Zuordnungs- und Entscheidungsstrategien. In: KLAUER, K.J. (Hrsg.): Handbuch der Pädagogischen Diagnostik, Bd.4. Schwann, Düsseldorf 1978, 919 – 951.
- WIEDL, K. H.: Die Fehlerlatenz als alternative Form der Leistungsbewertung bei Lernstprozeduren. Psychologische Beiträge 1978, 20, 440 – 455.
- WIEDL, K. H.: Lerntests: nur Forschungsmittel und Forschungsgegenstand? Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1984, 16, 245 – 281.
- WIEDL, K. H., BETHGE, H.-J.: Der CPM-PV-Test. Materialien zur Entwicklung und Anwendung einer dynamischen Version des farbigen Matrizen-tests von RAVEN. Psychologische Forschungsberichte aus dem Fachbereich 8 der Universität Osnabrück, Bd. 30. Osnabrück 1983.
- WIEDL, K. H., BETHGE, H.-J.: Zum gegenwärtigen Stand der Entwicklung dynamischer Testverfahren in der Intelligenzdiagnostik. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1984, 16, im Druck.
- WIGGINS, J.S.: Personality and prediction: Principles of personality assessment. Addison & Wesley, Reading, Mass. 1973.
- WILL, H.: Zur Bedeutung der Diagnosen leichter frühkindlicher Hirnschädigungen für die pädagogische und sonderpädagogische Praxis. Psychologie in Erziehung und Unterricht 1982, 29, 279 – 291.
- WIMMER, H. et al.: Die Bedeutung soziokultureller, sprachlicher und kognitiver Merkmale für die Schulleistung von Schulanfängern. Psychologie in Erziehung und Unterricht 1975, 22, 199 – 208.
- WINKELMANN, W.: Testbatterie zur Entwicklung kognitiver Operationen (TEKO). Westermann, Braunschweig 1975.
- WOLF, B.: Zum Einfluß der häuslichen Lernumwelt. In: ROST, D. H. (Hrsg.): Entwicklungspsychologie für die Grundschule. Klinkhardt, Bad Heilbrunn 1980, 172 – 186.
- WOLF, W.: Beurteilungsmodelle bei Klassenarbeiten. Bestandsaufnahme schulischer Praxis in der Sekundarstufe I. Zeitschrift für Pädagogik 1984, 30, 115 – 130.
- WOOD, R., SKURNIK, L.S.: Item Banking. National Foundation for Educational Research in England and Wales, 1969.
- WRIGHT, B. D.: Sample-free test calibration and person measurement. Invitational Conference on testing problems, Princeton 1967.
- WYGOTSKI, L.S.: Denken und Sprechen. Akademie Verlag, Berlin 1964/Fischer Taschenbuchverlag, Frankfurt/M. 1977.
- YSSELDIKE, J.E., SALVIA, J.: Diagnostic prescriptive teaching: Two models. Exceptional Children 1974, 41, 181 – 185.
- ZELLER, W.: Der erste Gestaltwandel des Kindes. Barth, Leipzig 1936/Hogrefe, Göttingen 1964<sup>2</sup>.
- ZIEGENSPECK, J.: Zensur und Zeugnis. Schroedel, Hannover 1973, 1976<sup>3</sup>.
- ZIELINSKI, W.: Beurteilung von Schülerleistungen. In: Funkkolleg «Pädagogische Psychologie» (Studienbegleitbrief 12). Beltz, Weinheim 1973.
- ZIELINSKI, W.: Die Beurteilung von Schülerleistungen. In: WEINERT, F.E. et al. (Hrsg.): Funkkolleg Pädagogische Psychologie, Bd.2. Fischer, Frankfurt/M. 1974, 879 – 900.
- ZIELINSKI, W.: Lernschwierigkeiten. Verursachungsbedingungen, Diagnose, Behandlungsansätze. Kohlhammer, Stuttgart 1980.
- ZIMMERMANN, K.W.: Über die Beziehung zwischen Schulnoten von lernbehinderten Sonderschülern. Zeitschrift für Heilpädagogik 1968, 29, 465 – 471.

## 8. Autorenverzeichnis

BARTNITZKY, HORST, Dipl.-Päd., Schulrat in Düsseldorf, Stellv. Vorsitzender des «Arbeitskreises Grundschule e. V. Frankfurt»

*Anschrift:* Heinrich Albrod-Str. 54, D-4100 Duisburg 28

BIRKEL, PETER, Dr.phil., Dipl.-Psych., Akad. Oberrat an der Pädagogischen Hochschule Weingarten

*Anschrift:* Turmstr. 8, D-7981 Unterankenreute

BÜSCHER, PETER, Dipl.-Psych., Oberstudienrat an der Wirtschaftsschule für Hörgeschädigte in Heidelberg

*Anschrift:* Fichtenstr. 20, D-6948 Wald-Michelbach

CHRISTIANI, REINHOLD, Dipl.-Päd., Ministerialrat im Kultusministerium Nordrhein-Westfalen in Düsseldorf (Grundschulreferent)

*Anschrift:* Friedrich Wilhelm-Str. 24, D-4100 Duisburg 1

EBERLE, GERHARD, Dr.phil., Dipl.-Psych., Professor am Fachbereich Sonderpädagogik der Pädagogischen Hochschule Heidelberg

*Anschrift:* Bahnhofstr. 11, D-7531 Kieselbronn

FINGERHUT, WALTER, Dr.phil., Dipl.-Psych., Abteilungsleiter an der DEKRA-Akademie Wart in Stuttgart

*Anschrift:* Bächlenweg 4, D-7270 Nagold

HELLER, KURT A., Dr.phil., Dipl.-Psych., Professor der Psychologie (Arbeitsbereich Pädagogische Psychologie und Psychologische Diagnostik) an der Ludwig-Maximilians-Universität München

*Anschrift:* Hauptstr. 6, D-8031 Wörthsee

HORN, RALF, Dipl.-Psych., Leiter der Testabteilung im Beltz-Verlag Weinheim

*Anschrift:* Daimlerstr. 49, D-6840 Lampertheim

KLINGEN, LEO, Dr.rer.nat., Oberstudiendirektor und Leiter des Helmholtz-Gymnasiums Bonn

*Anschrift:* Billrothstr. 2, D-5300 Bonn 1

KORMANN, ADAM, Dr.phil., Dipl.-Psych., Staatl. Schulpsychologe beim Ministerialbeauftragten für Gymnasien im Regierungsbezirk Niederbayern

*Anschrift:* Wilhelm-Hauff-Str. 20, D-8300 Landshut

KORNMANN, REIMER, Dr.phil., Dipl.-Psych., Professor für Diagnostik der Lernbehinderten an der Pädagogischen Hochschule Heidelberg

*Anschrift:* Im Schulzengarten 4, D-6903 Neckargemünd

KRAPP, ANDREAS, Dr.phil., Dipl.-Psych., Professor für Erziehungswissenschaft und Pädagogische Psychologie an der Hochschule der Bundeswehr München

*Anschrift:* Olmützerweg 1, D-8045 Ismaning

LANGFELDT, HANS-PETER, Dr.phil., Dipl.-Psych., Akademischer Rat am Seminar für Psychologie der Pädagogischen Fakultät an der Universität Bonn

