

## Inhaltsverzeichnis

<b>Einleitung</b> .....	11
<b>1 Vorüberlegungen auf der Grundlage bisherigen Wissens</b> .....	15
1.1 Messen als Voraussetzung naturwissenschaftlicher Erkenntnis .....	15
1.1.1 Forschungsmethoden in Psychologie und Physik .....	16
1.1.2 Die unterschiedliche Behandlung von Extremwerten .....	19
1.1.3 Fehlinterpretationen in der Psychologie .....	20
1.2 Intelligenz – Begabung – Begabungsstruktur .....	22
1.2.1 Definitionen von Intelligenz und Begabung .....	22
1.2.1.1 Zum Zusammenhang zwischen Intelligenz und Schulleistungen .....	24
1.2.1.2 Die Intelligenz, ein Konstrukt .....	25
1.2.1.3 Die Diskussion um den Begriff „Begabung“ .....	25
1.2.1.4 Intelligenzstruktur versus Begabungsstruktur .....	27
1.2.2 Intelligenztheorien .....	28
1.2.3 Die Konstruktion von Intelligenztests .....	33
1.2.3.1 Regeln für die Konstruktion von Intelligenztests als Messinstrumente .....	34
1.2.3.2 Die Erstellung der Normskala .....	35
1.2.3.3 Die Bedeutung von Altersnormen .....	37
1.2.4 Die Anlage-Umwelt-Kontroverse .....	38
1.3 Das gemeinsame Interesse von Physik und Psychologie an der Intelligenzstruktur .....	39
1.3.1 Die Begabungsstruktur als Voraussetzung für die Bedienung technischer Geräte .....	39
1.3.2 Die Begabungsstruktur als Voraussetzung der Eignung .....	40
1.3.3 Der I-S-T als Messverfahren im Sinne der DIN-Normen .....	41
1.4 Hinweise auf mathematisch-naturwissenschaftliche Begabungsschwerpunkte im Schulsystem .....	42
1.4.1 Hinweise auf diskrepante Begabungen in Biografien prominenter Persönlichkeiten .....	43
1.4.2 Hinweise auf diskrepante Begabungen nicht-prominenter Legastheniker .....	48
1.4.3 Hinweise auf diskrepante Leistungen in Intelligenztests von legasthenen Schülern .....	50
1.4.4 Hinweise auf diskrepante Intelligenzstrukturen in Notentafeln .....	56
1.4.5 Die Wahrnehmung diskrepanter Begabungen in der Reformdiskussion vor 1970 .....	59
1.5 Die Kritik an der schulischen Auslese im deutschen Schulsystem .....	62
1.5.1 Belege für das Übergewicht der sprachlichen Fächer aus der Reformdiskussion .....	62

1.5.2	Zahlenmäßiges Übergewicht sprachlich-geisteswissenschaftlicher Fächer .....	65
1.5.3	Die Überbewertung der Rechtschreibung .....	68
1.5.3.1	Die Bedeutung von schriftsprachlichen Leistungen in der schulischen Auslese .....	71
1.5.3.2	Das Gewicht der Rechtschreibung im deutschen Abitur .....	73
1.5.3.3	Ein Fallbeispiel aus dem Schuljahr 1999/2000 .....	74
1.6	Ursprünge der Überbewertung von Sprache und Schriftsprache .....	75
1.6.1	Das humanistische Bildungsideal und die daraus erwachsenen Vorurteile .....	77
1.6.2	Die Snow-Hypothese von den „zwei Kulturen“ .....	81
1.6.3	Die beiden Hirnhälften .....	85
1.6.4	Denken in Bildern und Denken in Worten .....	90
1.6.5	Abschließende Betrachtung zu den „Zwei Kulturen“ .....	94
<b>2</b>	<b>Theoretischer Teil: Der I-S-T als Messinstrument</b> .....	<b>97</b>
2.1	Einleitende Überlegungen zum I-S-T .....	97
2.2	Die Anwendung des I-S-T .....	98
2.2.1	Die Aufgabenstellung und die geprüften Funktionen .....	98
2.2.2	Der Gang der Auswertung .....	100
2.2.3	Die erzielbaren Ergebnisse .....	102
2.2.4	Grenzen der Aussagefähigkeit .....	104
2.3	Fragen an den I-S-T .....	105
2.3.1	Leistungsbereiche, die der I-S-T nicht erfasst .....	106
2.3.2	Nicht erfasste Arten des Denkens .....	108
2.4	Die Unsicherheiten in den Testaussagen .....	110
2.4.1	Messfehler und statistische Schwankungen .....	110
2.4.2	Auswirkungen zufälliger Messabweichungen einer Aufgabengruppe .....	111
2.4.3	Gesamtstandardwert und Fehlerfortpflanzungsgesetz .....	113
2.4.4	Messbereichsumfang für den Rohwert und Standardwert .....	114
2.4.5	Absolute und relative Messung .....	115
2.4.5.1	Absolute Messung .....	115
2.4.5.2	Relative Messung .....	115
2.4.6	Verringern zufälliger Schwankungen durch Mittelung .....	116
2.4.7	Aussagewert der Daten .....	117
2.4.8	Mögliche Fehlerquellen in der aktuellen Testsituation .....	118
2.4.9	Exkurs: Die Statistik in der Wissenschaft .....	119
2.5	Intelligenzstrukturen von Personen, Personengruppen, Berufen .....	120
2.5.1	Voraussetzung von Intelligenz-Strukturkurven: Bildung von Mittelwerten .....	121
2.5.1.1	Voraussetzung einer Mittelwertbildung .....	121
2.5.1.2	Das gewogene arithmetische Mittel .....	123

2.5.1.3	Mittelwertbildung bei stark voneinander abweichenden Standardwerten .....	124
2.5.1.4	Mittelwertbildung bei relativen Standardwerten .....	126
2.5.2	Schlussfolgerungen .....	126
2.6	Die Legasthenie als nicht der Norm entsprechendes Beispiel .....	127
2.6.1	Mehrheiten und Minderheiten im Beruf .....	129
<b>3</b>	<b>Praktischer Teil: Die Untersuchung und ihre Auswertung</b> .....	<b>131</b>
3.1	Beschreibung der untersuchten Stichprobe .....	131
3.1.1	Aufgabenstellung des Instituts .....	133
3.1.2	Einmaligkeit der Messung .....	133
3.1.3	Vertraulichkeit der Messergebnisse .....	134
3.2	Die tabellarische Darstellung der Ergebnisse .....	134
3.3	Auswertung und Analyse der I-S-T-Daten .....	141
3.3.1	Die Fragestellung .....	141
3.3.2	Aussagen der I-S-T-Daten über den Test selbst .....	141
3.3.2.1	Was wird gemessen? .....	141
3.3.2.2	Die Beantwortung der Fragebogen .....	143
3.3.2.3	Analyse der grafischen Darstellung richtig gelöster Aufgaben .....	143
3.3.2.3.1	Unterschiedliche Schwierigkeitsgrade innerhalb der Aufgabenreihe .....	143
3.3.2.3.2	Verminderung irregulärer Schwankungen durch Glättung .....	146
3.3.2.3.3	Unterschiede zwischen den Parallelförmigen .....	148
3.3.2.4	Analyse des Lösungsverhaltens in den verschiedenen Aufgabengruppen .....	150
3.3.2.4.1	Die Lösungsstrategien der Physiker .....	150
3.3.2.4.2	Lösungsverhalten bei den Aufgabengruppen ME und ZR .....	153
3.3.2.4.3	Vergleich des Lösungsverhaltens in 6 Aufgabengruppen .....	155
3.3.3	Aussagen der I-S-T-Daten über die Physiker des IAP .....	157
3.3.3.1	Die Intelligenzstruktur .....	157
3.3.3.2	Untergruppen der 34 Physiker des IAP .....	158
3.3.3.3	Intelligenzstrukturkurven von Wissenschaftlern der exakten Wissenschaften .....	160
3.3.3.4	Häufigkeitsverteilungen für die absoluten und relativen Standardwerte .....	163
3.3.3.5	Prüfung auf Normalverteilung der relativen Standardwerte für WA und RA .....	166
3.3.3.6	Leistungs- und Begabungsunterschiede .....	168
3.3.3.6.1	Begabungsunterschiede in den Daten von Amthauer .....	168
3.3.3.6.2	Leistungs- und Begabungsunterschiede in unseren Daten .....	171
3.3.3.6.3	Die Bedeutung von Begabungs- und Leistungstiefs .....	176
3.3.3.6.4	Die Bedeutung von Begabungs- und Leistungsspitzen .....	177

3.3.3.7	Diskrepanzen in der Intelligenzstruktur .....	179
3.3.3.7.1	Leistungsabweichungen von mehr als 11 Standardwerten .....	180
3.3.3.7.2	Leistungsabweichungen von mehr als 15 Standardwerten .....	181
3.3.3.7.3	Typische und nicht typische Intelligenzstruktur .....	183
3.3.3.8	Beziehungen zwischen den Aufgabengruppen .....	184
3.3.3.8.1	Die Konsequenzen sehr hoher bzw. niedriger Werte .....	187
3.3.3.8.2	Beziehungen zwischen höchsten und niedrigsten Standardwerten .....	187
3.3.4	Begabungsstrukturen: Berufe von AMTHAUER und Physiker unserer Untersuchung .....	189
3.3.4.1	Zusammenhang zwischen den Aufgabengruppen WÜ und WA .....	190
3.3.4.2	Standardwerte benachbarter Berufe liegen eng beieinander .....	196
3.3.4.3	Gegensätzliche Leistungen der Physiker in den Aufgabengruppen ..	199
3.3.4.4	Intelligenzhöhe (Gesamtstandardwert) und Aufgabengruppe .....	201
3.3.4.5	Intelligenzhöhe (Gesamtstandardwert) und Beruf .....	206
3.3.4.6	Das Zusammenwirken verschiedener Begabungen zu einem Beruf ..	208
3.3.4.7	Intelligenzstruktur-Band eines Berufes statt einer Intelligenzstruktur- Kurve .....	208
3.3.4.8	Zusammenfassung der Berufe zu Berufsgruppen .....	210
3.4	Auswertung der Abiturnoten der 34 Physiker .....	213
3.4.1	Aussageergebnisse der Abiturnoten .....	214
3.4.1.1	Die Abiturnoten der 34 Physiker und ihre Häufigkeitsverteilung ....	214
3.4.1.2	Häufigkeitsverteilung der Abiturnoten von Untergruppen .....	218
3.4.1.3	Vergleich der Mathematik- und Physiknoten mit den Sprachnoten ..	219
3.4.1.4	Noten der typisch und untypisch Physikbegabten .....	220
3.4.1.5	Zusammenhang zwischen IQ und Abiturnoten .....	221
3.4.1.5.1	Zusammenhang zwischen Abitur-Durchschnitt und Sprachnoten .....	222
3.4.1.5.2	Die Bedeutung der Sprache in der Schule .....	223
3.4.2	Abiturnoten und Leistungen in den Aufgabengruppen des I-S-T .....	224
3.4.2.1	Sprachnoten und Merkfähigkeit .....	224
3.4.2.2	Abiturdurchschnitt und Merkfähigkeit .....	226
3.4.2.3	Prüfung auf Normalverteilung relativer Standardwerte von ME und RA .....	227
3.4.2.4	Die Bedeutung von RA- und ME-Befähigung im Beruf des Physikers .....	229
3.4.2.5	Der Gegensatz von Selbständigkeit im Denken (SE) und Gedächtnis (ME) .....	230
3.4.2.6	Sprachliches versus bildhaftes Denken .....	233
3.5	Hinweise auf Auslesefaktoren in Schule und Berufsausbildung .....	233
3.5.1	Ausleseprinzipien und die absoluten und relativen Standardwerte .....	235
3.5.2	Ausleseprinzipien und die Streuung der absoluten Standardwerte.....	236
3.5.3	Zur Bedeutung der Reihenfolge der Ausleseprinzipien .....	238
3.5.4	Der Auslesefaktor Rechtschreibung .....	241
3.5.4.1	Intelligenzquotient und Rechtschreibleistung .....	242

---

3.5.4.2	Rechtschreibleistung und Schultyp .....	245
<b>4</b>	<b>Zusammenfassung – Diskussion – Schlussfolgerungen .....</b>	<b>247</b>
4.1	Zusammenfassung der Ergebnisse .....	247
4.1.1	Die Ergebnisse der Analyse des I-S-T als Messinstrument .....	247
4.1.2	Die Analyse der Testergebnisse der Untersuchung im IAP .....	250
4.1.3	Die Analyse der Abiturnoten .....	253
4.2	Diskussion der Ergebnisse .....	254
4.2.1	Geisteswissenschaftliche versus naturwissenschaftliche Begabung .....	255
4.2.2	Die Legasthenie als Erscheinungsbild diskrepanter Begabungsstruktur ...	258
4.2.3	Intelligenztheorien und Intelligenztests .....	261
4.3	Schlussfolgerungen .....	264
4.3.1	Für notwendig erachtete Veränderungen im Bildungssystem .....	264
4.3.2	Schlussfolgerungen zur Intelligenzforschung .....	266
4.3.3	Geisteswissenschaftliche – mathematisch-naturwissenschaftliche „Welt“ ..	268
<b>5</b>	<b>Fachausdrücke .....</b>	<b>271</b>
<b>6</b>	<b>Literatur .....</b>	<b>275</b>
<b>7</b>	<b>Register .....</b>	<b>281</b>

# Einleitung

Vor etwa 25 Jahren, Mitte der siebziger Jahre, stellte sich in der Angewandten Physik der Christian Albrechts Universität in Kiel die Frage nach brauchbaren Kriterien für die Zulassung zum Physikstudium. Angesichts des damals drohenden Numerus Clausus erschien es notwendig, die wirklich Geeigneten zum Studium zuzulassen. Die Abiturnoten aber hatten sich als eher ungeeignet erwiesen, ein erfolgreiches Physikstudium vorauszusagen.

Es war zu diesem Zeitpunkt schon bekannt, dass im Einstufungsverfahren nach dem Graduierten-Förderungsgesetz die Doktoranden, die das Vordiplom und das Diplom im Fach Physik mit „gut“ und „sehr gut“ bestanden hatten, und die in beiden Fachgutachten die Beurteilung „eignen sich für den Hochschulnachwuchs“ erhalten hatten, in den Abiturzeugnissen durchgehend oder überwiegend überraschend schwache Noten aufwiesen, auch in den naturwissenschaftlichen Fächern. Die Abiturzeugnisse waren den Gutachtern zunächst nicht bekannt. Das Entscheidungsgremium meinte, als es die Zeugnisse sah, dass es sich nicht um dieselben Personen handeln könne! Da aber nicht zu übersehen war, dass die Bestleistungen im Fach Physik und die schwachen Abiturzeugnisse doch von denselben Personen stammten, erklärte man sich die Diskrepanz zwischen Studienerfolg und Abiturnoten mit der Naturwissenschaftsfeindlichkeit der Schule.

Für den mit diesen Fragen konfrontierten Professor der Physik (Mitautor), der selber legasthene Schwächen und entsprechend „schlechte“ Schulnoten gehabt hatte, lag die Frage näher, ob es sich in diesen Fällen um legasthene Begabungen oder doch zumindest um Personen mit spezifischen Begabungsschwerpunkten handeln könnte. Um ein erfolgreiches Physikstudium vorauszusagen, müßte man genau diese Begabungsstruktur bestimmen, gegebenenfalls durch Tests oder Eignungsprüfungen.

Am Ende der 60er Jahre stieß die Psychologin (Mitautorin) bei einer Testuntersuchung in einer Schulklasse mit Intelligenz- und Schulleistungstests auf einen Schüler, dessen Testleistungen teilweise auf dem Niveau von Gymnasialschülern, bei schriftsprachlichen Aufgaben jedoch unter dem Niveau der Hauptschule lagen! Dieser Schüler sollte am Ende der 4. Klasse in die Sonderschule für Lernbehinderte überwiesen werden. Angesichts so großer Diskrepanzen im Leistungsbild des Schülers stellte sich die Frage, ob es sinnvoll und vernünftig sein könne, die weitere Schullaufbahn dieses Jungen durch die schwache Rechtschreibung bestimmen zu lassen oder das Niveau seines Denkvermögens zu berücksichtigen. Gleichzeitig wurde auch deutlich, was Legasthenie bedeutet: Ein diskrepantes Erscheinungsbild in den Schulleistungen, die Unterschätzung der intellektuellen Fähigkeiten des Schülers und damit Schullaufbahnprobleme!

Aus dieser Einsicht folgten Fragen an die Intelligenzdiagnostik: Ist es bei einem so diskrepanten Intelligenztestprofil sinnvoll, für die Summe aller Testwerte einen Gesamtwert abzulesen? Ein Durchschnittsmaß, um das es sich in diesem Fall handelt, soll als Orientierungshilfe dienen. Welche Orientierung aber hat man, wenn die Rechtschreibleistung eines Schülers dem Niveau der Sonderschule für Lernbehinderte zuzuordnen ist, seine mathematischen und Denkleistungen aber dem Niveau des Gymnasiums? Einen Durchschnittswert für zwei so extreme Leistungen zu berechnen, entspricht dem scherzhaften Beispiel aus Knaurs Buch der modernen Statistik: „Wenn jemand mit einem Fuß auf der Herdplatte steht und mit dem anderen im Eis-

kasten, dann sagt der Statistiker: Im Durchschnitt ist ihm angenehm warm.“ (SWOBODA, 1971, S. 39, Bildunterschrift)

In Diagnostik-Seminaren sprach man damals noch von einer „Schulintelligenz“, die sich in Konzentrations- und sprachlichen Intelligenztestaufgaben abzeichnete und die für die Prognose der Schullaufbahn herangezogen wurde. Die Psychologie als Wissenschaft vertrat offenbar die Auffassung: Wenn für ein Fortkommen im Schulsystem „Schulintelligenz“ die notwendige Voraussetzung ist, kann man für Schüler mit nicht-sprachlichen Begabungsschwerpunkten leider nichts tun! Aber steht die Rede von der „Schulintelligenz“ nicht schon im Widerspruch zum Konzept der Intelligenz als eines globalen Merkmals? Die mathematisch-naturwissenschaftlichen Begabungen kannte man in der Diagnostik (noch) nicht, obwohl einige wenige Schulpsychologen bereits begannen, sich für Schüler mit einer legasthenen bzw. diskrepanten Begabungsstruktur einzusetzen.