- und Privatdozent am Fachbereich Psychologie der Universität Marburg  
*Anschrift:* Küdinghovenerstr. 113, D – 5300 Bonn 3
- LANGHORST, ERICH, Dr.phil., Dipl.-Psych., Professor für Psychologie an der  
 Pädagogischen Fakultät der Universität Bonn  
*Anschrift:* Luisenstr. 2, D – 5300 Bonn 1
- NICKEL, HORST, Dr.phil., Dipl.-Psych., Professor der Psychologie (Entwicklungs- und Erziehungspsychologie) an der Universität Düsseldorf  
*Anschrift:* Berliner Str. 25, D – 5309 Meckenheim
- ROSEMANN, BERNHARD, Dr.phil., Dipl.-Psych., Professor für Pädagogische Psychologie an der Ruhr-Universität Bochum  
*Anschrift:* Mauerham 97, D – 8221 Taching a. See
- SCHRECKENBERG, WILHELM, Dr.phil., Oberstudiendirektor, Leiter des Bezirksseminars für das Lehramt am Gymnasium/Sekundarstufe II in Düsseldorf  
*Anschrift:* Redinghovenstr. 9, D – 4000 Düsseldorf
- SCHÜMANN, MICHAEL, Dipl.-Psych., Wissenschaftlicher Angestellter am Fachbereich Psychologie der Universität Hamburg  
*Anschrift:* Grindelallee 186, D – 2000 Hamburg 13
- WENDELER, JÜRGEN, Dr.rer.nat., Dipl.-Psych., Professor für Pädagogische Psychologie (mit besonderer Berücksichtigung der Heil- und Sonderpädagogik) an der Justus-Liebig-Universität Gießen  
*Anschrift:* August-Bebel-Str. 22, D – 6072 Dreieich
- WIECZERKOWSKI, WILHELM, Fil. Dr., Lic.psych., Professor der Psychologie (Entwicklungs- und Pädagogische Psychologie) an der Universität Hamburg  
*Anschrift:* Gorch-Fock-Weg 1, D – 2000 Hamburg-Norderstedt

## 9. Personenregister

- Aebli, H. 275, 317  
Ahrens, G. 257, 317  
Ahrens, H.J. 2  
Algina, J. 116, 324, 337  
Allinger, U. 38, 303, 305, 317  
Amidon 265  
Ammer 218  
Amthauer, R. 41  
Anastasi, A. 65, 317  
Anger, H. 154, 317  
Angermaier, M. 287, 317  
APA 113, 317  
Arbinger, R. 62, 317  
Asch, S.E. 230, 317  
Aschersleben, K. 84, 240, 317  
Atkinson, J.W. 23, 57f., 317  
Ator, N.A. 332  
Auckenthaler, A. 230, 317  
Augst, G. 287, 317  
Aurin, K. 303, 317  
Ammer, C. 317
- Baltes, P.B. 119f., 318  
Bandura, A. 216  
Barcher, P.R. 257f., 330  
Bargmann, R. 317  
Barr 265  
Bartnitzky, H. 37f., 273, 295, 298, 318  
Baur, R. 58, 318  
Beck, O. 36, 159, 247f., 250f., 259, 318, 339  
Beckmann, H.K. 16f., 318, 332  
Beetz, E. 231, 326  
Bellack 265  
Benner, D. 292, 318  
Bereiter, C. 117, 119, 121, 318  
Berk, R.A. 116, 318  
Berliner, D.C. 53, 269, 318, 322  
Bethäuser, H. 303, 307, 318  
Bethge, H.J. 119f., 203, 340  
Betz, D. 233, 318  
Biäsch, H. 230, 318  
Bibl, W. 162, 318  
Biggs, J.B. 252, 259  
Biglmaier, F. 288, 318
- Binet, A. 287  
Birkel, P. 35, 205f., 231, 233, 236, 318  
Block, J.H. 116, 318, 336  
Bloom, B.S. 29, 32, 35, 141, 143, 146f., 152ff., 163f., 318, 329  
Bock, R.D. 119f., 318  
Bolscho, D. 307, 319, 324, 335  
Bormuth, J.R. 113f., 319  
Bramble, W. 109, 327  
Brandstädter, J. 58, 319  
Bredenkamp, J. 124, 329, 332  
Breuer, H. 285f., 319  
Brickenkamp, R. 154, 161, 319  
Brombach, E. 238, 319  
Bronfenbrenner, U. 277, 319  
Brophy, J.E. 56, 211, 319  
Brown, A.L. 133, 319  
Brown, J.S. 131, 319  
Bruner, J.S. 132  
Budoff 200  
Büscher, P. 24, 27, 63, 162  
Buggle, F. 317  
Bull, G.M. 233, 319  
Burk, K.H. 335  
Burke, C.L. 289, 322  
Burton, R.R. 131, 319
- Cahan, L.S. 53, 318  
Campbell, D.T. 119f., 319f.  
Carlton, S.T. 320  
Carter, R.S. 233, 319  
Carver, R.P. 112, 319  
Cattell, R.B. 43f.  
Cherkes Julkowski, M. 131, 319  
Chiout, H. 142, 319  
Christiani, R. 37f., 273, 298, 318  
Clauss, G. 319  
Coan 43  
Cohen, R. 230, 234, 319  
Coleman, J.S. 60, 319  
Collis, B.E. 252, 259  
Combe, A. 319  
Comber, L.C. 60, 319  
Cook, T.D. 119f., 319  
Copei, F. 270, 320

- Coulson, D. B. 116, 324  
 Cox, R. 105f., 111, 320  
 Crambert, A. C. 113, 320  
 Cranach, M. v. 137, 320  
 Crano, W. D. 55, 320  
 Cronbach, L. J. 53, 65, 73, 96, 112, 119,  
 304, 320, 331  
  
 Dahl, T. 113f., 320  
 Dahrendorf, R. 58, 333  
 Damarian, F. 338  
 Dann, H. D. 62, 320  
 Dawydow, W. W. 133, 320  
 Denig, F. 20, 40, 320  
 Dettermann, D. K. 319, 333  
 Dicker, H. 242, 320  
 Diederich, P. B. 255, 320  
 Dörner, D. 129, 320  
 Donat, H. 222, 320  
 Drenth, P. J. D. 65, 320  
 Dressmann, H. 62, 320  
 Dumke, D. 211, 293, 320  
 Dyer, J. L. 55, 320  
 Dyson, A. P. 232, 320  
  
 Ebel, R. L. 100, 116, 168f., 177, 196, 321,  
 331  
 Eberle, G. 30, 64  
 Ebner, H. 319  
 Echterhoff, W. 119, 321  
 Eckensberger, L. 318  
 Eckstein, B. 330  
 Eells, W. C. 239, 321  
 Eggert, D. 285, 321  
 Eichler, W. 258, 321  
 Eisebitt, R. 302, 321  
 Elliot, E. C. 239, 242, 337  
 Endrigkeit, F. 328  
 Engemann, A. 236, 321  
 Erler, R. 326  
  
 Fall, S. 333  
 Fassnacht, G. 218, 220, 228, 321  
 Feger, H. 124, 329, 332  
 Feller, G. 157, 321  
 Fend, H. 48, 60, 321  
 Feuerstein, R. 198ff., 321  
 Fiedler, F. E. 61, 321  
 Fingerhut, W. 17, 20ff., 39, 83, 95, 207,  
 321, 338  
 Fippinger, F. 42f., 45, 88, 157, 321  
  
 Fischer, G. H. 65, 92f., 98, 119, 123, 321,  
 337  
 Flammer, A. 125, 200, 292, 298, 321f.  
 Flanders 265  
 Flitner, A. 17, 293, 322  
 Flügge, J. 231, 322  
 Formann, A. K. 119, 123, 321  
 Frazier, S. H. 233, 333  
 French, J. W. 320  
 French, L. A. 133, 319  
 Freud, S. 211  
 Freudenthal, H. 60, 322  
 Fricke, R. 93, 103, 107ff., 113, 116, 322,  
 327  
 Fritz, V. 318  
 Frostig, M. 127, 285, 322, 330  
 Funke, E. H. 20, 41, 322  
 Furby, L. 119, 320  
 Furck, C. L. 18, 322  
 Furst, M. 61, 335  
  
 Gaedike, A. K. 23, 46, 54, 57, 322, 325  
 Gage, N. L. 269, 319, 322, 331  
 Gagné, R. M. 22, 30, 132, 147, 201, 239,  
 322  
 Galanter, E. 135, 331  
 Garvin, A. D. 101, 322  
 Gaude, P. 17, 32, 169, 176f., 190, 309,  
 316, 322  
 Gehlen, A. 17, 322, 327, 336  
 Gerstein, H. 59, 322  
 Gerster, J. 203, 322  
 Gibson, E. 286, 322  
 Glaser, R. 99f., 129, 322, 325, 333  
 Gleser, G. C. 53, 96, 304, 320, 331  
 Glück, G. 290, 322  
 Gössmann, W. 259  
 Goldstein, H. 323  
 Good, T. L. 56, 211, 319  
 Goodman, L. A. 323  
 Goodman, Y. M. 289, 323  
 Graham, G. T. 111, 320  
 Grau, U. 251, 256, 328  
 Graumann, C. F. 210f., 215f., 323, 339  
 Grimm, M. 203, 323  
 Groen, G. J. 133, 323  
 Groffmann, K. J. 200, 228, 323f., 338  
 Gronlund, N. E. 163, 165, 167, 169f., 173,  
 175, 177f., 217, 323  
 Grzesik, J. 263, 265f., 268, 323  
 Gugler, B. 320

- Guilford, J. P. 23, 55, 210f., 323  
 Gulliksen, H. 65, 323  
 Guthke, J. 30, 119, 144, 197, 199ff., 203, 323  
 Guttman, L. 111, 113f., 122, 323  
  
 Haarmann, D. 335  
 Haase, H. 154, 323  
 Hacker, W. 132, 135, 323  
 Haenisch, H. 60, 323  
 Haertel, E. H. 48, 113, 331  
 Haertel, G. D. 323  
 Hahn, H. 252, 255, 323  
 Hambleton, R. K. 112f., 116, 323f.  
 Hammaker, M. 257, 324  
 Haneis, E. 324  
 Hanke, B. 56, 324  
 Hardesty, F. P. 287, 324  
 Harris, Ch. 338  
 Harris, C. W. 112, 324, 318  
 Hartog, P. 239, 324  
 Hartung, K. 56, 324  
 Hasemann, K. 211, 228, 324  
 Hastings, J. Th. 318  
 Haueis 258  
 Hausser, K. 17, 324  
 Heckhausen, H. 16f., 36, 49, 52, 57, 62, 262, 269, 317, 323f., 329, 339  
 Hehl, F. J. 118, 333f.  
 Heipcke, K. 262, 324  
 Heiss, R. 324, 331, 337  
 Heller, K. A. 38, 48, 53, 56, 58, 65, 78, 194, 212, 228, 282, 300ff., 305ff., 317f., 322, 324f., 332f., 335, 338f.  
 Helmreich, R. 121  
 Henry, N. W. 330  
 Hensler, H. 333  
 Hentig, H. v. 15, 17, 293, 325  
 Herbig, M. 115f., 325, 327  
 Hess, T. 277, 280, 325  
 Hetzer, H. 275ff., 281, 285, 325  
 Hilke, R. 125, 325  
 Hively, W. 112, 325  
 Höger, D. 41, 325  
 Hoepfner, R. 55, 323  
 Hofen, N. 159, 318  
 Hofer, A. 289, 325  
 Hofer, M. 62, 210, 322, 325  
 Holtz, K. L. 125, 325  
 Holzman, T. G. 129, 325  
 Hopf, C. 265, 326  
 Hopf, D. 299, 326  
 Hopkins, K. D. 169, 177, 197, 326  
 Horn, R. 143, 148f., 151f., 154, 161, 326, 334  
 Horn, W. 21, 31f., 42f., 45, 326  
 Horst, P. 65, 326  
 Huber, H. P. 118, 326  
 Hugow, K. 157, 321  
 Hummell, H. J. 53, 326  
 Husek, T. R. 103, 333  
 Husén, T. 60, 326  
 Huynh, H. 113, 326  
  
 Indermühle, K. 320  
 Ingendahl, W. 248, 326  
 Ingenkamp, K. 56, 94, 161, 170, 204, 233, 237f., 244, 321, 324, 326, 328, 333f., 336f.  
  
 Jackson, R. 111, 114, 326  
 Jacobs, B. 56, 326  
 Jäger, R. S. 53, 161, 276, 281, 326, 328  
 Jagodzinski, W. 54, 326  
 Janig, H. 332  
 Jantzen, W. 42, 326  
 Jencks, C. 60, 327  
 Jendrowiak, H. 272, 327  
 Johnson, D. L. 333  
 Jung, U. 287, 327  
  
 Kalbermatten, V. 320  
 Kalveram, K. T. 40, 327  
 Kant 29, 327  
 Kasztantow, U. 291, 334  
 Kautter, H. J. 158, 283, 327  
 Keeser, W. 120, 334  
 Keeves, J. P. 54, 56, 59f., 319, 327  
 Kehrbach, K. 327  
 Keil, W. 82, 335  
 Kemmler, L. 23, 42, 51, 59, 327  
 Kempf, W. H. 321  
 Kenny, D. A. 320  
 Kerlinger, F. N. 318  
 Kern, A. 275, 327  
 Kifer, E. 109, 327  
 Kirk, S. A. 317  
 Kirk, W. D. 317  
 Klafki, W. 17f., 263, 327  
 Klahr, D. 133, 327, 334  
 Klauer, K. J. 62, 98f., 115f., 118f., 127, 145, 151, 244, 262, 318, 321, 325, 327ff., 334ff., 340