Wegen der Betonung schriftsprachlicher Leistungen bei der Entscheidung über Versetzung oder den Zugang zu Schulformen mit höherem Anforderungsniveau entsprach die Beurteilung einer „Schulintelligenz“ zwar in der Tat den Gegebenheiten im Schulsystem und einer dadurch bestimmten Schullaufbahnprognose, im Grunde jedoch nicht dem Intelligenzkonzept, nach dem die Intelligenztests konstruiert waren: Intelligenz zu messen als Leistungsfähigkeit, die sich vor allem in der Lösung von Denk- und Problemlöse-Aufgaben ausdrückt und nicht vorrangig in schriftsprachlichen Leistungen.

Da die besseren Testleistungen des legasthenen Jungen aus der damals untersuchten vierten Klasse für mathematische und technisch-konstruktive Fähigkeiten sprachen, drängte sich zudem die Frage auf, ob legasthene Begabungen nicht häufiger in mathematisch-naturwissenschaftlichen Fächern vertreten sein müssten, also auch in der Physik? War nicht Albert EINSTEIN Legasthener und eine Reihe anderer Nobelpreisträger aus dem Fach Physik auch? Gibt es möglicherweise eine speziell für mathematisch-naturwissenschaftliche und technische Leistungen befähigende Intelligenzvariante, die zwar legasthene Schwächen einschließen kann, aber nicht notwendig mit ihnen verbunden sein muss? Werden Personen, die mit dieser Intelligenzvariante „begabt“ sind, vielleicht nur deswegen in der Schule auffällig, weil sie für die Sprachen weniger begabt sind, wie andere Schüler für Mathematik und Physik?

So trafen sich beide Autoren dieses Buches in einem gemeinsamen Interesse, das zu dem Vorhaben führte, die Physiker des Instituts für Angewandte Physik in Kiel mit einem Intelligenztest zu prüfen, der geeignet ist, Stärken und Schwächen im Leistungspotenzial zu erkennen. Ziel dieser Untersuchung war herauszufinden, ob sich eine spezifische Intelligenzstruktur unter den Teilnehmern gehäuft finden ließ, die als günstige Voraussetzung für die Studienleistungen im Fach Physik gelten kann und möglicherweise in einer Reihe von Fällen mit legasthenen Schwächen verbunden ist.

Zur Untersuchung einer solchen Fragestellung stand seinerzeit der I-S-T 70 (Intelligenz-Strukturtest in der überarbeiteten Fassung von 1970) vom AMTHAUER zur Verfügung. Die Ergebnisse dieser Untersuchung werden im dritten Teil dieser Arbeit ausführlich dargestellt.

Es geht also in diesem Buch um Intelligenz und Intelligenzmessung, speziell um die Vertiefung des Ansatzes von AMTHAUER, nicht einen Gesamtwert der Intelligenz, „den IQ“ bei Beurteilungen, Numerus-Clausus-Entscheidungen und Berufsberatungen zugrunde zu legen, sondern die *Intelligenzstruktur*, also Stärken und Schwächen im Leistungsprofil. Solche Überlegungen führten zwangsläufig dazu, ältere Konzepte der Intelligenz und Intelligenzmessung

zu überdenken. Die heutige internationale Diskussion über Intelligenzkonzepte hat genau diese Überlegungen aufgegriffen.

Es geht auch – speziell in der durchgeführten Untersuchung – um die Eignung des I-S-T 70 als „*Messinstrument*“ und um mögliche Korrekturen der Aufgabenstellungen des Tests. Zwar sind Veränderungen dieses Tests inzwischen erfolgt (I-S-T 2000, AMTHAUER, BROCKE und LIEPMANN, 1998 und AMTHAUER, BROCKE, LIEPMANN und BEAUDUCEL, 2001), doch wurde in der überarbeiteten Testform die tabellarische Darstellung von Berufsprofilen aufgegeben, so dass Vergleiche, wie sie in unserer Arbeit zwischen Berufsgruppen angestellt und kritisch beleuchtet werden, mit dem neuen Test (auch in der revidierten Form 2000 R) nicht mehr möglich sind. Auf den derzeitigen Stand der Intelligenz-Strukturforschung wird im Zusammenhang mit den Intelligenztheorien eingegangen.

Um die eingangs gestellte Frage nach der spezifischen Eignung für ein Physikstudium beantworten zu können, wird man allenfalls eine Übertragung der Ergebnisse dieser Untersuchung auf Einzelfall-Messungen mit dem I-S-T 2000 versuchen müssen. Die grundsätzlichen Überlegungen zur Messung mit dem I-S-T und zur Profil-Interpretation bleiben im Hinblick auf praxisnahe Berufs- oder Studienberatung aktuell, auch wenn mit dem neuen Instrument in dieser Hinsicht erst Erfahrungen gewonnen werden müssen.

Nicht zuletzt führt dieses Buch auch zu *bildungspolitischen Fragen*, die nach den Ergebnissen der TIMSS-Studien (BAUMERT et al. 1997, BAUMERT et al. 2000) und seit den Ergebnissen der PISA-Studie (Deutsches PISA-Consortium, 2001) besonders drängend sind. Will die Bundesrepublik Deutschland für die zukünftigen Anforderungen gerüstet sein, dann muss sie nicht nur mehr Lehrer einstellen, sondern auch die Anforderungen im Schulsystem verändern! Entsprechende bildungspolitische Fragen wurden in der Reformdiskussion um 1970 schon einmal angesprochen. Sie sind jedoch in Öffentlichkeit und Bildungssystem bis heute auf seltsame Weise wieder untergegangen. Gesamtschulen sollten vor allem die Durchlässigkeit im Schulsystem und einen qualitativ besseren Unterricht garantieren. Die TIMSS-Studie kam u.a. zu dem Ergebnis, dass die Leistungen von Schülern aus schwedischen Gesamtschulen zu den besten gehörten, während Schüler aus deutschen Gesamtschulen deutlich ungünstiger abschnitten. Die Frage nach dem „Warum?“ ist bis heute nicht beantwortet. Es gibt Hinweise darauf, dass Schülerinnen und Schüler in deutschen Schulen weniger zur Selbständigkeit erzogen werden (MESSNER, 1998, S. 61 ff., BRÄU, 1998, S. 202 ff.), was sich in den naturwissenschaftlichen Fächern und Mathematik stärker negativ auswirken dürfte, als in den sprachlichen Fächern.

Möglicherweise ist das deutsche Bildungssystem u.a. auch deswegen so wenig leistungsfähig, weil unterschiedliche Begabungsausprägungen der Kinder nicht wahrgenommen werden. Es dominiert immer noch die Vorstellung von einem Begabungsniveau, das bei allen Kindern für alle Fächer als gleich hoch angenommen wird. Wenn Menschen sich aber hinsichtlich ihrer Intelligenz nicht allein im Niveau, d.h. im Ausprägungsgrad intellektueller Leistungsfähigkeit unterscheiden, sondern auf gleichen Niveaustufen durch Begabungsschwerpunkte, muss das Schulsystem in einem demokratischen Staat dies nicht berücksichtigen? Darf es weiterhin einseitig sprachliche Begabungen, Schüler mit einer an das Bildungssystem angepassten „Schulintelligenz“ bevorzugen? Könnte das Schulsystem seine Leistungsfähigkeit nicht steigern, wenn es diese Begabungsunterschiede berücksichtigen würde, anstatt alle Kinder vor die gleiche, bei der Leistungsmessung dominierende, sprachlich-rechtschriftliche Anforderung zu

stellen? Wenn beispielsweise als „Grundtugenden des Physikstudiums“ Problemverständnis, Kreativität und Flexibilität genannt werden (DANIELMEYER 1992 b), dann fragt man sich, wo an deutschen Schulen diese Tugenden eingeübt werden und mit welchem Gewicht sie sich in Zeugnisnoten niederschlagen?

Im Zusammenhang mit der Computermesse CEBIT 2000 beklagten die großen Firmen Siemens und IBM den eklatanten Mangel an Fachkräften auf diesem Gebiet. In beiden Fällen werden Veränderungen im Schulsystem und in der Lehrerausbildung angemahnt. Die Erfahrungen von Siemens z.B. bestätigen, was bereits im Zusammenhang mit dem Numerus Clausus für das Medizinstudium befürchtet wurde: „Früher wurden die Bewerber mit Bestnoten genommen, und ausgerechnet von denen erweisen sich viele als unfähig.“ (FAZ vom 29.02.2000) Es könnte daran liegen, dass Bestnoten vor allem in den sprachlichen und geisteswissenschaftlichen Fächern erworben werden und dass selbst Bestnoten in Mathematik eher das Reproduzieren erlernter Lösungswege widerspiegeln als die Fähigkeit, mit mathematischen Problemen kreativ umzugehen. Das Bildungsideal des 19. Jahrhunderts, das unser Schulsystem und die Einstellung vieler Bildungspolitiker bis heute prägt – trotz der Kritik, die bereits im Verlauf des 19. Jahrhunderts laut wurde – reicht offenbar im 21. Jahrhundert nicht mehr aus, um ein zeitgerechtes Schulsystem zu gestalten. Jugendliche brauchen zugleich eine Allgemeinbildung traditioneller Art und eine Vorbereitung auf die Erfordernisse des heutigen Berufslebens. Das bedeutet – last not least –, sie für lebenslanges Lernen zu begeistern.

Die bildungspolitischen Fragen sollen in diesem Buch im Anschluss an die Ergebnisse der I-S-T-Untersuchung behandelt werden. Für die Zukunft zeichnet sich heute schon ein globaler Wettkampf auf den Weltmärkten ab. In diesem Wettkampf werden Aufgabenstellungen überwiegend durch interdisziplinäre Zusammenarbeit vor allem der mathematisch-technischen und naturwissenschaftlichen Disziplinen gelöst werden müssen (DANIELMEYER 1992 a). Daraus ergeben sich entsprechende Herausforderungen auch für eine notwendige Anpassung des Schul- und Ausbildungswesens und der Universitäten an die gesellschaftliche Realität. Auch die Green-Card Diskussion gehört in diesen Kontext.

# 1 Vorüberlegungen auf der Grundlage bisherigen Wissens

Die interdisziplinäre Zusammenarbeit zwischen Physik und Psychologie (soweit sie experimentell arbeitet) hat zwei gemeinsame Grundlagen: Psychologie und Physik benutzen Messungen, um zu wissenschaftlichen Aussagen zu gelangen und beide Fächer haben mit Menschen zu tun, die solche Messungen durchführen.

Menschen wenden Messinstrumente an und werten Messungen aus. Menschen machen Fehler. In beiden Wissenschaften ist es notwendig, Messverfahren zu entwickeln, bei denen die menschliche Unzulänglichkeit als Unsicherheitsfaktor bei Messungen möglichst weitgehend reduziert wird.

Im ersten Abschnitt dieser Vorüberlegungen wird daher ein Schwerpunkt auf generelle Probleme von Messungen, von ihren Auswertungen und auf Forschungsmethoden aus der Sicht der exakten Wissenschaften gelegt.

Der zweite Abschnitt bezieht sich auf den Stand des heutigen Verständnisses von Intelligenz, Begabung und Begabungsstruktur, auf Probleme der Intelligenzmessung und den theoretischen Hintergrund dafür.

Im dritten und vierten Abschnitt dieser Vorüberlegungen wird die mathematisch-naturwissenschaftliche Begabungsstruktur als ein Phänomen belegt, das sich in vielen Biografien bekannter Wissenschaftler und Künstler und in vielen Schullaufbahnen legasthener Schüler sowie in ihren späteren Lebensläufen abzeichnet.

Anschließend geht es um die Kritik an der als einseitig beschriebenen Auslese im deutschen Schulsystem und um die Frage nach den Ursachen für diese Einseitigkeit. Diese Frage führt schließlich im 6. Abschnitt zu den Aussagen der SNOW-Hypothese von den zwei Kulturen und zu einer neuen Hypothese, aufgrund der vielfältigen Belege für spezifische Begabungen.

## 1.1 Messen als Voraussetzung naturwissenschaftlicher Erkenntnis

Wenn man den festlichen Saal der Royal Institution in London betritt, so fällt der Blick auf eine Umschrift, die ein Zitat aus Lord KELVINS Werken wiedergibt. Sie besagt:

*„Wenn ihr das, wovon ihr spricht, messen und durch eine Zahl ausdrücken könnt, so wisst ihr etwas von euerm Gegenstand. Könnt ihr es aber nicht messen, könnt ihr es nicht in Zahlen ausdrücken, so sind eure Kenntnisse armselig und sehr ungenügend. Dann könnt ihr vielleicht am Anfang der Kenntnisse sein, seid aber in euren Gedanken kaum bis zur Wissenschaft vorgedrungen, gleichviel um was für einen Gegenstand es sich handelt.“*

(Lord KELVIN 1824-1907. Physikalische Blätter 27 Jahrg. Juni 1971, Heft 6, Seite 260)

Auch die Psychologie bemüht sich seit der zweiten Hälfte des 19. Jahrhunderts darum, eine messende Wissenschaft zu sein. Allerdings – darauf wird noch einzugehen sein – stellt die Konstruktion von Messinstrumenten für psychische Merkmale spezielle Anforderungen.

Über das Messen in der Physik sagte Max PLANCK in seiner Rede über die „Physikalische Gesetzlichkeit“ am 14. 2. 1926:

*„Das Wesen der physikalischen Gesetzlichkeit und der Inhalt der physikalischen Gesetze läßt sich nicht durch reines Nachdenken erschließen. Es gibt hierfür keinen anderen Weg, als den, sich vor allem an die Natur zu wenden, in ihr möglichst zahlreiche und vielseitige Erfahrungen zu sammeln, dieselben miteinander in Vergleich zu bringen und zu möglichst einfachen und weittragenden Sätzen zu verallgemeinern. Da der Inhalt einer Erfahrung um so reicher ist, je genauer die Messungen sind, die ihr zugrunde liegen, so versteht sich von selbst, dass der Fortschritt aller physikalischen Erkenntnis auf das engste verknüpft ist mit der Verfeinerung der physikalischen Instrumente und mit der Technik des Messens.“* (NEFF 1976)

Diese Aussage ist genau so auf Messungen in der Psychologie zu beziehen, obgleich es sich bei Intelligenztests nicht um psychophysische Messungen handelt, für die es technische Messgeräte gibt, sondern um Messungen mit nichttechnischen „Instrumenten“, die speziell konstruiert werden müssen. Die Devise der beginnenden empirischen Forschung in der ursprünglich geisteswissenschaftlich orientierten Psychologie lautete: „Messen, was messbar ist und messbar machen, was nicht messbar ist“! Dabei kommt – als ein besonders zu lösendes Problem – die Frage der Inhalte, die gemessen werden sollen, ins Spiel. Wenn ein Phänomen wie „die Intelligenz“ gemessen werden soll, bedarf es zunächst des Nachdenkens darüber, mit welchen Aufgabenstellungen intelligentes Verhalten erfasst werden kann – eine inhaltliche Frage –, bevor man mit der Konstruktion der Messinstrumente beginnen kann. Auf dieses Problem wird später näher eingegangen.

### 1.1.1 Forschungsmethoden in Psychologie und Physik

Wie in allen anderen empirischen Wissenschaften steht auch in der Psychologie wie in der Physik am Anfang jeden Forschens die Beobachtung eines Phänomens, über das man Genaueres wissen möchte. Es werden Annahmen zu diesem Phänomen entwickelt, die man überprüft und die je nach den Ergebnissen zu weiteren Annahmen führen. Lassen sich auch diese Annahmen in entsprechenden Überprüfungen stützen, führen sie letztlich zu einer Theorie. Um wissenschaftlich akzeptiert zu werden, muß diese Theorie sich wiederum in einer Reihe von Untersuchungen mit entsprechenden Fragestellungen anhand objektiver Kriterien bewähren. Zunächst geht es um eine theoretische Überprüfung, in der z.B. die folgenden Fragen gestellt werden: Welchen Informationsgehalt besitzt die Theorie? Ist sie präzise formuliert? Ist sie in sich logisch konsistent, mit anderen Theorien logisch vereinbar? Ist sie empirisch überprüfbar? In einem zweiten Schritt werden aus der Theorie einzelne Hypothesen, also Aussagen über bestimmte beobachtbare und damit überprüfbare Sachverhalte abgeleitet. Auch diese Hypothesen sind in empirischen Untersuchungen zu überprüfen. Stehen die Untersuchungsergebnisse im Einklang mit den Hypothesen, so wird die Theorie dadurch gestützt; erweisen sich die aus der Theorie abgeleiteten Hypothesen aufgrund der empirischen Resultate als falsch, muß die Theorie verändert oder verworfen werden.

Physik und Psychologie haben im Prinzip die gleichen Forschungsmethoden, insofern sie messende Verfahren einsetzen. Soweit sie Häufigkeiten zählen und vergleichen und graphische Darstellungen benutzen, um Informationen sichtbar zu machen, sind ihre Methoden iden-

tisch. Ob Messungen in beiden Wissenschaften verlässlich reproduzierbar sind, wird an anderer Stelle diskutiert.

Auch muss man in beiden Wissenschaften gleichermaßen eine Vorstellung davon gewonnen haben, was man eigentlich messen will – in der Psychologie das Problem der Validität oder inhaltlichen Gültigkeit von Tests. Danach geht es um die Frage, wie man den Gegenstand des Interesses messen kann. In der Physik stehen viele Messgeräte zur Verfügung. Dennoch stellt sich auch in dieser Wissenschaft, wie in der Psychologie, nicht selten das Problem, ein geeignetes Messgerät erst erstellen zu müssen. Allerdings arbeiten die Psychologen – anders als Physiker – mit so genannten psychometrischen Verfahren, das sind die nach besonderen Regeln konstruierten Tests, als „Messinstrumente“, bei denen vorwiegend mit Papier und Bleistift gearbeitet wird: Die Testperson muß Fragebogen ausfüllen oder Aufgabenlösungen ankreuzen. Aus der Anzahl der Lösungen pro Testperson in einer großen Stichprobe werden schließlich die standardisierten Werte einer Messskala errechnet, die fortan als Maßstab für diesen Test gelten.

Bei der Testkonstruktion interessiert allerdings nicht allein, wie genau bzw. zuverlässig der Test als Messinstrument misst, sondern – wie bereits angedeutet – zuerst die eigentlich wichtigere Frage, ob mit den vorgegebenen Aufgabenstellungen wirklich das Phänomen gemessen wird, das man inhaltlich erfassen will (siehe unter 1.2.3). Das scheint in der Psychologie ein schwierigeres Problem zu sein als in der Physik.

Letztlich aber läuft auch diese Art der „Messung“ auf dasselbe Prinzip hinaus: Es werden auch bei der Bearbeitung von psychologischen Tests Zahlen als Rohwerte gewonnen, auf die sich die weitere Auswertung stützt. Streng genommen, weiß man nur nicht, ob die „Sandkörnchen“, die man zählt, unterschiedlich groß sind?

Bei der Erforschung neuer Gebiete versucht man in der Physik, so weit wie möglich, keine Vor-Annahmen zu treffen. Da man die Informationen, die in dem neuen Gebiet auffindbar wären, noch gar nicht kennt, würde man mit einem Interesse, das durch Vor-Annahmen schon eingeschränkt ist, für unerwartete Informationen aus den Messergebnissen unter Umständen nicht mehr offen sein. Sie würden nicht wahrgenommen werden oder könnten verlorengelassen werden. Man versucht beispielsweise die gewonnenen Daten in einer Reihe unterschiedlicher Grafiken darzustellen und probiert aus, welche dieser Grafiken verwertbare Aussagen erkennen lässt. Auf diese Weise kann Neues, vorher Unbekanntes, entdeckt werden.

In der Psychologie wird im allgemeinen gefordert, „theoriegeleitet“ zu forschen. Das heißt, dass Fragestellungen aus bereits formulierten Theorien abgeleitet werden sollen. Man befragt in diesem Fall nicht eigentlich die Wirklichkeit. Vielmehr überprüft man die aus Modellvorstellungen (die notwendigerweise reduziert sein müssen) der vermuteten Wirklichkeit abgeleiteten Annahmen. Aufgrund solcher Annahmen, die für den in Frage stehenden Sachverhalt oder Zusammenhang aufgestellt wurden, werden bei Forschungsvorhaben Null- und Alternativhypothesen aufgestellt, die nach statistischer Auswertung der Daten akzeptiert oder abgelehnt werden. So versucht man, nach der Philosophie der statistischen Methoden, Irrtümer zu vermeiden. Dabei kann es allerdings vorkommen, dass ein Ergebnis nahe der statistischen Signifikanzgrenze vernachlässigt und die (der Realität entsprechende) Alternativhypothese zurückgewiesen wird. Dies ist vor allem dann möglich, wenn man Einflüsse, die sich beim Zustandekommen der Messdaten ausgewirkt haben, nicht kennt oder (noch) nicht durchschaut.

So sind psychologische Untersuchungen nicht selten allein darauf ausgerichtet, eine vorher formulierte, aus einer Theorie abgeleitete, Hypothese zu überprüfen. Lassen sich die Ergebnisse auf einem vorher gewählten statistischen Signifikanzniveau sichern, so werden sie interpretiert und vielleicht noch mögliche Verbesserungen der Versuchsanordnung diskutiert. Eine weitere Analyse der vorliegenden Daten mit der Frage „welche unbekanntes Informationen sind in den Daten enthalten?“ wird selten vorgenommen. Das heißt, man beschränkt sich auf die Beantwortung der vorab formulierten Fragestellung. Die Forscher scheinen in solchen Fällen nicht neugierig zu sein, zu erfahren, wie sich etwa randständige Fälle in einer Häufigkeitsverteilung von den Fällen unterscheiden, die im mittleren Bereich der Verteilung liegen.

Beispielsweise wird ein vorschulisches Training zur Vorbereitung der Lautunterscheidung und des Heraushörens von Lauten aus dem gesprochenen Wort (phonologische Bewusstheit) durchgeführt, und zwar mit unausgelesenen Kindern im letzten Jahr des Kindergartens. Man begleitet diese Kinder bis zum Ende der zweiten Klasse um zu sehen, wie sich ihre Leserechtschreibleistung im Vergleich zu nicht trainierten Kindern verhält (SCHNEIDER et al. 1998). Das Ergebnis ist zwar nicht überraschend, aber eindrucksvoll: Am Ende der zweiten Klasse liegen nur knapp 7% der trainierten Gruppe auf dem Niveau besonders schwacher Rechtschreibleistungen (Prozentrang unter 10), während in der nicht trainierten Gruppe rund 17% der Kinder als rechtschreibschwach auf dem Niveau der 15% schwächsten Leistungen auffallen. Auf keinem anderen Gebiet der Lese- und/oder Rechtschreibforschung ließ sich für ein gezieltes Training ein so deutlicher Erfolg nachweisen.

Wer aber neugierig ist, fragt weiter: Was unterscheidet die schwächsten 7% der trainierten Kinder von ihren erfolgreicherer Altersgenossen? Sind sie – was wenig wahrscheinlich ist – weniger intelligent? Kommen sie aus einem extrem benachteiligenden Milieu? Haben sie Handikaps der auditiven Wahrnehmung, die ihnen den Zugang zu dem Training erschwert haben? Gibt es andere Gründe für den geringeren Erfolg? Außerdem interessiert den Neugierigen, ob das Prozentrang-Niveau für die Einstufung als „leserechtschreibschwach“ mit der Schulwirklichkeit harmonisiert? Gibt es Kinder, die ein höheres Testniveau erreicht haben, die aber dennoch – eventuell erst im Gymnasium – wegen ihrer Rechtschreibung auffällig werden? Solche Fragen bedürfen der Klärung, damit man Einsichten für die Frühdiagnose von Handikaps oder speziellen Begabungen gewinnen und ggf. das Training für die betroffenen Kinder so verändern kann, dass auch sie davon profitieren.

Allerdings gibt es auch in der Psychologie die Möglichkeit, zunächst hypothesenfrei zu untersuchen, wenn ein so genanntes Erkundungsexperiment durchgeführt wird, dessen Ergebnisse der Hypothesengewinnung dienen. Problem bleibt allerdings auch in diesem Fall, dass psychische Merkmale oder Leistungen von sehr vielen Einflüssen abhängig sein können, die bei der Datenerhebung nicht immer bekannt sind oder nicht kontrolliert werden können.

Wegen der Komplexität der Phänomene, die in der Psychologie untersucht werden, z.B. „Intelligenz“, „Legasthenie“ oder „Motorik“, reichen einmalige Versuchsanordnungen nicht aus. Daher spricht man von Konstrukten, deren Aufklärung nur Schritt für Schritt erfolgen kann, indem man Teilaspekte dieser komplexen Phänomene ins Auge fasst. Die theoretischen Vorstellungen über die in Frage stehende Erscheinung werden durch die Teilergebnisse gewissermaßen „angereichert“, so dass die Ergebnisse solcher Forschung dazu dienen, die vorher entwickelten Modellvorstellungen zu korrigieren, zu erweitern und sie der Wirklichkeit immer mehr anzunähern. So formulieren beispielsweise Intelligenzforscher heute:

*„Die Bedeutung von Intelligenz ergibt sich aus der Gesamtheit der zu einem bestimmten Zeitpunkt wissenschaftlich erschlossenen Teilaspekte von Intelligenz.“ (AMTHAUER et. al. 1999, S. 6)*

Dieses Zitat ist zugleich ein Ausdruck der Bescheidenheit: Was wir nicht durch Testuntersuchungen über die Intelligenz erfahren können, bleibt unbekannt.

In den Anfängen experimenteller psychologischer Forschung galt noch der Satz „Statistics is an unvaluable servant but a bad master“, d.h. die Statistik hat dienende Funktion. In manchen Bereichen der Psychologie hat man heute den Eindruck, dass die Statistik inhaltliche Fragen dominiert. Das zeigt sich in dem scherzhaft gemeinten, aber der Wirklichkeit nicht ganz fernen Slogan „Wir können heute mit immer feiner ausgearbeiteten statistischen Methoden immer weniger inhaltlich relevante Fragen untersuchen“. Diesen Satz im Sinn, macht es für den Psychologen besonders interessant, sich auf die Auswertung von psychologischen Messergebnissen durch einen Physiker einzulassen und zu sehen, welche Erkenntnisse er aus den Daten gewinnt. Er erhält Informationen, die in psychologischen Auswertungen oft aus grundsätzlichen messtechnischen und statistischen Erwägungen in der Regel als „Informationsverlust“ hingenommen und nicht weiter beachtet werden.

### 1.1.2 Die unterschiedliche Behandlung von Extremwerten

Ein Beispiel für die forschungsmethodische Selbstbeschränkung in der Psychologie ist die Behandlung von Extremwerten. Im Vergleich zur experimentellen Physik vernachlässigt die Psychologie ihre Extremwerte, die sogenannten „Ausreißerwerte“. Für diese Werte wird zwar angenommen, dass sie durch spezielle Bedingungen zustandekommen. Im Rahmen der vorher gewählten Fragestellung interessieren diese besonderen Bedingungen jedoch nicht. Anders in der Physik: Beispielsweise bei der Messung der Knochendichte stießen Physiker in Zusammenarbeit mit Medizinern auch auf Ausreißerwerte, denen man in der weiteren Analyse der Daten nachging. Es stellte sich heraus, dass in diesen Fällen der Knochen wegen mangelnder Durchblutung z.B. nach einem Gefäßriss oder infolge einer Thrombose, sich in Auflösung befand.

Eine Parallele dazu stellen die Legastheniker dar, die in Bezug auf ihre Rechtschreibleistung im Vergleich mit Altersgenossen gleicher Intelligenzhöhe, Ausreißerwerte produzieren. Als man sich für den Zusammenhang zwischen Intelligenzhöhe und Rechtschreibleistung interessierte, wurden aus den Daten Korrelationen für die Enge des Zusammenhangs zwischen beiden Merkmalen errechnet. Wie sich herausstellte, hingen Intelligenztestwerte und Rechtschreibtestwerte etwa in gleicher Höhe miteinander zusammen, wie die Intelligenztestwerte aus zwei verschiedenen Tests. Für die Ausreißerwerte – schwache Rechtschreibleistungen bei hoher Intelligenz und relativ bessere Rechtschreibleistungen bei schwacher Intelligenz – interessierte man sich nicht. SCHLEE bezeichnete diese „randständigen“ Fälle als „normal“. Wegen des nur mittleren Zusammenhangs von Intelligenz und Rechtschreibung sei nichts „Auffälliges“ an Diskrepanzen zwischen beiden Leistungen. Aus diesem Grunde sprach er sich dagegen aus, die Intelligenz als Maßstab für schulisches Lernen zu wählen, wie dies in der Legasthenie-Diagnose geschieht (SCHLEE 1978, S. 22).

Hier stellt sich natürlich die Frage, welches Merkmal anstelle der Intelligenz die Schullaufbahn eines Kindes vernünftigerweise bestimmen sollte? Diagnostisches Kriterium für das

Vorliegen einer Legasthenie ist eine Diskrepanz zwischen besserer Intelligenz und schwächerer Rechtschreibleistung. Eine Schullaufbahn, so argumentierten die Legasthenie-Forscher, dürfe nicht wegen der zu schwachen Rechtschreibung scheitern und etwa in die Sonderschule für allgemein Lernbehinderte oder vom Gymnasium herunter zur Hauptschule führen. Vielmehr müsse die Intelligenz eines Kindes – ungeachtet seiner Rechtschreibleistung – und andere, bessere Schulleistungen den Ausschlag bei Versetzungsentscheidungen und der Wahl einer Schullaufbahn geben.

Die Argumentation von SCHLEE lässt außer Acht, dass diskrepante Leistungsverhältnisse sich im Schulalltag und in der Schullaufbahn auswirken: Schüler mit besserer Rechtschreibung erreichen bekanntlich höhere Bildungsabschlüsse als Schüler mit schwächerer Rechtschreibung. Wenn beide Leistungen nur in mittlerem statistischem Zusammenhang stehen, also Fälle mit diskrepanten Leistungen (Rechtschreibung < IQ und Rechtschreibung > IQ) gar nicht so selten vorkommen wie früher angenommen, muss das Schulsystem nicht gerade dann darauf Rücksicht nehmen?

In unserer Untersuchung der Kieler Physiker sind wir den Weg der Neugier, den Weg des Physikers gegangen, den die Frage bewegte „Welche Informationen stecken in den Daten?“ Die auf diesem Weg gewonnenen Erkenntnisse gehen weit über das hinaus, was bei einer ersten Auswertung mit statistischen Methoden gefunden wurde. Die Ergebnisse finden sich im Teil 3 dieses Buches.

### 1.1.3 Fehlinterpretationen in der Psychologie

Die drei Geisteswissenschaften, die sich mit dem Menschen befassen – Pädagogik, Psychologie, Soziologie – stehen bekanntlich weit stärker als mathematisch-naturwissenschaftliche Disziplinen in der Gefahr, ihre Ergebnisse aus der Sicht von vorgefassten Annahmen zu interpretieren. Selbst dann, wenn in diesen Wissenschaften empirisch geforscht wird, können sich bei der Interpretation der gewonnenen Ergebnisse immer noch ideologisch fixierte „Meinungen“ durchsetzen, wie die Argumentationen von ANGERMAIER (1976), SCHLEE (1978) und anderer Referenten beim Rundgespräch der Deutschen Forschungsgemeinschaft „Zur Lage der Legasthenieforschung“ am 1. und 2. November 1976 gezeigt hat.

Es gibt bis heute Beispiele dafür, dass Schulpsychologen zugunsten der in ihrem Land schulbehördlich vertretenen Auffassungen eine Legastheniediagnose nicht stellen oder nicht erkennen. So heißt es beispielsweise im Gutachten eines Schulberaters: „Er ist also das, was man bisher einen ‚echten Legastheniker‘ nannte“. Und ein Fallbericht aus dem Fernsehkolleg des Deutschen Instituts für Fernstudien an der Universität in Tübingen beschreibt eine klassische Legasthenie: Der Schüler zeigt uneinheitliche Schulleistungen in der 4. Klasse. Er bleibt die nächsten zwei Schuljahre auf der Hauptschule, wo sich die Leistungen kaum bessern, so dass die Hoffnung, den Übergang zur Realschule zu schaffen, gering ist. Der mit dem CFT 20 gemessene IQ beträgt 141 – nur 4 von 1000 Gleichaltrigen erreichen so hohe Werte. In der Rechtschreibung unterlaufen dem Schüler dagegen viele und zum Teil unverständliche Fehler. Er liest im Lesetest sehr fehlerhaft und ungenau, im Fach Englisch beherrscht er z.T. sehr seltene Vokabeln, kann aber einfachste Wörter nicht übersetzen. Neben Schwächen – traditionellerweise wird hier die schwache „Konzentration“ genannt – hebt der Untersucher erfreulicher-

weise auch die besonderen Stärken hervor. Sie entsprechen nach unseren Erkenntnissen genau der mathematisch-technischen Begabungsstruktur:

*„Am besten sind bei ihm konstruktive Fähigkeiten, anschauungsgebundenes Denken und technische Begabung ausgeprägt.“* (WEINERT 1982, S. 33-41)

Die abschließende Diagnose lautet jedoch ganz anders: Wegen der totalen Überwachung, der Betreuung der Schularbeiten durch die Mutter und des ständigen zusätzlichen Übens übernehme der Junge zu wenig Verantwortung für sich selbst und seine eigene Leistung! Das bedeutet eine klare Schuldzuweisung an die Mutter, die von ihrem Kind vermeintlich zu viel erwartet – bei einem IQ, der weit über dem Durchschnitt liegt und selbst bei zurückhaltender Beurteilung dem Bereich der Hochbegabung zuzuordnen ist! Die Frage, warum die Mutter sich genötigt sah, in so starkem Maße kontrollierend einzugreifen, stellt sich für den Diagnostiker nicht. Die Empfehlungen an die Eltern lauten entsprechend: Betreuung durch einen Nachhilfelehrer und Wiederholung der ersten Realschulklasse, um „Kenntnislücken“ aufarbeiten zu können.

Auch das ist eine Fehlinterpretation: die Rechtschreibleistung wird nicht durch den „Erwerb von Kenntnissen“ sondern als psychomotorische Leistung aufgebaut. Regelwissen hat erst zusätzlich Bedeutung. Es kann in einem solchen Fall demnach nicht um die Beseitigung von „Kenntnislücken“ gehen, die etwa in einem Schuljahr zu beheben wären. Außerdem wird aus den falsch interpretierten Untersuchungsergebnissen ein falscher Rat abgeleitet: Der Junge soll nicht in die nächsthöhere Klasse aufsteigen, er wird sich mit jüngeren Schülern in einer Klasse finden und wird allein wegen seiner erheblich ausgeprägten Lese-Rechtschreibschwäche intellektuell unterfordert bleiben.

Oft schließt sich in den Erziehungswissenschaften bei Themen, die ohnehin ideologisch besetzt sind oder an denen – wie z.B. im Schulsystem – unterschiedliche Gruppen interessiert sind, ein Meinungsstreit an, bei dem Fakten weniger eine Rolle spielen als unverrückbare Überzeugungen.

In den Naturwissenschaften können sich solche Überzeugungen nicht lange halten. Wissenschaftler, die an den gleichen Fragestellungen interessiert sind, führen Nachuntersuchungen unter genau gleichen Bedingungen durch und die Ergebnisse der Erstuntersucher lassen sich dadurch bestätigen oder eben nicht. Dann ist dem Meinungsstreit ein Ende gesetzt, und man versucht herauszufinden, worauf die unterschiedlichen Ergebnisse zurückzuführen sind. Die Ursache kann in methodischen Fehlern liegen, aber auch in einem eingeschränkten Geltungsbereich der anfänglich gefundenen Ergebnisse. Auf jeden Fall lassen sich Widersprüche, die in den Daten verschiedener Untersucher aufgetaucht sind, durch Nachuntersuchungen und durch kontrollierte Veränderung der Versuchsbedingungen über kurz oder lang klären.

In der Psychologie werden Nachuntersuchungen unter gleichen Bedingungen nur selten durchgeführt. Dabei scheint das hinderlichste Problem das der Validität der Untersuchungstichprobe zu sein: Konnte der Nachuntersucher wirklich eine Personenstichprobe gewinnen, die derjenigen aus der Erstuntersuchung in den wesentlichen Merkmalen vergleichbar ist? Es hat sich gezeigt, dass Ergebnisse unterschiedlich ausfallen, wenn in der Erstuntersuchung z.B. Kinder aller Intelligenzbereiche, in einer Nachuntersuchungen aber nur mit Intelligenzquotienten über 90 oder 95 erfasst wurden. Verallgemeinert man die Ergebnisse aus solchen Untersuchungen z.B. auf „die Legasthenie“, dann gelangt man zu einer fehlerhaften Erklärung für das komplexe Problem.