- Kleber, E. W. 161f., 196, 282, 327  
 Klein, S. P. 113, 327  
 Kleiter, E. F. 117, 119, 122ff., 132, 328  
 Kligen, L. 38, 274  
 Klinger, E. 56, 328  
 Kluwe, R. H. 62, 133, 327f., 334  
 Knapper, C. 209, 339  
 Kötter, L. 251, 256, 328  
 Koob, B. 280ff., 328  
 Kormann, A. 29f., 64, 117, 144, 198ff.,  
 203, 282, 328, 335  
 Kornadt, H. J. 334  
 Kornmann, R. 30, 82, 131, 133, 203f.,  
 281, 319, 328, 331, 334, 336  
 Kosecoff, J. P. 113, 327  
 Kozdon, B. 17, 328  
 Krapp, A. 17, 20, 22ff., 39, 46, 48, 50f.,  
 62, 124, 276, 281f., 304, 307, 327ff., 330,  
 332, 335, 337f.  
 Krathwohl, D. R. 141, 153, 329  
 Kratzmeier, H. 286, 329  
 Krauth, J. 119, 121, 329  
 Krepper, W. 219, 329  
 Kretschmann, R. 331  
 Kretschmer, J. 127, 336  
 Kreuz, A. 57, 336  
 Kreuzer, K. 272, 327  
 Krohne, H. W. 56, 329  
 Kropp, R. P. 147f., 329  
 Krug, S. 23, 48, 55, 329  
 Kubinger, K. 124, 203, 329  
 Küffner, H. 131, 203, 329  
 Kühn, R. 46, 53, 62, 329  
 Kuffner, H. 330  
 Kutzer, R. 30, 32, 132, 157, 290f., 330,  
 334  
 Kvale, S. 233, 330  
  
 Landa, L. N. 127, 330  
 Lang, A. 321, 335  
 Lange, B. 56, 330  
 Langeheine, R. 119, 121, 330  
 Langfeldt, H.-P. 17, 20ff., 24f., 39, 63, 83,  
 95, 207, 225, 305, 321, 330, 338  
 Langhorst, E. 33f., 205  
 Lautmann, R. 230, 330  
 Lavin, D. E. 46, 54, 58, 330  
 Lazarsfeld, O. 330  
 Lenk, H. 47, 330  
 Lens, W. 58, 317  
 Levine, H. 286, 322  
  
 Lewin, K. 23  
 Lichtenstein-Rother, I. 17, 330  
 Lienert, G. A. 65, 70, 72f., 77f., 81, 98,  
 105, 119, 121, 187f., 191, 194f., 232,  
 329f.  
 Livingston, S. A. 112, 330  
 Lockowandt, O. 127, 285, 330  
 Löschenkohl, E. 276, 281, 330, 332  
 Löser, G. 43, 337  
 Lohmöller, J. B. 324  
 Lompscher, J. 129f., 320, 330  
 Lotter, H. 242, 330  
 Lueg, C. H. 251, 330  
 Lüer, G. 329  
 Lühmann, R. 162, 318  
 Lukesch, H. 59f., 323, 330, 336  
  
 Madus, G. F. 318  
 Mager, R. F. 145, 153, 330  
 Magnusson, D. 65, 330  
 Mai, W. 96, 331  
 Malgadi, R. G. 257f., 330  
 Mandl, H. 46, 58, 62, 281f., 304, 324,  
 327ff., 330f.  
 Marjoribanks, K. 23, 58, 319, 331  
 Markowa, A. K. 320  
 Marshall, J. L. 113, 331  
 Masendorf, F. 121, 132, 331  
 Masia, B. B. 153, 329  
 Mayer, R. E. 128, 133, 331  
 McCarthy, J. H. 317  
 McKeachie, W. J. 233, 331  
 McNeil, J. D. 267, 331  
 Medley, D. M. 209, 331  
 Meehl, P. E. 73, 303, 320, 331  
 Mees, U. 211f., 223, 331  
 Meier, F. 117, 119, 331  
 Meis, R. 156, 276, 286, 331  
 Meister, H. 133, 196, 204, 319, 327f., 331,  
 334, 336  
 Mercer, C. D. 37, 284, 290, 331  
 Merkelbach, V. 253, 258f., 331  
 Merkens, H. 112, 331  
 Messick, S. 338  
 Meyer, W. U. 334  
 Michael, W. B. 303f., 331  
 Michaelis, W. 329  
 Michel, L. 26, 86, 96, 228, 323f., 331, 338  
 Mietzel, G. 16, 331  
 Miller, G. A. 135, 331  
 Miller, L. B. 55, 320

Mitzel, H. E. 209, 221, 331  
Möbus, C. 118ff., 124, 332  
Möller, C. 146, 332  
Moeller, M. L. 231, 332  
Monroe 286  
Morsh 265  
Moser, H. 262, 332  
Mühle, G. 55, 332  
Müller, A. 277, 280, 325  
Müller, R. 156, 287, 332  
Munz, W. 283, 327  
Murray, H. 211  
  
Nagl, W. 118ff., 124, 332  
Neeb, K. E. 336  
Nesselrode, J. 119f., 318  
Neubauer, W. F. 215, 332  
Neumann, J. 131, 332  
Neumayer, R. 329  
Nevermann, K. 326  
Newcomb, T. 268, 332  
Nickel, H. 37, 228, 256f., 273, 276ff.,  
281f., 286, 293, 307, 324f., 332, 339f.  
Niegemann, H. M. 62, 332  
Niemeyer, W. 286, 332  
Nipkow, K. E. 17, 332  
Nitko, A. J. 162, 332  
Novick, M. R. 112f., 332  
Nunnally, J. c. 168ff., 172, 177, 181, 183,  
187, 332  
  
Oeveste, H. z. 304, 340  
Oldenbürger, H.-A. 332  
Opp, K. D. 53, 333  
Osburn, H. G. 113, 333, 337  
Osterland, J. 224, 333  
  
Panchaud, G. 229, 333  
Parey, E. 336  
Pause, G. 61, 266, 333  
Pawlik, K. 29, 117, 276, 333  
Payrhuber, F. J. 251, 259, 318, 339  
Pedersen, D. M. 233, 333  
Peetz, H. 282, 335  
Peisert, H. G. 58, 333  
Pellegrino, J. W. 129, 325, 333  
Perlwitz, F. 329  
Perrez, M. 330  
Perrig-Chiello, P. 298, 322  
Petermann, F. 53, 117ff., 122ff., 132, 328,  
333f., 338f.  
  
Petri, G. 18, 333  
Pfistner, H.-J. 17, 333  
Piaget, J. 132, 289, 333  
Piontkowski, U. 62, 333  
Plickat, H. 203, 333  
Pokorny, A. D. 233, 333  
Popham, W. J. 103, 106ff., 162, 197, 267,  
322, 331, 333  
Porter 44  
Portmann, R. 336  
Preiser, S. 55, 232, 333f.  
Prell, S. 20, 329  
Pribram, K. H. 135, 331  
Priester, H. J. 287, 324  
Pritz, V. 233  
Probst, H. 30, 32, 132f., 157, 290f., 330,  
334  
Prohansky, H. 321  
Propham, W. J. 322  
Pulver, U. 321, 335  
Putz-Osterloh, W. 133, 334  
  