In manchen Fällen lassen sich unterschiedliche Ergebnisse sehr einfach klären, sofern man in der betreffenden Fragestellung kompetent genug ist, um zu erkennen, dass die anscheinend einander widersprechenden Befunde sich in Wirklichkeit ergänzen. Das war z.B. in der Legasthenie-Diskussion der siebziger Jahre der Fall. Die allerschwächsten Rechtschreiber (Prozentrang  $\leq 5$ ), so ergaben die Gruppenuntersuchungen der jüngeren Forschergeneration, waren überwiegend soziokulturell benachteiligte Schülerinnen und Schüler mit relativ schwacher, wenngleich schon durchschnittlicher Intelligenz. Auf diese Gruppe aber hatten die ersten Legasthenieforscher ihr Interesse gerade nicht gerichtet, sondern auf Schülerinnen und Schüler, deren schwache Rechtschreibleistung nicht durch soziokulturelle Benachteiligung oder niedrige Intelligenz erklärt werden konnte.

Im Schulsystem müssen selbstverständlich beide Gruppen gesehen und berücksichtigt werden. Die wissenschaftlichen Ergebnisse der klassischen Legasthenieforschung und die der jüngeren Forscher ergänzen demnach einander. Man muss bei jeder Argumentation also zunächst einmal klar machen, auf welche Gruppe schwacher Leser/Rechtschreiber man sich bezieht. Allerdings kann sich der Meinungsstreit auch nach einem solchen Klärungsversuch noch fortsetzen, wenn die eine Seite den Vorschlag zur Unterscheidung von Teilgruppen gar nicht erst in Betracht zieht.

Bei ganz neu auftauchenden Ideen, die sich nicht im Labor oder durch Berechnungen klären lassen, kann allerdings auch in den Naturwissenschaften ein Meinungsstreit lange andauern. Das war z.B. im Fall des Fundes vom Neandertal so. Der als Wissenschaftler höchst renommierte VIRCHOW erklärte den Knochenfund als neuzeitlich und durch Rachitis verändert. Es hat einiger Jahrzehnte und weiterer Funde bedurft, bis der Neandertaler als eine Frühform des Menschen seine endgültige Einordnung fand.

Ein zweites Beispiel ist die Theorie Alfred WEGENERS von der Kontinentalverschiebung. Sie stieß auf heftige und lang anhaltende Kritik. Er hat die Anerkennung der Grundprinzipien seiner Idee nicht mehr erlebt. Ähnlich erging es SEMMELWEIS, der herausfand, dass die Wöchnerinnen im Kindbett starben, weil sie von Ärzten untersucht wurden, die vorher Leichen sezirt und danach ihre Hände nicht desinfiziert hatten!

## **1.2 Intelligenz – Begabung – Begabungsstruktur**

Intelligenz ist ein Phänomen, das vorwissenschaftlich und umgangssprachlich nur den Menschen mit höherer Intelligenz zugesprochen wird: Jemand ist intelligent oder unintelligent. Wissenschaftlich gesprochen ist jeder Mensch intelligent, nur gibt es unterschiedliche Ausprägungsgrade, die in einer Skala von weit unterdurchschnittlich bis weit überdurchschnittlich reichen. Nach früheren Auffassungen über eine Global-Intelligenz, die zu annähernd gleich hohen Leistungen in allen Bereichen befähigen sollte, wendet sich heute das wissenschaftliche Interesse unterschiedlichen „Intelligenzen“ zu, d.h. Befähigungen in verschiedenen Anforderungsbereichen.

### **1.2.1 Definitionen von Intelligenz und Begabung**

Intelligenz ist nicht direkt beobachtbar, sondern nur indirekt an intelligenten Verhaltensweisen ablesbar. Ein erstes Problem stellt bereits die Einigung darüber dar, was denn „intelligente Ver-

## 2 Theoretischer Teil: Der I-S-T als Messinstrument

### 2.1 Einleitende Überlegungen zum I-S-T

Der Intelligenz-Struktur-Test (I-S-T) kommt dem Bedürfnis der Menschen entgegen, Eignungen der einzelnen Personen für die verschiedenen Berufe, soweit bis jetzt möglich, nach Art der modernen exakten Naturwissenschaft zu beschreiben. Das bedeutet, die Zusammenhänge, die zu diesen Aussagen führen, messbar, wiederholbar und nachvollziehbar darzustellen.

Die Entwicklung unserer Naturwissenschaft hängt eng mit der Entwicklung des Messwesens zusammen und mit dem Bedürfnis, die Natur in ihre Grundbestandteile zu zerlegen, um so die Zusammenhänge zu erkennen und daraus Nutzen zu ziehen.

Auch vor der Entwicklung der modernen exakten Naturwissenschaft hatten die Menschen eine Vorstellung von der Natur, die sie umgab, in der sie leben und überleben mussten. Die kurze Lebenserwartung, die Belastung durch Erarbeitung des Lebensunterhalts und die Unfähigkeit, Erfahrungen wegen fehlender Möglichkeiten durch Messungen auf Richtigkeit zu überprüfen, führten zu einer vom heutigen Wissensstand aus gesehen archaischen Wissenschaft, in der zum Beispiel Steine wachsen (ELIADE), Gott zum Menschen spricht (Grundlage so mancher Religion) usw.. Das Hauptkriterium für die Güte dieser archaischen Wissenschaft war, dass sie es den Menschen ermöglichte, wenn auch unter Umständen auf Kosten des Einzelnen, als Art zu überleben. Für das Überleben war die Einfachheit der Vorstellungen von großem Vorteil, auch wenn die Richtigkeit darunter litt. Auch falsche Vorstellungen können Erfahrungen beschreiben und erklären und eine Hilfe für das Leben sein. Dass die Sonne sich um die Erde bewegte und nicht die Erde um die Sonne, war z.B. eine Vorstellung, die das Überleben der Menschen nicht beeinträchtigte.

Der zunehmende Wohlstand gerade in den letzten hundert Jahren ermöglicht es den Menschen, sich den Luxus zu leisten, das Messwesen immer weiter zu entwickeln und auszubauen und damit die Erfahrung immer mehr durch exaktes Wissen zu ersetzen. Erfahrung ist die nicht abgesicherte Form des Wissens. Da nach AMTHAUER (1970, S. 14 Mitte) bisher jeder Anwender des I-S-T andere Erfahrungen sammeln kann, denn „für die Interpretation selbst läßt sich keine verbindliche Regel aufstellen“, ist die wissenschaftliche Absicherung zumindest für dieses Messinstrument noch nicht weit genug gediehen. Für einen Physiker ist die Möglichkeit, dass jeder andere Erfahrungen sammeln kann, unerträglich, weil in diesem Fall dem Subjektiven und damit der Willkür keine Grenzen gesetzt werden. Leider ist es im allgemeinen so, dass ein Urteil um so sicherer gefällt wird, je geringer das Wissen des Urteilenden ist. Man passt die Welt dem eigenen Verständnis an. Wissen belastet, weil es dieses einfache Weltbild stört. Die „Tragik“ der modernen exakten Wissenschaft besteht nach Thomas Henry Huxley, in der „Ermordung“ schöner, auf persönlicher Erfahrung beruhender Vorstellungen durch „hässliche Tatsachen“.

In dieser Arbeit soll nun in dem eben aufgeführten Sinne versucht werden, die mit dem I-S-T gewonnenen Messwerte durch graphische Auswerteverfahren so weiterzuverarbeiten, dass die in den Messwerten enthaltenen, aber bisher unbekanntes, Tatsachen und Informatio-

nen erkennbar werden, die das allgemeingültige Wissen vermehren und die bisherigen Erfahrungen ersetzen.

Aus der Sicht des Physikers werden mit dem I-S-T, wie bei vielen anderen Messverfahren auch, der Testperson Signale in Form von zu lösenden Aufgaben angeboten. Als Testergebnis werden die Signalantworten in Form der Anzahl richtig gelöster Aufgaben in vorgegebener Zeit registriert. Auf das Testergebnis wirken viele Einflüsse ein, neben den zu testenden Eigenschaften der Testperson auch die Güte der Testaufgaben, das Verhalten der den Test durchführenden Person, das Testumfeld, aber auch Verständnisschwierigkeiten auf Grund besonderer fachlicher Vorbildung usw.. Aus den Folgen dieser Einflüsse setzt sich die Anzahl der gelösten Aufgaben als Messergebnis zusammen. Deshalb müssen die Nebeneinflüsse als systematische und zufällige Messabweichungen soweit möglich erfasst, analysiert und berücksichtigt werden. Dieses ist eine weitere wesentliche Aufgabe dieser Arbeit.

Gemessen wird die Anzahl der gelösten Aufgaben in vorgegebener Zeit. Der reziproke Wert ist die Bearbeitungszeit pro gelöster Aufgabe und damit eine Zeitgröße. Damit sind die Grundbegriffe der Messtechnik nach DIN 1319 anzuwenden.

Die Fragestellungen für Psychologen als Praktiker einerseits und für den Naturwissenschaftler andererseits sind sehr unterschiedlich. Für den Psychologen als Diagnostiker ist der I-S-T das Werkzeug, das ihm Auskunft über den Menschen gibt. Der Naturwissenschaftler dagegen untersucht den I-S-T unter allen nur erdenklichen Fragestellungen. Für ihn stehen nicht die Leistungsmerkmale des Einzelnen oder der Gruppe im Vordergrund des Interesses, sondern die Frage, welche Auskunft der I-S-T über sich selbst und über den Menschen allgemein überhaupt liefern kann. Erst dieses Wissen erlaubt es dann dem Psychologen, den I-S-T für seine Fragestellungen optimal anzuwenden.

## **2.2 Die Anwendung des I-S-T**

Für die Durchführung des I-S-T gibt AMTHAUER die folgenden Anweisungen: „Es ist ein geeigneter Raum zu benutzen und die Zeit so zu wählen, dass die Teilnehmer frisch und leistungsfähig sind. Mit wenigen Worten wird der Zweck der Durchführung des Tests erklärt. Es wird versucht, eine Atmosphäre der Aufnahmebereitschaft und Lockerheit zu schaffen. Um vergleichbare Ergebnisse mit denen anderer Personen zu erhalten, werden alle Anweisungen, Erläuterungen und Beispiele vorgelesen. Die Teilnehmer benutzen Testhefte, in denen sie die mündlich gegebenen Anweisungen mitlesen können und in denen die Aufgaben stehen, sowie Antwortbögen. Die zu bearbeitenden Aufgaben sind aufgeteilt in 9 Aufgabengruppen zu je 20 Aufgaben. Für jede Aufgabengruppe ist die Durchführungzeit genau einzuhalten. Die Gesamtdurchführungszeit beträgt ungefähr 90 Minuten“.

### **2.2.1 Die Aufgabenstellung und die geprüften Funktionen**

Jede einzelne Aufgabengruppe stellt an die Teilnehmer des Tests ganz spezifische Anforderungen, die so beschrieben werden können:

- SE (Satzergänzung)  
Ein unvollständiger Satz muß sinnvoll ergänzt werden durch Ankreuzen einer von fünf Alternativlösungen.
- WA (Wortauswahl)  
Von fünf vorgegebenen Wörtern muß dasjenige angekreuzt werden, das nicht zu den übrigen passt.
- AN (Analogien)  
Aus fünf vorgegebenen Wörtern soll das richtige angekreuzt werden, das parallel zu einem anderen passt, zum Beispiel „Wald zu Baum wie Wiese zu ...“?
- GE (Gemeinsamkeiten)  
Von 6 Wörtern gehören 2 zu einer wesentlichen Gemeinsamkeit, diese zwei Worte sind anzukreuzen.
- RA (Rechenaufgaben)  
Eingekleidete Rechenaufgaben verschiedenster Art, die die vier Grundrechenarten und den Dreisatz voraussetzen, sind zu lösen, und das Ergebnis ist anzukreuzen.
- ZR (Zahlenreihen)  
Die Gesetzmäßigkeit von Zahlenreihen ist zu erkennen und die nächstfolgende Zahl anzukreuzen.
- FA (Figurenauswahl)  
Von 5 vorgegebenen geometrischen Figuren ist jeweils eine verschieden aufgeschnitten und ihre Teile verstreut angeordnet. Es ist anzukreuzen, welche Teile welche Figur ergeben.
- WÜ (Würfelaufgaben)  
Je einer von fünf vorgegebenen Würfeln mit unterschiedlichen auf den drei sichtbaren Seiten angegebenen Markierungen wird bei den einzelnen Aufgaben in gedrehter Lage wiedergegeben. Es ist anzukreuzen, um welchen es sich handelt.
- ME (Merkaufgaben)  
Es sind mehrere Serien von Wörtern, die innerhalb der Serie begrifflich zusammenhängen, auswendig zu lernen, und nach einer vorgegebenen Pause ist innerhalb einer begrenzten Zeit anzukreuzen, welcher der je 5 Vorschläge mit einem der gelernten Wörter zusammenhängt.

Die Ergebnisse, wie viele der je 20 Aufgaben aus den neun Aufgabengruppen richtig gelöst wurden, geben Auskunft darüber, welche Funktion mit welcher Leistungsfähigkeit bei der Lösung der Aufgaben beteiligt war. Diese aufgabengruppenspezifischen Funktionen beschreibt AMTHAUER wie folgt:

- SE (Satzergänzung)  
**„Urteilsbildung“**, „common sense“, Akzent im „Konkret - Praktischen“, „Wirklichkeits-sinn“, „Selbständigkeit im Denken“.
- WA (Wortauswahl)  
**„Erfassen von sprachlichen Bedeutungsgehalten“**, „Sprachgefühl“, „induktives sprachliches Denken“, „Einfühlungsfähigkeit“, „rezeptive“ Komponenten.

AN (Analogien)

„**Kombinationsfähigkeit**“, „Beweglichkeit und Umstellfähigkeit im Denken“, „Erfassen und Übertragen von Beziehungen“, „Klarheit und Folgerichtigkeit im Denken“, „Widerstand gegen Ungefährlösungen“ – eine, vielleicht die bedeutsamste Voraussetzung für wissenschaftliche Studien.

GE (Gemeinsamkeiten)

„Sprachliche **Abstraktionsfähigkeit**“, „Begriffsbildung“, „sprachlogisches Denken“.

ME (Merkaufgaben)

„**Merkfähigkeit**“, „gelernte Wörter behalten können“, „längerfristiges Behalten“, „Gedächtnis“.

RA (Rechenaufgaben)

„**Praktisch - rechnerisches Denken**“, „sachlogisches, mathematisches Denken“, „reasoning“, „schlussfolgerndes Denken“.

ZR (Zahlenreihen)

„**Theoretisch - rechnerisches Denken**“, „induktives Denken mit Zahlen“, „Beweglichkeit und Umstellfähigkeit im Denken“, „rhythmische“ Komponenten.

FA (Figurenauswahl)

„**Vorstellungsfähigkeit**“, „Vorstellungsreichtum“, „anschaulich - ganzheitliches Denken“, „gestaltend-konstruktive“ Komponenten.

WÜ (Würfelaufgaben)

„**Räumliches Vorstellen - Können**“, „technisch - konstruktive“ Komponenten, „analytische Momente“ können beteiligt sein; von konventioneller Bildung weitgehend unabhängig.

### 2.2.2 Der Gang der Auswertung

Die einzelnen Personen reagieren bei der Lösung der je 20 Aufgaben der neun Aufgabengruppen mit unterschiedlicher Leistung. Diese Unterschiede hängen zum einen von der individuellen Leistungsfähigkeit bei der Lösung der Aufgaben der einzelnen Aufgabengruppen ab, deretwegen der Test ja durchgeführt wird, zum anderen aber auch vom biologischen Alter der am Test teilnehmenden Person. Es ist daher notwendig, die Anzahl der richtig gelösten Aufgaben der Aufgabengruppe  $i$ , die mit  $RW_i$  (Rohwert der Aufgabengruppe  $i$ ) bezeichnet wird, und den Gesamtrohwert  $RW$  als Summe aller richtig gelösten Aufgaben aller neun Aufgabengruppen für eine möglichst eng begrenzte Altersgruppe zu „normieren“. Normieren bedeutet, in eine Form bringen, die es ermöglicht, die Werte verschiedener Personen unterschiedlichen Alters und unterschiedlicher Berufsgruppen in Beziehung zu setzen. Erst dann ist die im jeweiligen Normwert ausgedrückte Leistungsfähigkeit in den verschiedenen geprüften Funktionen unabhängig von personenspezifischen Daten, und es kann angegeben werden, wie viel Prozent der Bevölkerung diese Leistung in gleicher Höhe erbringen kann. Mit den Normwerten können dann auch die Leistungen verschiedener Personen und verschiedener Altersgruppen miteinander verglichen werden.

Zu diesem Zweck sind von AMTHAUER eine Reihe von Begriffen eingeführt oder benutzt worden, die aus den Rohwerten durch einfache Gleichungen errechnet oder durch teils altersabhängige, teils altersunabhängige Tabellen ermittelt werden können. Wie diese Begriffe zusammenhängen und hergeleitet werden können, ist in Abb. 7 zu sehen.

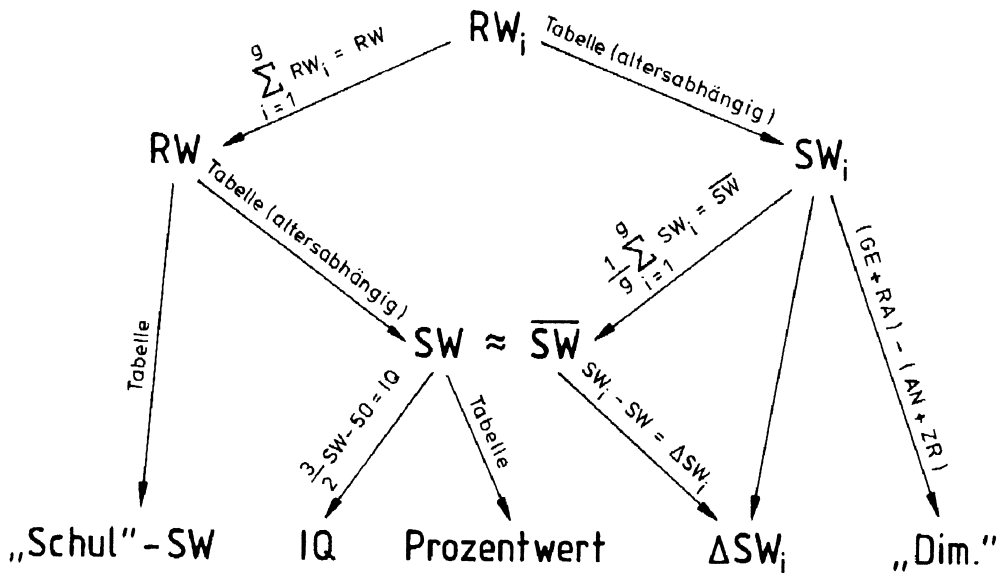


Abbildung 7: Die Aussagemöglichkeiten aus den mit dem Intelligenz-Struktur-Test gewonnenen Ergebnissen, dargestellt in einem Flussdiagramm. Hierin bedeutet  $\Sigma$  das Summenzeichen für die Summe aller  $RW$  oder  $SW$  mit dem laufendem Index  $i$  von  $i = 1$  bis  $i = g$ , in diesem Falle  $g = 9$ .

In Abb. 7 werden folgende Begriffe benutzt:

$RW_i =$  Roh-Wert = Anzahl der richtig gelösten Aufgaben der Aufgabengruppe  $i$   
 $SW_i =$  (absoluter) Standardwert: tabellierte Größe, die auf Grund der Normierung des Tests dem Rohwert einer Person unter Berücksichtigung ihres Alters zugeordnet wird, für die Aufgabengruppe  $i$ . Durch den Standardwert werden die Leistungen der Person mit denen anderer Personen unterschiedlichen Alters vergleichbar.

$\overline{SW} = \frac{1}{9} \sum_{i=1}^9 SW_i =$  individuelles arithmetisches Mittel der (absoluten) Standardwerte einer Person in den neun Aufgabengruppen.

$\Delta SW = SW_i - \overline{SW} =$  relativer Standardwert = Abweichung eines (absoluten) individuellen Standardwertes der Aufgabengruppen  $i$  von dem arithmetischen Mittel der (absoluten) individuellen Standard-Werte aller neun Aufgabengruppen.

$Dim. = (SW_{GE} + SW_{RA}) - (SW_{AN} + SW_{ZR}) = (GE + RA) - (AN + ZR) =$  Dimension „Festigung - Flexibilität“. Es handelt sich hier um einen relativen individuellen Standardwert, der als Hinweis dafür dient, ob die Person eher zur Flexibilität oder eher zur Festigung neigt. Mit dem Alter findet bei Erwachsenen eine Umstrukturierung der Intelligenz von größerer „Flexibilität“ zu größerer „Festigung“ statt (AMT-HAUER 1970, S. 14).

$RW = \frac{1}{9} \sum_{i=1}^9 RW_i =$  Gesamt-Rohwert = Summe der Rohwerte der neun Aufgabengruppen

„Schul“- $SW =$  Schul-Standard-Wert: Tabellierte Größe in Abhängigkeit vom Gesamtrohwert einer Stichprobe von Personen mit entsprechendem Schulabschluss. Der Schul-Standardwert

wird unabhängig vom Alter ermittelt. Durch den Schul-Standardwert wird die Leistung der Person mit den Leistungen der Personengruppe mit gleichem Schulabschluss in Beziehung gesetzt.

SW = (absoluter) Gesamt-Standardwert: tabellierte Größe, die dem Rohwert unter Berücksichtigung des Alters der Person zugeordnet wird

IQ =  $\frac{2}{3} SW - 50 = \text{Intelligenzquotient.}$

Prozentwert (PW) oder Prozentrang (PR): Tabellierte Größe in Abhängigkeit von dem Gesamtstandardwert der getesteten Person. Der Prozentwert gibt an, wie viel Prozent der Bevölkerung (über alle Altersgruppen hinweg) einen niedrigeren oder gleichen Gesamtstandardwert haben als die getestete Person. Die Tabellen für verschiedene Normskalen, wie IQ-Skala, SW-Skala, PR-Skala u.a. gehen auf die standardisierte Normalverteilung der Ursprungsdaten einer Eichstichprobe zurück. Daher können alle diese Werte einer Vergleichstabelle entnommen und ineinander umgewertet werden.

### 2.2.3 Die erzielbaren Ergebnisse

Der I-S-T ist nach AMTHAUER (1970, S. 28) „in Bezug auf Gültigkeit, Zuverlässigkeit und Objektivität gesichert“. Das ist Voraussetzung dafür, Untersuchungen durchzuführen und Ergebnisse zu gewinnen. Aber ist sie wirklich gegeben?

„Intelligenz wird“ nach AMTHAUER (1970, siehe Seite 7 oben) „aufgefasst als eine Sonderstruktur im Gesamt der Persönlichkeitsstruktur eines Menschen.“ Sie sei „eine strukturierte Ganzheit von seelisch-geistigen Fähigkeiten, die in Leistungen wirksam werden“ (1973, S. 7). Intelligenz bleibe eng mit dem Antriebs-, Willens- und Gemütsleben verbunden und erweise sich auch nicht als identisch mit intellektuellen Prozessen allein.

AMTHAUER beschreibt den I-S-T als ein „Verfahren zur Bestimmung der Intelligenz“, mit dem Untersuchungen durchgeführt und Ergebnisse gewonnen werden können. Der Test ermöglicht seiner Meinung nach

- 1) einen „zuverlässigen Befund über das Intelligenz - Niveau“
- 2) einen „Einblick in die Struktur der Intelligenz“.
- 3) eine exakte Methode, die das Problem der
  - a) Intelligenz
  - b) Begabung
  - c) psychologisch-diagnostischen Untersuchungen lösen kann.
- 4) den Schwerpunkt der Begabung deutlich zu erkennen,
- 5) Eignungsuntersuchungen und Beratung durchzuführen für
  - a) Berufswahl
  - b) Veränderungen im Beruf
  - c) Umschulung

Mit dem I-S-T werden Ergebnisse gewonnen, die die Grundlage für anschließende Interpretationen sind:

- 1) Gesamtstandardwert
- 2) Schul-Standardwert
- 3) Festigung – Flexibilität

- 4) Intelligenz-Profil
- 5) Abweichung der Standardwerte der einzelnen Aufgabengruppen vom mittleren Standardwert

Nach AMTHAUER läßt sich für die Interpretation selbst keine verbindliche Regel aufstellen (1970, S. 14), was aus der Sicht eines Physikers noch zu bestreiten sein wird.

Die Ergebnisse der individuellen Leistung in den neun Aufgabengruppen ermöglichen (AMTHAUER S. 14 Mitte):

- 1) die Beurteilung des Intelligenz-Niveaus in Beziehung
  - a) zur schulischen Entwicklung
  - b) zur beruflichen Entwicklung
  - c) zur sozialen Herkunft
  - d) zum derzeitigen Milieu des Probanden
  - e) zum Vergleich mit den Leistungen gleichaltriger Personen
- 2) den Vergleich mit Personen gleicher Schulausbildung
- 3) die Beurteilung der Dimension = Flexibilität - Festigung (eher gefestigt - eher flexibel), und darüber hinaus der geistigen Entwicklungsfähigkeit
- 4) die Interpretation des individuellen Profils im Vergleich mit Berufsprofilen usw..

Die Ergebnisse in den ersten vier Aufgabengruppen (SE, WA, AN, GE) lassen je nach Profil zwei Interpretationen zu. Hat das Ergebnisprofil die Form eines:

M dann handelt es sich um eine mehr sprachlich theoretische Begabung

W dann handelt es sich um eine mehr praktische Begabung. Gemeint ist die naturwissenschaftlich-technische Begabung.

Vermutlich wird der Standardwert für ME (oder unter Auslassung von ME der Standardwert für RA?) hier mit einbezogen, da sich aus den ersten vier Werten (AMTHAUER S. 15) kein M und W vollständig ergibt. Zudem werden diese Angaben durch die Strukturkurven in AMTHAUER 1980 relativiert

Für die Eignung ist nicht die Höhe des Intelligenzquotienten maßgebend, sondern das Intelligenz-Profil. Die wesentlichsten Eigentümlichkeiten des I-S-T - Profils werden nicht erst in der beruflichen Praxis erworben, sondern sind schon vorher vorhanden. Liegt Artverwandtschaft zwischen Berufen vor, so kommen, wenn auch auf verschiedenem Niveau, ähnliche Profile vor (zum Beispiel: Diplom - Ingenieur und technischer Handwerker). Größere Schwankungen in der I-S-T - Struktur, die nicht mehr allein als Strukturmomente verständlich sind, können auf neurotische Tendenzen oder – in schwächer ausgeprägten Fällen – auf eine relativ labile Persönlichkeitsstruktur hinweisen:

*„Selbstverständlich können die Tableaus durch diejenigen Schulabschluss- bzw. Berufsgruppen- Profile erweitert werden, die in der Praxis des Fachpsychologen vorkommen. ...Über den Gewinn an zusätzlicher Information wird man wahrscheinlich erstaunt sein.“*  
(AMTHAUER 1970, S. 17 und 20)

Vergleiche von Intelligenz-Struktur-Profilen machen deutlich,

*„warum mit dem Intelligenz-Niveau über die Intelligenz eines Menschen weniger ausgesagt wird als durch Ermittlung des ihm eigentümlichen Profils mit Einblick in seine Struktur.“ (AMTHAUER 1970, S. 21 unten)*

### 2.2.4 Grenzen der Aussagefähigkeit

Ein mittlerer Prozentwert oder Standardmittelwert der Intelligenzverteilung der Bevölkerung oder der 1-Sigma-Wert trifft nur für Europäer zu. Andere Völker haben andere Verteilungen, deren Maximum (bei Verwendung des gleichen Testinstruments) nicht bei 100 liegen muß (SCHILCHER S. 281 - 283). Um zu reproduzierbaren und zu vergleichbaren Aussagen zu kommen, muß die Auswertung der Testfragebögen immer nach den gleichen Kriterien erfolgen. Jede richtige Lösung zählt einen Punkt bei der Rohwertermittlung. Weitere Markierungen neben der richtigen in derselben Zeile des Antwortbogens machen die Lösung dieser Aufgabe ungültig. Die trotzdem auftretenden Unsicherheiten in den ermittelten Ergebnissen werden von AMTHAUER wie folgt beschrieben:

*„Hierbei ist zu beachten, dass man <individuellen> Unterschieden zwischen den in den verschiedenen Aufgabengruppen erreichten Standardwerten, die weniger als 0,5 Sigma = 5 Standardwertpunkte betragen, nicht zu viel Gewicht beimessen sollte und sie als weniger sicher betrachten muß.“ (AMTHAUER 1970, S. 15)*

Diese Auffassung wird weiter unten kritisch zu bewerten sein.

Da bei der „Dimension = Festigung – Flexibilität“ die Differenz aus einer Summe von Werten gebildet wird, die mit Messunsicherheiten behaftet sind, ist das Ergebnis mit einer nach der Fehlerrechnung berechneten noch größeren Messunsicherheit behaftet. So sind nach AMTHAUER bei der „Dimension = Festigung – Flexibilität“ Werte unter  $\pm 10$  weniger sicher. In den Untersuchungen von Brocke et al. (1998, S. 94) hat sich erwiesen, dass diese Dimension, vor allem ihre Korrelation mit dem Alter, sich nicht nachweisen ließ. Wir haben in unseren Auswertungen daher auf die Berücksichtigung dieser Dimension verzichtet.

Bei der Bildung relativer Standardwerte durch Differenzbildung zwischen (absoluten) individuellen Standardwerten der Aufgabengruppe  $i$  und dem arithmetischen Mittel der (absoluten) individuellen Standardwerte aller neun Aufgabengruppen heben sich die systematischen Messabweichungen, die alle neun Standardwerte in Größe und Vorzeichen in gleichem Maße betreffen, auf. Der tatsächliche oder Absolutwert ist viel anfälliger gegenüber Einflüssen von außen wie Erfolg, Indisponiertheit, Umwelteinflüssen usw., als die Relativwerte. So wie eine Balkenwaage unabhängig von den Umwelteinflüssen (zum Beispiel auf dem Mond wie auf der Erde) stets den gleichen Wert oder Betrag der Masse anzeigt, so ergibt sich immer die gleiche Intelligenzstruktur, egal welche Faktoren die Messung insgesamt beeinflussen – nur die absolute Höhe ist dann verschieden.

Im folgenden werden die Testlösungen daraufhin analysiert, welche Informationen sie enthalten. Die an die Probanden gestellten Fragen selbst werden nicht untersucht. Indirekt kann die Anzahl der richtig beantworteten Fragen auch, ohne dass der Auswertende die Fragen im einzelnen kennt, Auskunft über die Qualität der einzelnen Aufgaben geben. Wenn Testpersonen z.B. eine Aufgabe nicht lösen, obwohl sie nachfolgende, schwierigere Aufgaben lösen können, dann spricht das dafür, dass die fragliche Aufgabe missverständlich formuliert ist. Es kann

sein, dass die Formulierung dem Verständnis geisteswissenschaftlich geschulter Personen entgegenkommt, so dass der Schwierigkeitsgrad dieser Aufgabe für sie angemessen ist. Wer aber die Aufgabe als naturwissenschaftlich vorgebildeter Mensch liest, kann die Formulierung entweder nicht verstehen (und dann die Lösung nicht finden) oder er löst sie im Sinne der Absicht der Testkonstrukteure falsch, aus spezialisiertem naturwissenschaftlichen Wissen jedoch richtig. Dieses Problem ist in unserer Untersuchung mehrfach aufgetreten (siehe Zitat von MAIER-LEIBNITZ: Die Sprache der Physik in: DUZ, 19, 1982, S. 22-21).

## 2.3 Fragen an den I-S-T

Jemand, der andere beurteilen muß, braucht Entscheidungshilfen, so auch bei Berufungen für ein Amt, Einstellungen, Vergabe von Zensuren, Eignungsurteilen usw.. Vor allem im beruflichen Bereich nützen alle Menschen gern die Hilfe, die Tests, Zensuren etc. geben, schon um sich für die Beurteilung auch rechtfertigen zu können. Objektive quantitative Angaben, wie sie der Intelligenzquotient zu geben scheint, sind hilfreich und erwünscht.

Da die Zahlenwerte des Intelligenz-Quotienten um 100 streuen und in ganzen Zahlen angegeben werden, wird eine Genauigkeit von um 1% suggeriert, die die Objektivität des Intelligenz-Quotienten noch zu erhöhen scheint. Wenn auch nicht jeder vorgibt, genau zu wissen, was Intelligenz ist, so meint doch jeder, eine detaillierte Vorstellung davon zu haben. Insofern kann der erhobene Intelligenzquotient eine Person auch stigmatisieren.

Ist der Intelligenzquotient aber für sich allein aussagekräftig für die Leistungsfähigkeit einer Person? Ist der Intelligenzquotient als Entscheidungshilfe geeignet? Oder ist die Gesamtbegabung, die der IQ suggeriert, ein leerer Begriff, und allein die Leistungsfähigkeit in einzelnen Leistungsbereichen, also die Intelligenzstruktur als ein Maß der Eignung für Berufswege sinnvoll und aussagekräftig?

Wenn die Intelligenz nur durch eine einzige Zahl, den IQ beschrieben wird, so kann diese Zahl nur als eine undifferenzierte Aussage, allenfalls als grobe Orientierung gelten. Die Intelligenzstruktur aus neun Zahlen ermöglicht demgegenüber in den verschiedenen Kombinationen der neun Zahlen miteinander und untereinander komplexere und differenzierte Aussagen über die zu beurteilende Person. Das sollte mehr genutzt werden!

Was mit dem Intelligenz-Struktur-Test gemessen wird, geht aus einer Analyse des mit dem I-S-T gewonnenen Zahlenmaterials hervor. An die Testpersonen werden Fragen gestellt. Die Testperson erscheint gewissermaßen als „schwarzer Kasten“, das heißt, der Mechanismus, nach dem sie reagiert, ist nicht direkt erkennbar und schwer beschreibbar. Die Testperson reagiert auf die Fragen (20 zu lösende Aufgaben pro Aufgabengruppe), die die *Ursache* für ihre Reaktion sind, und liefert als Reaktion in vorgegebener Zeit, d.h. Antworten in Form der Anzahl gelöster Aufgaben pro Aufgabengruppe, die als *Wirkung* auf die Fragen aufgefasst werden können (DIN 1319, Teil 3, Anhang: Ursache, Wirkung). Die Ergebnisse der einzelnen Aufgaben lassen sich untereinander und in Verbindung mit weiteren Ursachen, wie Alter, Ausbildungsstand, Beruf usw., in verschiedenster Form graphisch darstellen. Da für alle Personen die Anzahl der zu lösenden Aufgaben, deren Inhalt sowie die Lösungszeit pro Aufgabengruppe dieselbe ist, ist die Ursache für alle Personen stets dieselbe. Was variiert, sind die unterschied-

# 3 Praktischer Teil:

## Die Untersuchung und ihre Auswertung

Im praktischen Teil dieser Arbeit wird über die Untersuchung mit dem Intelligenzstrukturtest im Institut für Angewandte Physik berichtet.

### 3.1 Beschreibung der untersuchten Stichprobe

Im Frühjahr 1973 waren an dem Institut für Angewandte Physik der Universität Kiel 34 Angehörige (Studenten mit Vordiplom, Diplomphysiker, Promovierte und Professoren) bereit, an einem Intelligenztest teilzunehmen und ihre Abiturnoten anzugeben. Die Untersuchung wurde unter Decknamen anonym durchgeführt. Anlass war, wegen einer zu erwartenden Einführung des Numerus Clausus für Physik in Kiel, neue Eingangskriterien zu erkunden.

Die Teilnehmer waren alle männlich, weil zu dieser Zeit keine Frau im Institut arbeitete. Studenten mit der Vordiplomnote gut oder sehr gut konnten 1973 auch ohne Diplom promoviert werden, was bald danach nicht mehr möglich war. Staatsexamenskandidaten mußten eine in der Prüfungsordnung nicht vorgesehene Staatsexamensvorprüfung in Physik und Mathematik ablegen oder das Vordiplom vorlegen.