Quack, L. 46, 334  
  
Radatz, H. 131, 133, 203, 289ff., 334  
Rajaratram, N. 320  
Ramseger, J. 292, 318  
Rasch, G. 27f., 93, 108ff., 122, 334, 337  
Ravitsch, M. M. 337  
Raynor, J. D. 317  
Reichenbecher, H. 303, 307, 318  
Reinartz, A. 322, 328  
Reinartz, E. 322  
Renn, H. 121, 334  
Repp, B. H. 321  
Resnick, L. B. 130, 133, 323, 334  
Reuchlin, M. 238, 334  
Revenstorf, D. 120, 334  
Rey, R. E. 319  
Rheinberg, F. 23, 48, 55, 57, 59, 329, 334  
Rhodes, E. C. 239, 324  
Richter, H. 326  
Rieder, O. 157, 334  
Robeck, M. C. 288f., 334  
Roeder, B. 331  
Roeder, P. M. 21, 133, 301, 335  
Rosemann, B. 28, 32, 63ff., 78, 143f., 228,  
301, 303ff., 307, 325, 335  
Rosenberg, L. 339  
Rosenshine, B. 61, 335  
Rost, D.-H. 282, 328, 340

Rost, J. 118, 123, 335  
 Roth, H. 285, 317, 324, 335  
 Roudabush, G. E. 108, 335  
 Rubinstein, S. L. 129, 201, 335  
 Rüdiger, D. 117, 124, 276, 280ff., 335  
 Rüegg, T. 298, 322  
 Rupprecht, H. 116, 327  
  
 Sader, M. 82, 335  
 Saldern, M. v. 62, 317  
 Salvia, J. 127, 340  
 Samstag, K. 157, 276, 335  
 Sander, A. 335  
 Sander, E. 51, 335  
 Sander, H. 328  
 Sang, F. 16, 335  
 Schäfer, H. 286, 335  
 Schallberger, U. 335  
 Schaller, K. 17, 335  
 Scheerer-Neumann, G. 286, 291, 335  
 Schell, C. 55, 332  
 Schenk-Danzinger, L. 132, 276f., 335  
 Schiefele, H. 124, 328, 332, 337f.  
 Schlee, J. 133, 204, 319, 328, 331, 334, 336  
 Schlesinger, J. M. 113f., 323  
 Schmack, E. 297, 335  
 Schmalohr, E. 286, 335  
 Schmalt, H.-D. 334  
 Schmid, F. W. 321, 335  
 Schmid, H. 200, 321  
 Schmidbauer, M. 58, 336  
 Schmidt, H. J. 296, 336  
 Schmidt, H. P. 288, 336  
 Schmidt, P. 53, 333  
 Schmidt, R. 335  
 Schmidt-Denter, U. 219, 336  
 Schmolders, G. 17, 336  
 Schneewind, K. A. 59, 330, 336  
 Schneider, K. 57, 336, 339  
 Schneider, W. 54, 336  
 Schnitzer, A. 17, 336  
 Schnotz, W. 30, 132, 144, 200ff., 204, 231, 336  
 Schofenegger, J. 332  
 Schorb, A. O. 58, 336  
 Schott, F. 116, 127, 327, 336  
 Schrand, H. 158, 336  
 Schreckenberg, W. 31, 36f., 64, 138, 206, 260, 265f., 272, 336  
 Schröter, G. 257ff., 336  
 Schruppf, B. 250, 336  
 Schubert, M. T. 329  
 Schümann, M. 35f., 206  
 Schümer, R. 230, 319  
 Schulz, W. 265, 336  
 Schwartz, E. 296, 336  
 Schwarz, E. 118, 120, 336  
 Schwarzer, C. 51, 196, 224, 228, 307, 319, 324, 327, 335f.  
 Schwarzer, R. 53f., 56, 196, 224, 228, 292, 327, 330, 336f.  
 Scriven, M. 29, 64, 337  
 Seidel, G. 327  
 Seidenberg, B. 321  
 Seitz, W. 43, 221, 337  
 Selg, H. 331  
 Shafto, M. 324  
 Shannon 201  
 Shavelson, R. J. 112, 337  
 Shinedling, M. M. 333  
 Shoemaker, D. M. 113, 337  
 Signer, M. 286, 337  
 Sikorski, P. B. 131, 329  
 Simons, H. 22, 54f., 122f., 337  
 Sixtl, F. 65, 319  
 Skurnik, L. S. 147, 340  
 Snow, R. E. 53, 320  
 Sommerla, G. 235, 241, 244, 337  
 Spada, H. 118, 123, 133, 327f., 334f.  
 Spache, G. D. 288, 337  
 Sporer, S. L. 200, 328  
 Stanley, J. C. 119f., 169, 177, 197, 319, 326  
 Starch, D. 239, 242, 337  
 Stark, G. 18, 336f.  
 Steffens, W. 142, 319  
 Stegmüller, W. 47, 337  
 Steinhagen, K. 53, 337  
 Stene, J. 27, 337  
 Stelzl, I. 93, 118, 121, 337  
 Sternberg, R. J. 333  
 Sternberger, R. J. 319  
 Stoker, H. W. 147f., 329  
 Stourfer, S. A. 323  
 Storz, L. 158, 327  
 Straka, G. A. 62, 337  
 Strasser, S. 137, 337  
 Strittmatter, P. 56, 326  
 Subkoviak, M. J. 113, 337  
 Süllwold, F. 20, 81, 337f.  
 Süssenbacher, G. 332  
 Swaminathan, H. 113, 116, 324, 338  
 Szeminska, A. 289, 333

Tagiuri 210  
 Tausch, A. 338  
 Tausch, R. 266, 338  
 Tent, L. 42ff., 56, 241, 276, 281, 285, 325, 338  
 Teschner, W. P. 17, 32, 169, 176f., 190, 309, 316, 322, 336  
 Teusch, M. 17, 338  
 Tewes, U. 287, 338  
 Thomae, H. 212, 338  
 Thorndike, E. L. 209, 338  
 Thorndike, R. L. 52, 338  
 Thornton, G. R. 233, 338  
 Thurmer, F. 258, 321  
 Tiedemann, J. 51, 55, 338  
 Tilmann, H. 272, 338  
 Timmermann, G. 119, 328  
 Tismer, K.-G. 217, 228, 338  
 Trabasso, T. 324  
 Travers, R. M. W. 331, 335  
 Treiber, B. 31, 52ff., 60ff., 332f., 338f.  
 Treumann, K. 21, 301, 335  
 Trost, G. 56, 338  
 Trudewind, C. 59, 338  
 Tscherner, K. 221, 338  
 Tucker, L. R. 119f., 338  
 Twellmann, W. 330  
 Tyler, R. W. 163, 177, 338  
  
 Überla, K. 40, 338  
 Ulshöfer, R. 245, 259, 338  
  
 Valette, R. M. 288, 338  
 Vargas, J. S. 105f., 320  
 Voigt, J. 336  
 Voigt, M. 317  
 Volpert, W. 135, 338  
  
 Wahl, D. 52, 62, 338  
 Walberg, H. J. 323  
 Walter, H. 142, 339  
 Walter, R. 56, 326  
 Wang, M. C. 100, 339  
 Warr, P. B. 209, 339  
  