Die am Test teilnehmenden Personen waren ausnahmslos persönlich bekannt, und so konnte darauf geachtet werden, dass alle Teilnehmer, deren Testergebnisse zur Auswertung kamen, sich mindestens durch das Vordiplom als erfolgreich im Fach Physik erwiesen hatten oder darüber hinaus die für das Staatsexamen obligatorische Staatsexamensvorprüfung absolviert hatten und mit einer Staatsexamensarbeit befasst waren. Ein nichtwissenschaftlicher Mitarbeiter hat auf eigenen Wunsch am Test teilgenommen. Seine Werte wurden ausgesondert und nicht berücksichtigt. Da als Hochschulabschluss nur nach dem Vordiplom, Diplom oder der Promotion gefragt wurde, konnten diejenigen, die nur die Staatsexamensvorprüfung bestanden hatten, keine Angaben machen. Bei diesen drei Personen steht deshalb „ohne Angabe“.

Diejenigen, die am Test teilnahmen, waren wie folgt, am Institut beschäftigt:

- 12 mit Vordiplom als Hilfsassistenten, Diplomanden oder Doktoranden,
  - 1 mit Staatsexamensvorprüfung,
  - 5 mit Diplom als Doktoranden oder als wissenschaftliche Angestellte,
  - 13 mit Promotion als wissenschaftliche Angestellte oder Beamten und
  - 3 ohne Angabe zum Hochschulabschluss, vermutlich Staatsexamenskandidaten.
- 34 Summe

Das Institut für Angewandte Physik bestand zu der Zeit aus:

- 4 hauptamtlichen Professoren AH3/AH4
- 9 wissenschaftlichen Beamten A13 bis A15
- 12 wissenschaftlichen Angestellten BAT Ib/IIa
- 29 Doktoranden, Diplomanden, Staatsexamenskandidaten

11 wissenschaftlichen Hilfskräften

65 Summe

Mithin haben mehr als 50% der Institutsangehörigen an der Untersuchung teilgenommen.

In Abb. 8 ist die Alterszusammensetzung der am Test beteiligten Personen dargestellt. Die Altersverteilung zeigt ein Maximum der Examenskandidaten bei 25 Jahren und ein breites Maximum bei 31-34 Jahren von wissenschaftlichen Mitarbeitern.

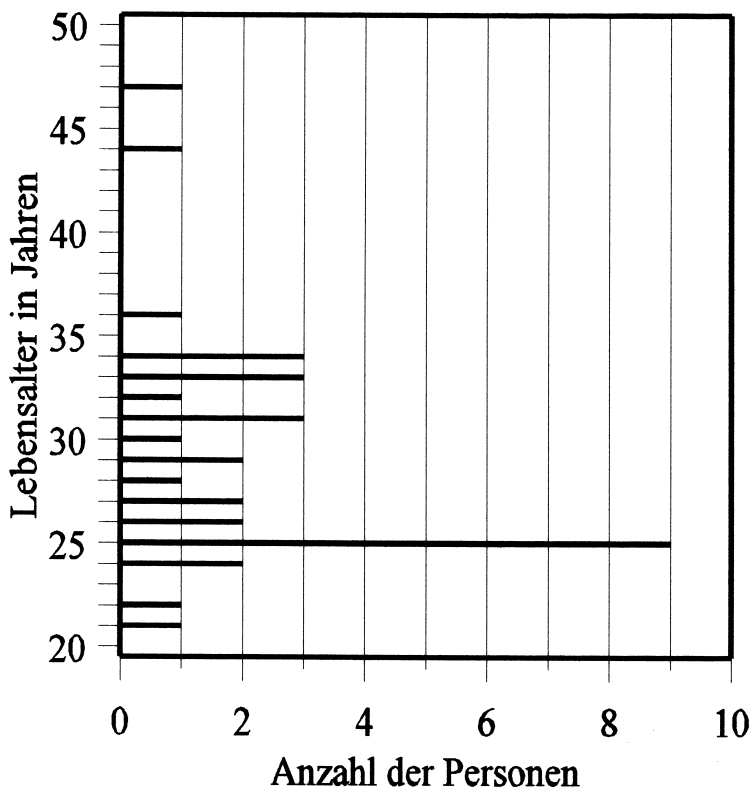


Abbildung 8: Altersverteilung der am Test teilgenommen habenden Personen

Mithin liegt der Modalwert (der häufigste Wert) bei 25 Lebensjahren, der Median (Zentralwert, 50% - Wert) bei 28 Jahren (DUMMER 1981).

Eine Reihe von Institutsangehörigen waren zum Zeitpunkt der Untersuchung nicht im Institut anwesend, weil sie in der vorlesungsfreien Zeit zu Tagungen gefahren waren oder Urlaub hatten. Eine weiterer Teil der Institutsangehörigen verweigerte die Teilnahme an der Untersuchung. Als Begründung wurden eine Reihe von Befürchtungen genannt, wie:

Die Person mag ihre Persönlichkeit nicht preisgeben.

Sie hat Angst, sich selbst, ihre innerste Persönlichkeit kennenzulernen.

Sie befürchtet Missbrauch der erhobenen Daten.

Sie fürchtet ein Versagen.

Sie möchte den Intelligenzquotienten lieber nicht erfahren, aus der Befürchtung heraus, ihr könne „Dummheit“ bestätigt werden.

Sie hat Prüfungsangst, der sie entfliehen will.

Diese Befürchtungen wurden verkläusuliert durch Argumente wie: Intelligenztests sind nicht aussagefähig, nicht geeignet, den Beruf zu beschreiben, können die Intelligenz nicht gerecht wiedergeben.

Von den 34 männlichen Personen konnten 27 vergleichbare deutsche Abiturnoten vorlegen, während die anderen 7 auf anderem Weg die Studienberechtigung erlangt hatten, nämlich auf dem beruflichen Bildungsweg, durch ein Abitur in der ehemaligen DDR, durch ein Notabitur während des Krieges oder ein Abitur, das im Ausland abgelegt worden war.

### 3.1.1 Aufgabenstellung des Instituts

Die Forschungsarbeiten des Instituts für Angewandte Physik der Universität Kiel 1973 behandeln Messverfahren und Analysetechnik in der Physik und ihre Anwendung in nichtphysikalischen, experimentell arbeitenden naturwissenschaftlichen Disziplinen wie Ozeanographie, Biologie, Medizin und physikalischer Chemie. Den Arbeitsgruppen gemeinsam ist die Beschäftigung mit elektronischer Mess- und Datentechnik, wobei anwendungsorientierte Fragen der Messtechnik, Schaltungstheorie, Systemtheorie und Datenerfassung und Interpretation der Messwerte selbst Gegenstand der Forschung sind.

Physikalische Messtechnik in Anwendung auf nichtphysikalische Gebiete schließt die Frage ein: Was ist messbar? Es interessiert der physikalische Hintergrund der Messwerte, die Erfassung systematischer Fehler, die Frage, was mit den gewonnenen Messwerten ausgesagt werden kann, die Projektierung und Entwicklung neuer Messverfahren und die Erschließung neuer Erkenntnisse auf dem jeweiligen nichtphysikalischen Gebiet.

### 3.1.2 Einmaligkeit der Messung

Jedes physikalische Institut hat sich auf bestimmte Arbeitsgebiete spezialisiert, die es mit den dem Institut eigenen Arbeitsweisen löst. Daher kann davon ausgegangen werden, dass die Mitglieder verschiedener physikalischer Institute im Mittel unterschiedliche Begabungsstrukturen aufweisen, also Varianten der Physik-Begabungsstruktur. Es ist anzunehmen, dass die Anforderungen der theoretischen Physik und die der experimentellen Physik sich durchaus unterscheiden und Personen mit unterschiedlicher Begabungsstruktur benötigen. Selbst die mittlere Begabungsstruktur der Angehörigen ein und desselben Institutes kann sich im Laufe der Jahre ändern, da sich die Arbeitsrichtung und die Arbeitsmethoden auf Grund der sich ändernden wissenschaftlichen Fragestellungen und der technologischen Entwicklung der Messgeräte und der Auswertemöglichkeiten ändern und den Gegebenheiten anpassen. Das führt dazu, dass Personen mit etwas anderer Begabungsstruktur sich angezogen fühlen und im Institut wegen der veränderten Aufgabenstellungen auch benötigt werden. Unsere Messung mit dem I-S-T ist aus diesem Grunde einmalig. Eine Wiederholung zum heutigen Zeitpunkt brächte unter Umständen andere Ergebnisse.

Ein weiterer Grund für die Einmaligkeit der Messung ist die Art der Selektion der Schule. Die Auslesebedingungen vor allem in Gymnasien haben sich in den vergangenen dreißig

Jahren mehrfach verändert. So fallen offenbar Schüler mit weniger ausgeprägter sprachlicher Begabung in Gesamtschulen nicht so schnell der Auslese zum Opfer wie in Gymnasien. Wer ausgeprägte nichtsprachliche Begabungen hat, kann zudem heute über die Fachoberschule und ein begonnenes Studium an einer Fachhochschule bei entsprechend guten Leistungen ohne Abitur in die Universität gelangen, was früher bei einem solchen Bildungsweg nur über das Abitur an einem Abendgymnasium oder mit mittlerer Reife und Examen an der Fachhochschule mit der Note „gut“ und besser möglich war. Seit den siebziger Jahren hat sich zudem der prozentuale Anteil der Abiturienten an der Altersgruppe wesentlich erhöht.

Es ist anzunehmen, dass eine Reihe für die Physik begabter Schüler wegen der Ausleseverhältnisse in den fünfziger Jahren nicht zum Abitur gelangten und nicht Physik studieren durften. Nur wer mindestens befriedigende Leistungen in den sprachlichen Fächern erreichte, den Sprachbegabten also nicht nachstand, konnte das Abitur erreichen. Bei entsprechender Begabungsstruktur – mathematisch-naturwissenschaftliche Begabung ausgeprägter als sprachliche – wären damit automatisch diejenigen Physikbegabten zum Studium gelangt, die im nichtsprachlichen Bereich hochbegabt waren. So könnten Physiker der Stichprobe von AMTHAUER im Vergleich zu der unseren allein durch die Auslese der Schule der Tendenz nach intelligenter sein! Dies könnte erklären, warum die Physiker in den Darstellungen von AMTHAUER in den sprachlichen Untertests SE, AN und GE und selbst im nichtsprachlichen Testteil in ZR und WÜ auf höhere Mittelwerte kommen im Vergleich zu unserer Stichprobe, bei gleich niedriger Leistung im Untertest ME.

### 3.1.3 Vertraulichkeit der Messergebnisse

Die Messergebnisse müssen aus vielen Gründen anonym bleiben. (Sie könnten sonst zur Beurteilung herangezogen werden: etwa bei der Zuteilung des Arbeitsplatzes, bei Prüfungen, bei Hierarchie-Reibereien innerhalb des Institutes usw.).

Daher müssen bei der Auswertung viele Fragen unbeantwortet bleiben, die sich bei der Datenanalyse ergeben und die durch Zusatzinformationen hätten geklärt werden können.

Die Auswertung muß so sein, dass ein Rückschluss auf die einzelne Person sehr erschwert ist. Völlig auszuschließen ist dies wohl nicht. Jedoch gewährleistet die starke Fluktuation im Institut und eine Verzögerung der Veröffentlichung um 29 Jahre eine Vertraulichkeit der Messwerte.

## 3.2 Die tabellarische Darstellung der Ergebnisse

Die Testergebnisse werden im Überblick in den Tabellen 10a und 10b sowie 11 wiedergegeben. Eine Übersicht gibt auch die Tab. 16. In Tab. 10a und in Tab. 16, obere Darstellung, sind die Standardwerte in absoluten Werten eingetragen, die als Ausgangswerte allen weiteren Tabellen, Berechnungen und graphischen Darstellungen zugrunde liegen.

Diese absoluten Standardwerte enthalten alle im Test erfassten Informationen. Aber nicht alle der in diesen Zahlen enthaltenen Informationen sind für die vorgesehene Datenanalyse von Interesse. Das Ziel dieser Untersuchung ist, die Intelligenzstruktur sichtbar zu machen (Betonung auf Struktur, der Intelligenzquotient IQ interessiert in diesem Falle dagegen **nicht**), und deren Feinstruktur beziehungsweise Unterstrukturen für diesen speziellen Berufszweig.

# 4 Zusammenfassung – Diskussion – Schlussfolgerungen

## 4.1 Zusammenfassung der Ergebnisse

Die im vorangegangenen Text erarbeiteten Ergebnisse werden hier noch einmal zusammenfassend wiedergegeben, um aus der Vielzahl der Einzelbefunde die wesentlichen hervorzuheben. So können dem Leser noch einmal Zusammenhänge verdeutlicht werden, die in dem doch recht umfangreichen Text leicht übersehen oder nicht als Ergebnis erkannt worden sein könnten. Die mit dem I-S-T durchgeführten Tests sind Messungen im Sinne von DIN 1319. Messungen enthalten Informationen. Es war Ziel dieser Arbeit, die in den hier beschriebenen Messungen enthaltenen Informationen sichtbar zu machen. Um dies zu erreichen, wurden die verschiedensten graphischen Darstellungsmöglichkeiten herangezogen.

Am I-S-T nach AMTHAUER haben **34 Physiker** des IAP teilgenommen. Ob die Zahl der teilnehmenden Personen für wissenschaftliche Aussagen ausreichend ist, kann durch Vergleich mit der Anzahl der Personen festgestellt werden, die den Werten der 21 Berufe bei AMTHAUER zu Grunde liegen, siehe Tab. 18. Bei den Architekten beträgt  $N = 28$ , den Geographen  $N = 29$ , den Bankbeamten  $N = 31$ , den Historikern und den Volkswirten  $N = 35$ , usw.. Da sind die 34 Physiker des IAP mit 34 Personen gut repräsentiert.

### 4.1.1 Die Ergebnisse der Analyse des I-S-T als Messinstrument

Die Aussagemöglichkeiten aus den mit dem Intelligenz-Struktur-Test gewonnenen Ergebnissen werden von AMTHAUER in einem zeitlichen Nacheinander und dann auch noch an verschiedenen Stellen beschrieben. Für jemanden, der in Bildern denkt, wie viele Angehörige der exakten Wissenschaften, ist die Beschreibung auf diese Art sehr unübersichtlich und schwer verständlich. Die Darstellung in einem **Flussdiagramm** (siehe Abb. 7), das aufzeigt, wie aus den Rohwerten RW im Gang der Auswertung die anderen Werte und damit Aussagen gewonnen werden, ermöglicht mit einem Blick, die entsprechenden Zusammenhänge zu erkennen. Man sieht, welche Ergebnisse auf welche Weise gewonnen wurden und wo Ungereimtheiten zu vermuten sind. So müsste es sich bei dem individuellen arithmetischen Mittel der Standardwerte  $\underline{SW}$  und dem Gesamt-Standardwert SW um denselben Wert handeln. Größtenteils weichen diese beiden Werte auch nur geringfügig voneinander ab, obwohl der individuelle Gesamtwert entsprechend den Vorgaben im Testhandbuch aus der Summe aller Rohwerte gebildet und für diese Summe in einer Tabelle der Standardwert abgelesen wird. Es handelt sich also nicht um einen exakten Mittelwert. Wenn diese Ungereimtheit auch belanglos zu sein scheint, ist doch nicht auszuschließen, dass dadurch geringfügige Ungenauigkeiten der Messung entstehen, die sich in einer Auswertung wie der unseren summieren können. Wir haben daher stets mit dem Mittelwert aus den 9 Untertest-Standardwerten gearbeitet.

Der Unterschied zwischen den **absoluten Standardwerten** (SW) und den **relativen Standardwerten** ( $\Delta SW$ ) ist der, dass bei den  $\Delta SW$  die allen Aufgabengruppen gemeinsamen Einflüsse auf das Messergebnis wegfallen. Denn je mehr Fakten in den Messwerten enthalten sind, um so schwieriger ist es, aus ihnen etwas herauszulesen. Die  $\Delta SW$ , als Abweichungen

vom individuellen Mittelwert, für Angehörige einer Berufsgruppe, sind demgegenüber frei von Einflüssen, die an die Personen gebunden sind, aber mit ihrer Begabungsstruktur nichts zu tun haben, wie z.B. Belastungen der sozialen Situation. Daher geben sie im Rahmen der Messwertstreuung die Eigenschaften eines Berufes frei von diesen berufsfremden Einflüssen wieder. Zu den berufsfremden Einflüssen, die alle 9 Aufgabengruppen in gleicher Weise beeinflussen können, gehören: Intelligenzniveau – tageszeitliche Schwankungen in der Leistung – seelische Belastung (Scheidung, Kündigung usw.) – Vorbildung – berufsbedingtes Sprachverständnis – Störungen – usw.

Beim I-S-T wird in Schritten von je **einem Rohwert RW** gemessen. Wie wirkt sich das auf den relativen Standardwert  $\Delta SW$  aus? Bei der Wortauswahl WA, siehe Tab. 8, bedeuten bei 12-Jährigen 1 RW gleich 4,2  $\Delta SW$ , bei 51- bis 60-Jährigen gleich 3,3  $\Delta SW$ , also eine kontinuierliche Abnahme der Werte, während bei den Würfelaufgaben die entsprechenden Werte 3,1  $\Delta SW$ , 3,0 und 2,8  $\Delta SW$  sind. Das bedeutet, dass bei WA im Laufe des Lebens (Schulbildung) ein Lerneinfluss das Lösen der Aufgaben erleichtert, während dies bei WÜ praktisch nicht vorliegt. Auch bei der Aufgabengruppe WÜ ist ein Lerneinfluss nicht zu erkennen, wenn gefragt wird, wie viele Aufgaben (RW) gelöst werden müssen, um 100 SW oder 130 SW zu erreichen. Daraus kann man schließen, dass Begabungen, die den Lösungen für WÜ-Aufgaben zugrundeliegen, im Gegensatz zu WA im Bildungssystem keine Bedeutung haben und nicht gefördert werden.

Aus der Summe der 9 Rohwerte wird über eine Tabelle (altersabhängig) der Gesamtstandardwert SW gewonnen, aus dem über die Gleichung  $\sqrt[3]{\frac{1}{2}SW} - 50 = IQ$  der **Intelligenzquotient IQ** errechnet wird. Indirekt handelt es sich um eine Mittelwertbildung, die etwas über die Intelligenzleistung des einzelnen aussagen soll. Die Natur hat es aber so eingerichtet, dass die Menschen in der Regel in bestimmten Begabungsbereichen spezielle Begabungshöhen und in anderen Begabungsmängel aufweisen, siehe Tab. 9. Ein bedeutender Komponist wie Mozart muss nicht ein Finanzgenie sein und ein Finanzgenie nicht ein bedeutender Komponist und jemand, der in allen 9 Aufgabengruppen Mittelmäßiges leistet, kann einen höheren IQ erreichen als einer mit Spitzenleistungen in einem Spezialfach (z.B. ein berühmter Komponist), bei einer Minderbegabung in dem einen oder andern für seine Tätigkeit uninteressanten Bereich. Es sei an den Klaviervirtuosen erinnert, dem das Abitur versagt blieb (Wicke, 1998).

Die Intelligenzleistung eines Menschen in einer einzigen Zahl zu erfassen, ist daher unangemessen und nicht sinnvoll. Dieses Vorgehen in den meisten Intelligenztests entspricht der Ideologie von einer „Globalintelligenz“, die sich in allen Schulleistungen gleichermaßen zeigen soll. Eine solche Globalintelligenz gibt es sicher nicht, wenngleich sich zeigen lässt, dass die Intelligenzstrukturkurven auf individuell unterschiedlichen Niveaus liegen. Diese Niveauunterschiede der Intelligenz werden aber eher in den Lebensleistungen der Menschen sichtbar, als in Gesamt-IQs von Tests. Und sie kommen in den Untertests zum Tragen, in denen sich die Spitzenleistungen der individuellen Begabungsstruktur ausdrücken.

Da der Text in fast allen Aufgabengruppen **sprachlich abgefasst** ist, treten beim Verstehen der einzelnen Aufgaben Probleme auf, die das Testergebnis negativ beeinflussen können, vgl. Abb. 09 und 10:

- 1.) Die **Sprachintelligenz** beeinflusst das Verstehen und damit das Lösungsverhalten aller sprachlich abgefassten Aufgaben.

- 2.) Die **Bedeutungsinhalte** der benutzten Wörter können in den verschiedenen Berufen merklich unterschiedlich sein und von denen der Umgangssprache abweichen. Es kommt damit zu Missverständnissen im Verstehen der Aufgabe, was sich im Lösungsverhalten negativ bemerkbar macht.
- 3.) Die **Definitionsschärfe** gibt an, wie exakt in den verschiedenen Berufen die in den Textaufgaben vorkommenden einzelnen Begriffe definiert sind, und kann zu erheblichen Missverständnissen führen und damit das Lösungsverhalten negativ beeinflussen.
- 4.) Die **Parallelform B2** führt bei einer Reihe von Aufgabengruppen zu besseren Ergebnissen. Sie stellt demnach eine leichtere Anforderung dar, als die Parallelform A1.

Aufschlussreich ist das Verhalten bei den Markierungen der Lösungsfragebogen mit richtig, falsch und nicht markiert, siehe Tab. 12 und Abb. 12. Die Markierungen erfolgten offensichtlich gewissenhaft, das heißt, es wurde beim Abbruch des jeweiligen Untertests nicht noch schnell wahllos der Rest der noch ungelösten Aufgaben angekreuzt, um die Möglichkeit auszunutzen, zufällig noch einige Aufgaben richtig anzukreuzen.

Bei den nicht markierten Aufgaben fallen zwei deutlich voneinander getrennte Aufgabengruppen auf.

- 1.) Bei drei Aufgabengruppen wurden mit Abstand die höchsten Aufgabenzahlen (174 bis 207) von insgesamt 680 Aufgaben **nicht markiert**. Bei ME (207 Aufgaben) wurde zu wenig im Gedächtnis behalten, dies aber auch erkannt, bei ZR (174 Aufgaben) wurden Falschlösungen gleichfalls als solche erkannt und daher nicht markiert, obwohl die richtige Lösung nicht gefunden wurde. Demnach bestand bei der Bearbeitung beider Aufgabengruppen Einsicht in das eigene Unvermögen. Dies dokumentiert sich in den niedrigsten Anzahlen von Falschmarkierungen im Vergleich mit den 9 anderen Aufgabengruppen. Bei WÜ (199 nicht markierte Aufgaben) gab es demgegenüber neben den fehlenden Markierungen eine beachtliche Zahl von Falschlösungen. Somit wurde häufig falsch markiert in der Annahme, die Lösung sei korrekt.
- 2.) Bei sechs Aufgabengruppen wurden nur 59 bis 93 Aufgaben **nicht markiert**. Vier dieser Aufgabengruppen weisen zugleich die höchsten Anzahlen von Falschlösungen auf: GE (130), SE (158), AN (160) und WA (206). Zum Vergleich: ME (40), ZR (42) und RA (53). Bei den Aufgabengruppen GE, SE, AN und WA wurde offensichtlich in der irrigen „Gewissheit“, die Aufgabe richtig gelöst zu haben, falsch markiert, die **Falschmarkierung** demnach als solche nicht erkannt.

Es fällt auf, dass in einer Aufgabengruppe manche Aufgaben von vielen Personen falsch gelöst wurden, während nur sehr wenige Testteilnehmer in anderen Aufgabengruppen falsch markiert haben. Das zeigt, dass einzelne Aufgaben innerhalb einer Aufgabengruppe für viele Physiker so **irreführend formuliert** sein müssen, dass sie eine falsche Lösung als richtig ansehen. Diese Fehlbeurteilung mag berufsbedingt typisch für den Physiker sein. Das bedeutet aber nicht, dass ein Physiker in dem Merkmal, das die Aufgabengruppe messen soll, unbegabt ist. Vielmehr sind die dort in der Frage vorkommenden Wörter in seiner Fachsprache mit anderen Inhalten belegt und auf Grund seines technischen Verständnisses anders zugeordnet und definiert als bei demjenigen, der die Fragen ausgearbeitet hat. Diese Voraussetzungen führen zu **Fehlleistungen zu Ungunsten der Physiker**. Auf diese Erscheinung machen auch SCHMIDT-ATZERT, HOMMERS und HEB (1995) aufmerksam, allerdings für die Aufgabengruppe WA. Aus diesem

Grund sind in unserer Untersuchungsstichprobe beim I-S-T faktisch keine Messabweichungen zu positiven sondern nur zu negativen Ergebnissen hin zu erwarten. Das bedeutet, Angehörige der exakten Wissenschaften werden durch diesen Test zumindest in einigen Aufgabengruppen benachteiligt.

#### 4.1.2 Die Analyse der Testergebnisse der Untersuchung im IAP

Beim **Vergleich der Intelligenzstrukturen der Physiker nach AMTHAUER und des IAP** (siehe Abb. 33) fällt auf, dass die Werte bei RA für Physiker des IAP um 8 SW höher, bei WÜ um 5 SW niedriger und bei ME um 4 SW höher liegen bei sonst gleichartigem Verlauf der Kurven. Der Unterschied in den beiden Strukturen kann am zeitlichen Abstand beider Messungen liegen: Die Physiker des IAP wurden 1973 und die Physiker von AMTHAUER 20 Jahre früher getestet. Die Berufsanforderungen haben sich geändert, die Auslese in den Schulen erfolgt nach anderen Kriterien (nachweislich durch die Einführung der Studienstufe) und die Fachrichtungen beider Gruppen können unterschiedlich sein. Daraus folgt, dass die Intelligenzstruktur der einzelnen Berufe in ganz bestimmten Aufgabengruppen stärkeren Schwankungen unterliegt. Daher ist es naheliegend, statt von Intelligenzstruktur-Kurven von -Bereichen oder -Bändern auszugehen.

Zum Vergleich der **Intelligenzstruktur von Untergruppen** (siehe Abb. 14 und 15), wurden zwei Untergruppen gegenübergestellt: Physikstudenten nur mit Vordiplom gegenüber Physikern mit bestandener Hauptdiplomprüfung oder Promotion sowie Physiker bis zu einem Alter von 27 Jahren und Physiker mit 28 und mehr Jahren (jeweils  $N = 17$ ). Die Unterschiede in den Intelligenzstrukturkurven sind sehr gering. Das spricht dafür, dass auch Untergruppen von  $34:2 = 17$  Personen nicht von der Intelligenzstruktur der 34 Physiker abweichen. Das bedeutet, der I-S-T ist ein sehr robustes Messinstrument. Die Testaufgaben führen bei Personen desselben Berufes immer wieder zu nahezu demselben Ergebnis. Das spricht einerseits für die Qualität des I-S-T. Es zeigt andererseits, dass die ausgewählten Personengruppen die gleiche Intelligenzstruktur haben wie die Gesamtgruppe.

Beim Vergleich der Intelligenzstrukturkurven von AMTHAUER für Mathematiker, Physiker und Ingenieure mit der Intelligenzstrukturkurve der Physiker des IAP (Abb. 16) stellt sich eine engere Beziehung zwischen den Mathematikern und den Physikern des IAP heraus, enger als zu den Physikern von AMTHAUER. Das lässt auf veränderte Anforderungen im Beruf, zumindest in der angewandten Physik schließen, wobei auch Auslesefaktoren eine Rolle spielen könnten. Bereits 1973 war die Physik stärker mathematisiert als zwanzig und mehr Jahre vorher.

Vergleicht man die **Häufigkeitsverteilungen** der absoluten und relativen Standardwerte, so fallen die doch sehr unterschiedlichen Verteilungen auf, siehe Tab. 16. Die Werte in einer Aufgabengruppe können als Zufallsverteilung, schiefe Verteilung oder Rechteckverteilung vorliegen. Jede dieser Verteilungen gibt Auskunft darüber, welche Bedeutung diese Aufgabengruppe im Rahmen der Bildungsselektion hat: Eine **Zufallsverteilung** in den **absoluten** Werten weist darauf hin, dass die Leistungen, die diese Aufgabengruppe repräsentiert, im Rahmen des Bildungsweges für die Selektion entscheidend ist. Allerdings werden die Verteilungen der absoluten Werte durch die oben erwähnten sozial bedingten Einflüsse mitbestimmt. Es kristallisiert sich immerhin eine annähernde Zufallsverteilung für die Aufgabengruppen SE und ME heraus, auf deren Bedeutung im Text näher eingegangen wurde. Eine schiefe Verteilung weist

auf einen „Deckeneffekt“ hin: Es konnten z.B. nicht mehr Aufgaben gelöst werden als der Test vorgibt. Damit setzt der Test eine obere Grenze für Bestleistungen. Eine Rechteckverteilung zeigt auf, dass die Leistungen in dieser Aufgabengruppe auf dem Bildungsweg oder für den Beruf uninteressant sind und so jede Leistung über den ganzen Bereich möglich ist. Etwas anders sind die Verteilungen der relativen Werte zu beurteilen: Hier sind die Verteilungen frei von nicht-relevanten Einflüssen. Eine Zufallsverteilung (wie in der Aufgabengruppe RA) besitzt höchste Bedeutung für die spezifische berufliche Anforderung. Auch für WÜ zeigt sich eine annähernde Zufallsverteilung ab, allerdings auf niedrigerem Niveau. Nicht zuletzt scheint auch die Leistung in AN für den Beruf des Physikers von Bedeutung zu sein. Dass diese Verteilung nach unten hin abbricht, spricht wiederum für den Einfluss der schulischen Auslese.

Vergleicht man die gegensätzlichsten Intelligenzstrukturen der Berufe nach AMTHAUER, nämlich die Philosophen und die Dipl. Ingenieure (Abb. 20), so zeigt sich bei den Ingenieuren unter allen 21 Berufen der geringste Begabungsunterschied innerhalb ihrer Intelligenzstruktur von nur 7 Standardwerten (-4 bei ME und +3 bei SE, AN und WÜ). Der höchste **Begabungsunterschied** findet sich bei den Philosophen mit 20 Standardwerten (-8 bei WÜ und +12 bei GE), siehe Abb. 20. Bei keinem der Berufe von AMTHAUER kommt eine „Rundumbegabung“ vor, d.h. annähernd gleich hohe Werte in allen der gemessenen Intelligenzbereiche. Am nächsten kommen dem noch die Angehörigen der exakten Wissenschaft und von ihnen wiederum die Diplom Ingenieure. Die Annahme einer „einseitigen Begabung“ trifft nach diesen Ergebnissen also eher für Geisteswissenschaftler und weniger für Naturwissenschaftler zu.

Auch bei den Physikern des IAP gibt es kein Individuum mit einer „**Rundumbegabung**“, siehe Tab. 11, Tab. 19 bis 21. Bei dem Physiker mit dem geringsten Begabungsunterschied innerhalb seiner Intelligenzstruktur betrug der Abstand zwischen dem höchsten und niedrigsten Wert 15 Standardwerte (Abb. 21 und 22 a). Bei dem Physiker mit dem höchsten Begabungsunterschied waren es 41 Standardwerte. Der häufigste Begabungsunterschied wurde mit 26-30 Standardwerten bei 10 von den 34 Physikern des IAP gefunden. So ausgeprägten Begabungsunterschieden kann eine Schule mit vorwiegend sprachlicher Anforderung in keiner Weise gerecht werden.

Begabungstiefs müssen unter Umständen kompensiert werden. Da der Betroffene nicht weiß, wie der „Normalbegabte“ denkt, sieht er das durch Kompensation erzielte Ergebnis als die Leistung eines „Normalbegabten“ an. Aber Kompensation ist eben ein Hilfsmittel, eine Gedächtniskrücke oder eine Umweg-Lösung. Allerdings sprechen solche Lösungen auch für Kreativität, eine Anlage, die von diesen Personen notgedrungen intensiver trainiert wird.

Bei den Physikern des IAP war zu klären, ob die Leistungen in den einzelnen Aufgabengruppen in irgendeiner Form mit den Leistungen in anderen Aufgabengruppen zusammenhängen, siehe Abb. 24 und 25 und Tab. 24. Wo fand man die 5 Besten und die 5 Schwächsten einer Aufgabengruppe in den anderen Aufgabengruppen wieder? Trägt man ihre Intelligenzstrukturen getrennt in ein Diagramm ein (Abb. 24), so zeigen sich eine ganze Reihe sehr deutlicher Zusammenhänge. So stehen hohe Leistungen in einer Aufgabengruppe mit niedrigen in einer anderen in Verbindung (SE-ME) oder hohe in einer Aufgabengruppe wiederum mit hohen in einer anderen (ME-WÜ). Es ist also ein Zusammenhang gegeben! Für die 21 Berufe nach AMTHAUER zeigen sich ähnliche, wenn auch zum Teil andere Zusammenhänge.

Für beide Gruppen, die Physiker des IAP und die Berufe nach AMTHAUER, siehe Abb. 25, zeigt sich: wenn hohe Leistungen in WA vorliegen, dann sind sie in FA niedrig und umge-

kehrt. Diese Aussage scheint allgemeingültig zu sein. Somit schließen sich zwei Begabungen gegenseitig aus. Man könnte daher im I-S-T auf eine der beiden Aufgabengruppen verzichten. Dieses Ergebnis schließt auch die Möglichkeit einer Rundumbegabung aus, sowohl nach den Daten von AMTHAUER als auch nach den unseren.

Was für die 5 höchsten und 5 niedrigsten Werte zutrifft, wurde in Abb. 28a für alle Personen der Parallellform B2 für WA und GE in einem Diagramm aufgetragen und überprüft, indem die einzelnen Nummern der Physiker bei WA mit denen bei GE durch einen Strich miteinander verbunden wurden. Es zeigte sich, dass hohe Werte in einer Aufgabengruppe mit niedrigen Werten in der anderen Aufgabengruppe verbunden sind, mittlere Werte in der einen Aufgabengruppe auch mit mittleren in der anderen und niedrige in der einen mit hohen in der anderen. Diese Bedingungen treffen, wie zu erwarten, „cum grano salis“ also im Rahmen einer schwachen individuellen Streuung zu. Für die 21 Berufe nach AMTHAUER ergibt sich ein etwas anderes Bild (Abb. 28b). Es kommen überwiegend gleichsinnige Leistungen vor. Doch ist hier zu berücksichtigen, dass die Werte über viele Einzelpersonen gemittelt wurden. Wie sich zeigen ließ, heben sich bei einer solchen Mittelung gegenläufige Einzelwerte gegeneinander auf, vor allem, wenn sich in der Gesamtgruppe unerkannte Untergruppen verbergen.

Trägt man die relativen Standardwerte von WÜ gegen WA für die 21 Berufe in ein Koordinatensystem ein und fasst die Werte von Berufsgruppen zusammen, so zeigt sich, wie eng die Werte innerhalb der Berufsgruppen beieinander liegen (Abb. 27) und wie deutlich sich die Berufsgruppen in der Abbildung gegeneinander abheben. Trägt man den WÜ-WA-Mittelwert für die 34 Physiker des IAP ein, so ergibt sich deutlich, dass sie zu den exakten Wissenschaften gehören, wie zu erwarten war. Es ist überraschend und erstaunlich, dass schon die beiden Werte WÜ und WA einen Beruf eindeutig einer Berufsgruppe zuordnen.

Die Intelligenzstrukturen aus den absoluten Standardwerten geben Auskunft, wie die einzelnen Berufe in ihrem Intelligenzniveau zueinander liegen, siehe Abb. 45a, 45b und 46. Die beiden in ihren Intelligenzstrukturen gegensätzlichsten Berufe sind nach AMTHAUER die Neuphilologen und die Physiker. Nur in ME und WA liegen die Physiker um 2 bzw. um 1 Standardwerteinheiten niedriger als die Neuphilologen, sonst aber erreichen sie höhere Werte, zum Teil sogar erheblich höhere wie in ZR mit 11, in RA und FA mit 12 und WÜ mit 13 Standardwerteinheiten. Da der IQ für Einzelpersonen im I-S-T aus den Rohwerten aller 9 Untertests gebildet und in einer Tabelle abgelesen wird, muss man bei der Bestimmung eines IQ-Aequivalents für Personengruppen einen anderen Weg gehen: Der mittlere Intelligenzquotient für eine Berufsgruppe wird aus den Mittelwerten aller 9 Untertests dieser Berufsgruppe summiert und durch 9 dividiert. Dabei ergibt sich, dass der Gesamt-IQ der Physiker sehr viel höher liegt als der der Neuphilologen. Ähnlich große Unterschiede zeigen sich zwischen den Berufsgruppen der exakten Wissenschaften und der Geisteswissenschaften. Die exakten Wissenschaften weisen die erheblich höheren IQ-Werte auf.