 Wasna, M. 56, 339  
 Watzlawick, P. 230  
 Weber, A. 254, 339  
 Wechsler, D. 80, 287, 339  
 Weede, E. 53, 339  
 Weiner, B. 48, 339  
 Weinert, F. E. 31, 52, 57f., 62, 240, 322, 324f., 332f., 337ff.  
 Weinläder, H. 325  
 Weinstein, T. 323  
 Weis, V. 20, 40, 320  
 Weiss, R. 237, 239, 242, 339  
 Wendeler, J. 37, 109, 131f., 159, 203, 206, 254, 256, 273, 287f., 339  
 Wetzell, H. 317  
 Weuffen, M. 285f., 319  
 Wieberg, H. J. W. 336  
 Wiczercowski, W. 35f., 203, 206, 256f., 304, 332f., 339f.  
 Wiedl, K. H. 119, 200, 203, 340  
 Wiggins, J. S. 38, 303, 340  
 Wilhelm, M. 317  
 Will, H. 56, 340  
 Wilson, J. A. R. 288f., 334  
 Wimmer, H. 48, 340  
 Winkelmann, W. 289, 340  
 Winther, A. 219, 329  
 Wohlrab, U. 203, 323  
 Wolf, B. 59, 340  
 Wolf, W. 244, 340  
 Wood, R. 147, 340  
 Wright, B. D. 109, 215, 340  
 Wurst, E. 329  
 Wygotski, L. S. 30, 129, 198, 200, 340  
  
 Ysseldike, J. E. 127, 340  
  
 Zängle, M. 54, 326  
 Zeller, W. 275, 340  
 Ziegenspeck, J. 16, 237, 340  
 Ziegler, R. 53, 326  
 Zielinski, W. 18, 51, 62, 239, 325, 340  
 Zimmermann, K. W. 20, 41, 340

# 10. Sachregister

- Abitur 38
- ability training approach 127
- Ähnlichkeitsfehler 211
- Äquivalenzkoeffizient 232
- Alternativantwort 166ff.
- Analyse kognitiver Komponenten 128ff.
- Anforderung 30, 41, 54, 125ff., 278f.
- Anforderungsanalyse 30, 64, 125ff.
  - -struktur 129f.
- Angst (Ängstlichkeit)
  - Auswirkungen auf die Leistung 53
  - und Schulleistung 53
- Anspruchsniveau 138ff., 269f., 312
- Antworten 172
  - freie 86f.
  - gebundene 86f., 166
  - nicht gebundene 166, 174ff.
- approximative Expertenabstimmung 303
- ATI-Forschung 24, 53, 63
- Aufgaben/analyse 86ff., 125ff., 183ff.
  - -bewertung 182ff.
  - -gültigkeit 107ff.
  - -konstruktion 32, 63, 113f., 150ff., 162ff.
  - -reihung 181
  - -schwierigkeit 87, 104f.
    - (s. a. Schwierigkeit)
  - -trennschärfe 105f.
    - (s. a. Trennschärfe)
- Aufsatz/beurteilung 35f., 143, 206, 242, 245ff.
  - -güte 250f.
- AUKL 301
- Auslese s. Selektion
- Auslesediagnostik 37f.
- Auswahl/antwort 166
  - -verfahren 222f.
- Auswertungsobjektivität s. Objektivität
  
- Baseline 120
- Begabung 42f.
- Begabungsdiaagnose 38, 198ff., 299f.
- Beobachtung 211ff.
  - Gelegenheits- 214
  - in standard. Situationen 214
  - systematische 211ff.
- Beobachtungs/pläne 212f.
  - -skalen 223f.
  - -systeme 218ff.
  - -verfahren 211ff.
- Beratung 273f., 304f.
- Beschreibung 211, 215ff.
  - Simultan- 216
  - standardisierte 226ff.
- best answer type 169
- Beurteilung 211, 229ff.
  - gebundene 222ff.
  - vorformulierte 222ff.
  - vorstrukturierte 222
- Beurteilungen/bögen 220
  - -fehler 208ff.
  - -formen 221f.
  - -kriterien 29, 237ff., 251ff., 316
  - -schritte 221
- Bildungsberatung 37
- Bildungsweginformation 299f.
  
- Chi-quadrat-Analyse 107
- Clusteranalyse 119, 122f.
- cognitive components approach 125ff.
- concurrent validity 72
- confidential limits 78
- construct validity 73
- content validity 72
- correct answer type 169
- criterion level 100
- curriculare Validität s. Validität
- cut-off-point 100
  
- Denken 129f.
  - divergentes 55
  - konvergentes 55
- Deutsch 40, 241f., 245ff.
- Diagnose-Prognose-Ansatz 38, 299ff.
- Diagnostik
  - anforderungsorientierte 125ff.
  - strukturorientierte 30, 126, 132ff.
- Differenzwerte 117ff.
- Diktat 241f.
- Distraktoren 114, 169ff., 183, 186
- dosierte Diskrepanz 125ff.
- Durchführungsobjektivität s. Objektivität

- Einschulungsdiagnostik s. Schuleingangsdiagnostik
- Einzelfalluntersuchung 118
- Empirisches Relativ s. Relativ
- Entwicklungsbedingungen und Schulleistung 49ff., 57
- Ergänzungsantworten 166, 175f.
- error of central tendency 211
- Erwartung 210f.
- Erwartungseffekt 210f.
- Erziehung/stil 49, 58f.
- -ziele 18, 138
- essay-test 176f.
- extended response question 177
- Evaluation
  - formative 29
  - summative 29
- facet-design-method 114
- Faktorenanalyse 40, 82, 119f.
  - von Zeugnissen 40ff.
- Fähigkeit 46, 109f., 276
- Fähigkeits/diagnostik 37, 284ff.
  - -konstrukt 276
- Fehler 46, 130, 208ff.
  - konstanter 69
  - systematischer 69
- Fehler/analyse 126, 130ff., 156, 273, 287ff.
  - -theorie 69f.
  - -wert 69ff., 75
- Förderung (individuelle) 17, 64, 96, 131, 146
- Förderungsdiagnostik 37, 203, 273, 282ff.
- Fremdsprachen 32, 35, 158, 243
- Führungsstil 61
- Funktionsmodelle (von Lehrzielen) 127
- Gelegenheitsbeobachtung 212
- Generalisierungstheorie (CRONBACH) 112
- generosity error 211
- Gesamtschule 308f.
- Grundschulgutachten 302f.
- Grund- und Leistungskurse 310ff.
- Gültigkeit 68, 72f. (s. a. Validität)
- Gütekriterien 68
- GUTTMAN-Skala 111, 122
- Halbierungs-Reliabilität s. Reliabilität
- Halo-Effekt 209, 231
- Handeln 134ff., 248f.
  - Ordnungsdimensionen 135
  - Qualitätskriterien 135
- Handlung 135
- Handlungskompetenz 134
- Hof-Effekt 209
- Implizite Persönlichkeitstheorie 210f., 231
- Informationsverarbeitung 209
- Informelle Tests s. Tests
- Inhaltsvalidität s. Validität
- Interaktion 137, 278ff.
- Interpretationsobjektivität s. Objektivität
- Interpretationsübung 166, 177ff.
- interpretive exercises 177
- Intervallskala 25, 67, 84
- Items
  - änderungssensitive 121
  - unterrichtssensitive 115
- Item/analyse 33, 86f., 104ff., 163, 183ff.
  - -charakteristikkurve 110
  - -homogenität 107
  - -konstruktion 113f., 144, 162ff.
  - -Lehrziel-Kongruenz 114f.
  - -muster 113f.
  - -revision 186ff.
  - -selektion 186ff.
- Kategoriensystem 218ff.
- Klassen/arbeiten 35, 237ff.
  - -bildung (Skalentransformation) 67
- Klassifikations/konsistenz 112
  - -skala 25
- Klausur 311, 316
- Koeffizient, LIVINGSTONSCHER 112
- Kompetenz 126, 131, 254
- Konfigurationsfrequenzanalyse 119, 121
- Konsistenz/analyse 189ff.
  - -reliabilität s. Reliabilität
- Konstruktvalidität s. Validität
- Kontingenztafeln 121
- Kontrastfehler 211, 231
- Korrelation 52, 70ff., 87
- Korrelationskoeffizient 70ff.
- Kreativität 23, 55
- Kriteriums/begriff 100f.
  - -bezogene Validität s. Validität
- Kurssystem (Sekundarstufe) 38, 274, 308ff.
- Kurzantworten 166, 175f.
- Kurzaufsatzantwort 166, 176f.

- Längsschnitt 51
- Latent-trait-Modell 123
- Legasthenie 155, 285ff.
- Lehr/-Lern-Forschung 31, 62, 133
  - -Lern-Prozess 51, 268
  - -plangültigkeit s. Validität
  - -stoffstrukturen 62
  - -strategien 36
  - -ziele 99, 143, 145
- Lehrer 41, 56, 60f., 138
- Lehrer/beurteilung 36, 64, 260ff., 270ff.
  - -urteil 143, 305f.
  - -verhalten 33, 49, 134, 140, 265f.
- Leistungs/Abfolge-Systeme 122
  - -bedingungen (schulische) 24, 59
  - -begriff 16f.
  - -beurteilung 15, 145, 164
    - objektive Verfahren 31ff., 34, 38, 143f.
    - subjektive Verfahren 33, 38, 143f., 205ff.
    - kriteriumsbezogene 27, 143, 274
    - kriteriumsorientierte 29, 63, 99ff.
    - lernzielbezogene 145
    - normorientierte 29, 63
    - strukturorientierte 30, 126
  - -bewertung 28f., 308
  - -diagnostik 15, 38, 63f., 299f.
    - Aufgaben (Funktionen) 18ff., 28, 37f.
  - -erfolgskontrolle 294f.
  - -fähigkeit 199
  - -feststellung 28f., 164, 241
  - -fortschritt 19, 38, 294
  - -gruppe 140
  - -motivation 23, 49
  - -prinzip 15, 17
  - -produkt 15
  - -prozess 15, 19, 23, 46, 117, 126, 282
  - -schwierigkeiten 33
  - -standard 101f.
  - -struktur (subjektive) 129f.
  - -test 27 (s. a. Test)
  - -verhalten 15
- Lern/hierarchie 147
  - -kontrolltest 29, 32, 64, 162
  - -kriterium 100
  - -leistungsstandard 100
  - -messung 29f., 117ff.
  - -steuerungstest 29, 64, 162
  - -test 30, 64, 144, 198ff.
- Lernweg-Lernstand-Diagnostik 122
- Lernziele 31, 99, 145ff., 178
  - operationalisierte 145ff., 162ff.
  - Spezifikationstabelle der 32, 162ff.
- Lernziele nach BLOOM
  - Analyse 147ff.
  - Anwendung 147ff., 151
  - Evaluation 148ff.
  - Synthese 147, 149
  - Verstehen 147ff.
  - Wissen 147ff.
- Lernziel/analyseschema 143
  - -bestimmung 33
  - -definition 31, 145
  - -diagnostik 37, 284, 289
  - -operationalisierung 31f., 143, 145ff.
  - -taxonomie 31, 147f., 164, 312
- Lesen 30, 89f., 154f., 285ff.
- Linear-logistische Modelle 119, 123
- LIVINGSTONscher Koeffizient 112
- mastery test 113
- matching exercises 172
- Mathematik 40, 55, 90f., 141, 147, 157, 242f., 289ff.
- Mehrfachwahl/antwort 166, 168ff.
  - -aufgabe 151
- Messen 25, 65f.
- Meß/fehlertheorie 24, 63, 75
  - -funktion 102
  - -instrumente 70
  - -theorie 27, 68f.
  - -verfahren 28
  - -wert 66ff.
  - -wiederholung 70f.
- Messung
  - kriteriumsbezogene 99ff.
  - normbezogene 99ff.
- Mildefehler 211
- Modelltest 110
- Moderatoransatz 53, 302
- multiple choice form 166, 168ff.
- Naturkunde 32
- Neigungsdifferenzierung 308ff.
- Nominalskala 25, 67
- Norm 84f., 293
  - bei Schultests 83, 191ff.
- Normierung 33, 83, 163
- Notengebung s. Zensurierung
- Numerisches Relativ s. Relativ

- Objektivität 25, 68f., 73ff., 82ff., 231f.
  - Auswertungs- 74, 82f., 172, 253f.
  - Durchführungs- 74, 82f.
  - Interpretations- 74, 83
  - von Schultests 94
- Ogive 193
- Ökologie 278ff.
- ökopsychologischer Ansatz (Schulreife) 277ff.
- Ordinalskala 25, 67, 84
- Ordnungsantworten 166, 172ff.
- Overachievement 52
  
- Pädagogisierungsphase 30, 199
- Paralleltest-Reliabilität s. Reliabilität
- Persönlichkeit und Schulleistung 48ff., 54ff.
- Personenmessung, itemfreie (RASCH) 109
- Pfadanalyse 53f., 59, 122
- Populationsabhängigkeit (Testtheorie) 93
- Positionseffekt 211
- predictive validity 72
- Projektionsfehler 211
- Protokoll
  - Gedächtnis- 216f.
  - Simultan- 216
  - zusammenfassendes 217
- Prototyp-Item 108
- Prozentränge 85f., 192ff.
- Prozentrang-Normen 84f., 192ff.
- Prozeßdiagnostik 29, 64, 144, 280
- Prüfung
  - Kriterienkataloge 239, 251ff.
  - Kritik 230ff.
  - mündliche 33ff., 205f., 229ff., 316
  - schriftliche 33ff., 150, 205f., 235, 237ff., 245ff.
  
- Querschnitt 51
  
- Rangreihen-Umwandlung 194ff.
- Rangskala 25
- RASCH-Modell 27f., 93, 109f., 122
- Ratewahrscheinlichkeit 168, 171f.
- Ratingskala 33f., 223f.
- Rationalskala 25
- Rechtschreibung 30, 91f., 155f., 241f., 285ff.
- Regressionsanalyse 52, 121
- Referat 244, 316
- Reifungskonzept 275
  
- Relativ
  - empirisches 65
  - numerisches 66
- Reliabilität 25f., 33, 68ff., 75ff., 87, 189ff., 232
  - Halbierungs- 77
  - Interrater- 232
  - Konsistenz- 77f., 189ff.
  - Parallel- 71, 76f.
  - Retest- 76f.
- Reliabilitätsbestimmung bei Schultests 95, 163, 189ff.
  - bei kriteriumsbez. Messungen 111ff.
- Repräsentativität von Schultests 94f.
- restricted response questions 177
- Retest-Reliabilität s. Reliabilität
- Rohpunktskala 84, 190
  
- Schätzskalen 33f., 222ff.
- Schereneffekt 58
- Schüler/beobachtung 208ff.
  - -beurteilung 37, 208ff., 299
  - -verhalten 33, 208ff., 265ff.
- Schul/angst 23, 56
  - -aufgaben 35, 143, 206
  - -eignung 299ff.
  - -eingangsdagnostik 273, 275ff.
  - -erfolg 23, 38f., 46f.
    - Prognose des 38, 275, 299ff.
  - -fächer 40ff., 88ff., 154ff., 241ff.
  - -fähigkeit 276
  - -laufbahnberatung 19f., 38, 273, 299ff.
- Schulleistung
  - Definition der 15f., 46ff.
  - Indikator der 16, 40ff.
  - Konstrukt der 16, 40ff.
  - Prognose der 38, 51, 56
- Schulleistung (und)
  - Bedingungsfaktoren 20, 39, 46ff., 276, 301f.
    - Strukturmerkmale 23, 39, 46ff., 58
    - Prozeßmerkmale 23, 39, 46ff., 58
  - Bedingungsmodell 22, 47ff.
  - Dimensionen (Faktoren) 20ff., 39, 42
  - Familie 48ff., 58ff.
  - gesellschaftlicher Hintergrund 48ff.
  - Intelligenz 21f., 41, 54f.
  - kognitive Faktoren 54ff.
  - Kreativität 23, 55
  - Lehrverhalten 49, 60f.
  - nicht-kognitive Faktoren 23, 56ff.

- Persönlichkeit 23, 48ff., 54ff.
- soziales Lernumfeld 23
- soziokulturelles Niveau 53, 58
- Schule 24, 48ff., 59f.
- Schulnoten 40f.
- Unterrichtsatmosphäre 62
- Schulleistungs/analyse 305
  - -fähigkeit 46ff.
  - -verhalten 46ff.
  - -schwierigkeiten 37, 127
- Schul/noten 42ff., 56, 80f.
  - als Urteile 240f., 245ff.
  - Objektivität der 94
  - Reliabilität der 95
  - Validität der 95f.
- reife 37, 273, 275ff.
- system 59f., 278
  - Modell der 37, 278ff.
- tests 24, 27f., 63, 65, 87ff., 94ff., 99, 125
  - formelle 94f., 154ff.
  - informelle 63, 162ff.
  - Klassifikation von 154ff., 160f.
  - Konstruktion 162ff.
  - Übersicht 160f., 200ff.
- versagen 23, 46f.
- Schwierigkeit (Aufgaben-) 26, 57, 87, 109f., 163
- Schwierigkeitsgrad 183f.
  - index 104, 183
- Selektion 276, 283, 304
- serieller Effekt 211
- Signifikanztest 119
- Skala 66ff.
  - Rating- s. Ratingskala
  - Schätz- s. Schätzskala
- Skalen/niveau 16, 25
  - -transformation 67
- Skalogrammanalyse 111
- soziale Koedukation 37, 293
- Spezifikationstabelle 162ff.
- Stabilitätskoeffizient 232
- Stamm (des Items) 169f.
- Standard/-Äquivalent-Normen 84f.
  - -beobachtung 214
  - -meßfehler 78
  - -norm 84f.
  - -normalverteilung 84f.
  - -schätzfehler 81
  - -schulnoten 84, 86
  - -situation 214
  - -werte 194ff.

- Status/diagnostik 29, 64
  - -test 199
- Strengfehler 210
- Summenprozentkurve 193
- Task analysis 30, 125ff.
- Taxonomie (BLOOM) 31, 35, 147ff., 162 ff.
- Test
  - eindimensionaler 110
  - informeller 144, 162ff.
  - kriteriumsbezogener 101ff., 125
  - lehrzielorientierter 99
  - lernzielorientierter 99
  - normbezogener 101f., 125
  - Schulleistungs- s. Schulleistungstest
  - standardisierter 25, 75, 143f., 154ff.
- Test/analyse bei kriteriumsbez. Tests 102ff., 111ff.
  - -batterien 88
  - -eichung, personenfreie 109
  - -gütekriterien 25
  - -heft 163, 180ff.
  - -instruktion 180
  - -konstruktion 102, 162ff.
  - -modell
    - logistisches 109
    - stochastisches 93
  - -theorie
    - Axiome der 69
    - klassische 24, 28, 63f., 65ff., 68ff.
    - Kritik der 26f., 92f., 103f.
  - -vorform 32, 180ff.
- Testen 65
- Transformation
  - Flächen- 85
  - lineare 85
- Trennschärfe 26, 87, 183ff.
  - -index 87, 105f., 185
  - klassische 106
  - lehrzielorientierte 106
- true-false-items 166
- Übereinstimmungskoeffizient 108f.
- Umordnungsantwort 166, 174
- Underachievement 19, 52
- Unterricht 18, 31, 46, 62, 126, 134f., 233ff., 278
  - Individualisierung des 37, 293
- Unterrichts/arrangement 139f., 266ff..
  - -beobachtung 264ff.

- beurteilung 31, 36f., 64, 134ff., 205f., 260ff., 270ff.
  - Gütekriterien 36f., 138f., 260
  - Gewichtungproblem 140ff.
- leistung 260ff.
- planung 32f., 36, 139f., 145, 147ff., 261ff.
  - Funktion 261f.
  - Gütekriterien 262ff.
- steuerung 18, 267
- strukturen 263f.
- ziele 36, 262
- Untertests 54, 88
- Urteilsfehler 33, 56, 208ff.
  
- Validität 25f., 33, 68f., 72f., 75, 87, 191, 232
  - curriculare 72, 79f., 96
  - differentielle 53, 301f.
  - inhaltliche 72, 79, 107, 232
  - Konstrukt- 73, 81f.
  - kriteriumsbezogene 72, 80
  - Übereinstimmungs- 70, 72, 233
  - von Schultests 79, 95f., 163, 191
  - Vorhersage- 72, 80f., 95f.
- Validitätsbestimmung bei kriteriumsbez. Meßverfahren 113ff.
- Varianzanalyse 52, 119
- Veränderungsmessung 29, 64, 117ff.
  - Dilemmata 118
  - graph. Methoden 118ff.
  - Methoden 118ff.
- Verhaltens/analyse 125, 220
  - beobachtung 33ff., 143, 205, 208ff. (s. a. Beobachtung)
  - beschreibung 34, 205, 216f.
  - beurteilung 205f., 208ff.
- Verhältnisskala 25, 67
- Verteilungsfunktion, kumulierte, logistische 110
- Vertrauensgrenzen 78
- Vierfelder/analyse 106
  - häufigkeit 108
- Vorhersage
  - klinische 303
  - statistische 303
- Vorhersagevalidität s. Validität
- Vorkenntnisse 22, 55f.
- Vortest-Nachtest-Design 106, 199
  - Differenzindex 106, 120
- Vorurteil 210
  
- Zeichensysteme 218
- Zeitreihenanalyse 119f.
- Zensierung (Notengebung) 38, 94, 143, 237ff., 314f.
  - von Schulaufsätzen 258f.
- Zentraltendenz 211
- Zettelarbeit 243
- Zeugnis 40ff., 292ff.
  - zensurenfrei 38, 273, 294ff.
- Zeugnis/bericht 297
  - brief 297
  - form 296ff.
- Zuordnungsantworten 166, 172ff.
- Zuverlässigkeit 68 (s. a. Reliabilität)