Da die Standardwerte der einzelnen Aufgabengruppen für Berufsgruppen in einem gewissen Bereich schwanken, je nachdem, aus welchen Untergruppen des Berufes die einzelnen Personen kommen, und da die Berufe im Laufe der Zeit einem Wandel unterworfen sind, liegt es nahe, diejenigen Berufe zusammenzufassen, deren Werte in den einzelnen Aufgabengruppen innerhalb dieses Schwankungsbereichs liegen, siehe Abb. 34, 35 und 36. So können die Architekten und Künstler zusammengefasst werden – die Bankbeamten, Kaufleute und Volkswirte – die Ärzte, Biologen und Psychologen – die Chemiker, Ingenieure, Mathematiker und Physiker

zu den exakten Wissenschaften – die Altphilologen, Germanisten, Historiker, Juristen, Neuphilologen, Pädagogen, Philosophen und Theologen zu den Geisteswissenschaften – nur die Geographen haben eine Intelligenzstruktur, die mit keiner anderen in Verbindung gebracht werden kann. Diese Zusammenfassung nach Berufsgruppen ist sinnvoll, da wegen der Schwankungsbreite der Standardwerte eine weitere Unterteilung eine Gewissheit vortäuscht, die nicht begründet ist.

Wie sehr die Schule einseitig aussondert, geht auch aus Tab. 35 und aus den Abb. 47 und 48 hervor: Wenn auch die meisten Schüler (in der 5. Klassenstufe) mit ihren Fehlern um die Regressionsgerade streuen, so gibt es doch zwei kleine Gruppen von Schülern, die sich aus dieser Verteilung herausheben. Die eine Gruppe weist trotz niedriger Intelligenz eine Spezialbegabung für Rechtschreibung auf (trotz niedrigem IQ fast keine Fehler). Es sind 8 von 102 Schülern. Einer von diesen besucht das Gymnasium und drei die Realschule. Die andere Gruppe macht trotz hohem IQ sehr viele Rechtschreibfehler. Es sind 9 von den 102 Schülern. Einer besucht das Gymnasium, zwei die Realschule. Offensichtlich liegt hier eine dritte Selektion nach der Rechtschreibleistung vor. Bekräftigt wird diese Vermutung einer dritten Selektion durch die Darstellung der Rechtschreibfehler in Abhängigkeit von der Intelligenz für die Schüler der drei Schulformen Hauptschule, Realschule und Gymnasium, siehe Abb. 48. Dieses Diagramm zeigt, dass indirekt die Rechtschreibleistung über die Schulform entscheidet, die der Schüler besuchen darf.

### 4.1.3 Die Analyse der Abiturnoten

Von den 34 Physikern des IAP hatten 27 vergleichbare Abiturzeugnisse. Die Mittelwerte der Abiturnoten zeigen, dass die Physiker, wie zu erwarten, in Physik die beste Leistung aufweisen, in Mathematik die zweitbeste, aber in Französisch die schwächste und kurz danach folgend in Englisch und in etwas größerem Abstand in Deutsch, siehe Abb. 37.

Eine schiefe Verteilung zu schwachen Noten hin betrifft die drei Sprachen und zwar besonders deutlich Englisch, aber auch Deutsch und Französisch. Offensichtlich haben die Sprachen mit dem Physikverständnis nichts zu tun. Die internationale Verständigung auch in Physik ist zwar auf die Sprachen angewiesen, bedarf aber zunächst des Verständnisses der Sachverhalte. Die internationalen Fachsprachen haben anscheinend sehr wenig Gemeinsamkeiten mit dem Schulenglisch. Daher stellt sich die Frage, ob es zu rechtfertigen ist, Höchstbegabungen in Physik wegen mangelhafter Leistungen in den Sprachen nicht das Abitur bestehen zu lassen (Selektion). Man denke an EINSTEIN, FEYNMAN und LINNÉE.

Werden die Mittelwerte der Noten der einzelnen Physiker für die naturwissenschaftlichen Fächer und getrennt davon für die geisteswissenschaftlichen Fächer gebildet, siehe Tab. 30, und nach Anzahl der Personen aufgetragen, so ergibt sich für die naturwissenschaftlichen Fächer, wie zu erwarten, eine schiefe Verteilung zu den besten Noten hin, während bei den geisteswissenschaftlichen Fächern eher eine Rechteckverteilung vorliegt, ein Hinweis darauf, dass diese Fächer für die Physikbegabung ohne Bedeutung sind.

Wird von der Physiknote die Note einer der Sprachen abgezogen, so zeigt sich, dass keiner in einer der Sprachen eine bessere Note als in Physik hat, jedoch 2 Physiker in Französisch und einer in Englisch um 4 Noten schwächer sind als in Physik. Auch sonst liegen bei allen Physikern die Noten in den Sprachen im Mittel um 2 Noten schwächer als in Physik (Tab.31).

Die Vermutung, dass die Noten in Deutsch und Englisch als ein Hinweis auf die Abiturnote gelten können, lässt sich schnell überprüfen, siehe Abb. 40. Wird die Abiturdurchschnittsnote in Abhängigkeit von dem Notenmittelwert aus Deutsch und Englisch wiedergegeben und die Regressionsgerade eingezeichnet, so ist ein deutlicher Zusammenhang zu sehen. Bestätigt wird das durch den Vierfelder  $\chi^2$ -Test. Danach sind beide Notengruppen, Abiturnoten und Notenmittelwert aus Deutsch- und Englischnote, mit einer Wahrscheinlichkeit von mindestens von 99,9% voneinander abhängig. Da aber in die Abiturnote auch die Noten der naturwissenschaftlichen Fächer eingehen, ist dieses Ergebnis nur verständlich, wenn die naturwissenschaftlichen Noten auch von den sprachlichen Leistungen in den naturwissenschaftlichen Fächern abhängen. Wer in den schriftlichen Leistungen der naturwissenschaftlichen Fächer sprachlich nicht gut ist, wird, ganz gleich, wie seine fachliche Leistung aussieht, schlechter benotet. Das regelt seit der Einführung der Studienstufe schon der Punkteabzug für Rechtschreibfehler in allen schriftlichen Arbeiten. Es heißt ja: Wer nicht Deutsch kann, kann nicht denken. Da die Physiker aber nach den Werten, die AMTHAUER mitteilt, einen erheblich höheren Intelligenzquotienten haben als die Germanisten und Neuphilologen, muss es sich wohl bei diesem Ausspruch um ein Wunschenken dieser Gruppe handeln. Die Auswirkungen sind für die Physiker ungünstig. Es zeigt sich, wie groß der negative Selektionsdruck für die Physiker ist.

Wie hängen nun der Mittelwert der Noten aus Deutsch und Englisch mit der Merkfähigkeit ME zusammen? Werden diese Werte in ein Diagramm eingetragen und die Begrenzungslinien für die Hauptmasse der Werte eingezeichnet, so ist ein Zusammenhang deutlich zu sehen, siehe Abb. 41. Der Vierfelder  $\chi^2$ -Test bekräftigt diese Aussage, dass die ME-Werte und der Sprachnotenmittelwert aus Deutsch und Englisch mit einer Wahrscheinlichkeit von mindestens 99% zusammenhängen. Nun haben aber die Physiker nach AMTHAUER und die des IAP gerade in ME sehr niedrige Werte.

Es ist daher naheliegend, die Abiturdurchschnittsnoten für Deutsch und Englisch mit den relativen Standardwerten für ME in Beziehung zu setzen, siehe Abb. 41. Der Vierfelder  $\chi^2$ -Test sagt über die Wahrscheinlichkeit aus, dass die Abiturdurchschnittsnoten mit der Merkfähigkeit (gelernte Wörter behalten zu können, längerfristiges Behalten, Gedächtnis) mit mehr als 97,5% zusammenhängt. Um zu einem guten Abitur zu gelangen, sind offenbar Gedächtnisleistungen Ausschlag gebend. Der Verdacht, die Schule setze vor allem auf rezeptives Lernen und Reproduktion des Gelernten, bestätigt sich in unseren Ergebnissen. Auch die Ergebnisse der PISA-Studie weisen – mehr als 25 Jahre später – in die gleiche Richtung.

## 4.2 Diskussion der Ergebnisse

Die Diskussion der Ergebnisse der Untersuchung mit dem I-S-T, wie der Befunde aus dem Literaturstudium zu dieser Arbeit, führt uns wieder zu den unterschiedlichen Begabungen zurück. Den Begabungsbegriff benutzen wir hier im Sinne von Begabungsschwerpunkten, die sich in Intelligenzstrukturkurven wie in den Abiturnoten abgezeichnet haben und für die sich auch in Notenentwicklungen und Intelligenztests legasthener Schüler Parallelen abgezeichnet haben. In diesem Zusammenhang sei nochmals betont, dass Ausdrücke wie „geisteswissenschaftliche“ oder „sprachliche“ Begabung ebenso wie „mathematische“ und „technische“ Begabung nur grobe begriffliche Annäherungen an die Realität darstellen können. Darüber hinaus gibt es wei-

# 7 Register

## A

Abiturnotenmittelwert 214  
Abiturprüfung 216  
Absolutwert 104  
abstrakte Theorie 266  
Abstraktionsfähigkeit 100  
adressenorientiert 229  
Altersabhängigkeit 112, 193  
Altersverteilung 132  
analytisch 33, 80, 86-89  
Anforderungsniveau 12, 56, 142, 245  
anonym 123, 131, 134, 188, 198  
Anonymität 118, 163, 171, 193  
Anpassungssyndrom 256  
Arbeitsweise 31, 119, 178, 228, 273  
archaische Wissenschaft 97  
Arroganz 128, 177  
assoziativ 229  
Aufgabenstellung des Instituts 133  
Ausdrucksfähigkeit 222  
Ausgangsdaten 235  
Auslassungen 150  
Auslesefaktor 233, 241, 250, 261  
Auslesekriterien 233f, 237, 241  
Ausleseprinzip 207, 236, 238, 240  
Ausleserelevanz 245  
Ausreißer 135, 166, 236  
Ausreißerwert 135, 186  
Außenseiter 256, 257  
Auswendiglernen 44, 51, 189

## B

Bearbeitungstempo 144, 152  
Bearbeitungszeit 98, 144  
Begabung  
    Global- 28, 127, 182  
    Rundum- 171, 251, 252  
    Spezial- 117, 208, 236  
Begabungs-  
    schwerpunkte 13, 27, 37, 40, 42, 44, 178,  
    194, 258, 264f  
    spitzen 177, 210, 258  
    tief 176  
    unterschiede 13, 60, 111, 168f, 171, 175, 219  
    varianten 265  
Begriffsbildung 50, 100, 168f, 173, 231  
behindert 51, 129, 189, 196  
Bereichsumfang 166

Berufsberatung 24, 124, 126, 129, 160, 188, 208  
Berufserfolg 263  
Berufsfelder 124  
Berufsgruppendarstellung 210  
berufunspezifische Leistungen 207  
Berufsvoraussetzung 167  
Beweglichkeit 100, 157, 237  
Bildererkennen 267  
bildhaft-figural 86  
Bildungsideologie 75  
bildungspolitische Fragen 13, 28, 94  
Blockschaltbild 233

## C

CFT 20, 30, 31, 267  
Chancengleichheit 60  
Computerwissenschaften 261

## D

Datenanalyse 120, 134f, 261, 266  
Dauerstress 128  
Definitionsschärfe 249  
Denkbahnen 230  
Denken  
    Analogie- 163, 184  
    eindimensionales 108  
    in Bildern 90f, 189  
    in Worten 90, 189  
    kreatives 32  
    mehrdimensionales 108  
    räumliches 189  
    sprachliches 93, 99, 194f, 255  
    Strategie- 107  
    vernetztes 108f, 267  
Denkhaltungen 255, 268  
Denkschemata 230  
Denkstrategien 239  
Denkvorgänge 189, 229  
Denkweisen 239, 269  
Detailerfassung 30ff, 88, 267  
Dichotomie 32, 89  
Differentialgleichungen 232  
Dyskalkulie 28, 268  
Dyslexie 42, 260

## E

Eichstichprobe 34, 102, 262  
Eidetiker 91

Eindeutigkeit 142  
 Einfühlungsfähigkeit 61, 99, 194  
 Einseitigkeit 15, 64, 124, 189  
 Einzelkämpfer 264  
 Entdeckung 228, 266

**F**

Fächernoten 221  
 Fachgymnasium 48f, 59, 65, 67, 80  
 Faktorenanalyse 29, 30f, 262  
 Fehlerabschätzung 106, 118  
 Fehlerfortpflanzungsgesetz 113  
 Fehlerquellen 106, 114, 118  
 Fehlerzahl 70, 73, 241, 244  
 Fehlleistungen 69, 128, 210, 223, 249  
 Feindifferenzierungen 32, 151  
 Feinstruktur 134  
 Figurenauswahl 99f, 201  
 Flexibilität 14, 101-104, 109, 257  
 Flussdiagramm 101, 233, 247  
 Folgerichtigkeit 100, 153, 157, 208, 237  
 Förder-Privilegien 260  
 Formelsprache 93  
 Formulierungen 151, 157, 196, 232  
 Freiheitsgrade 190  
 Funktionspapier 166

**G**

ganzheitlich 33, 70, 80, 86-91, 229, 255  
 Gaußsche Verteilung 120, 163, 166  
 Gedächtnis 69, 100, 108, 144, 178, 222f, 225f, 228-231, 249, 254  
 Gedächtniskrücke 251  
 gegenläufige Tendenz 189, 190  
 Gegenüberstellung 79, 154, 206, 219  
 Gehirn 32, 38, 41, 85f, 267f  
 Geisteswissenschaft 84f  
 Geltungsbereich 21, 228  
 Generalfaktoren 31  
 gescheitert 48, 76, 128  
 Geschicklichkeit 49, 61, 231, 264  
 Glättung 146  
 Gleichheitsideologie 257  
 gleichmäßige Verteilung 165ff, 212  
 Gleichungen 91f, 100, 193, 228, 232  
 gleitender Durchschnitt 146  
 Glockenform 216  
 Grammatik 41, 74, 77, 86  
 Graphiken 185, 206, 232  
 graphische Auswerteverfahren 97  
 graphische Darstellung 16, 120, 135, 147, 151, 187, 193, 196, 220  
 Gültigkeit 17, 23, 35, 63, 102

Gütekriterien 34, 37

**H**

Häufigkeitssummenkurve 166f  
 Hauptschule 11, 20, 36, 38, 50, 52, 56, 58f, 72f, 78, 241f, 244, 253  
 HAWIK 51, 267  
 Hilfsmittel 92, 196, 251  
 Hirnforschung 32, 85f, 90  
 Hirnhälften 32f, 85f, 90, 255, 268  
 Hochbegabten 38, 42f, 89, 207, 255ff  
 Hochschulabschluss 131, 136ff, 261  
 Hochschulzugang 195  
 HPI 31  
 HSU 265  
 Humanismus 77  
 Hypothesen 16, 141

**I**

individuelle Belastungen 135  
 Informatik 47, 59, 65, 74, 232, 261, 265  
 Information 103, 109, 116, 220, 229, 238  
 inhaltsleer 167  
 Intelligenz  
   Allgemein- 43, 61, 75, 204  
   -alter 35  
   -bereiche 27, 38, 111, 233  
   -diagnostik 11, 260, 267  
   Global- 248, 266  
   Grund- 31, 262  
   -höhe 19, 201, 206, 266  
   -konzepte 13  
   -profil 28, 116, 262  
   soziale 32, 110  
   Sprach- 248  
 Intuition 24, 109, 178, 231  
 intuitiv 87, 163, 177  
 Irrtümer 17, 118, 150  
 I-S-T 2000 13, 36f, 148, 262, 267  
 I-S-T 2000 R 37, 262, 267

**K**

kasuistische Methode 108  
 Kaufman-ABC-Tests 32  
 Kombinationsfähigkeit 100, 157, 178, 237  
 Kompensation 49, 128, 176, 178, 189, 251  
 Kompensationsmethode 176  
 kompensieren 84, 119, 129, 176, 178  
 komplexe Strukturen 107  
 Konkurrenzkampf 129  
 konkurrieren 235  
 Konstrukt 25, 37

Kontingenz 190  
 Korrelation 24, 71, 104, 184, 188  
 Kurzzeitspeicher 30, 51, 53ff

**L**

Lebensalter 35, 53, 113f  
 Lehr-Erfolgskontrollen 263  
 Lehrplan 233  
 Leistungsbewertung 265  
 Leistungsdifferenz 149, 171  
 Leistungsmängel 176  
 Leistungsmotivation 59  
 Leistungsschwächen 173, 184  
 Leistungsschwerpunkte 141  
 Lernbehinderung 28, 36  
 Lern-Erfolgskontrolle 263, 268  
 Leselernprozess 263, 265, 268  
 Linkshändigkeit 43, 86, 259ff  
 Literaturstudium 228, 230f, 254  
 Lösungen  
     Falsch- 151ff, 249  
     falsche 150, 152, 156  
     Pseudo- 229f  
     Ungefähr- 100, 151f, 157, 208, 237  
 Lösungshäufigkeiten 144, 154  
 Lösungsverhalten 153-157, 248f  
 Lösungswege 14, 163, 256  
 markierte Aufgaben 143, 151, 249

**M**

Mathematik-Olympiaden 207, 257  
 Mehrheitsverhältnisse 256  
 Mengenlehre 232  
 Merkaufgaben 99f, 230  
 Messgenauigkeit 34f  
 Messung  
     absolute 116  
     Einmaligkeit der 133  
 Messwolke 135  
 Milieu 18, 31, 60, 103, 259  
 milieubenachteiligt 265  
 Millimeterpapier 166  
 Minderheit 42f, 78f, 127, 129, 208, 256ff  
 Mittel  
     arithmetisches 101, 123, 206, 271  
     gewogene arithmetische 122f  
 Modellvorstellung 31, 37, 177  
 Muttersprache 69, 73, 77, 196, 257

**N**

Nachdenken 16, 23, 92, 94, 177f, 228  
 Näherung 107, 120, 126, 135, 222

Nebeneinander 232  
 Nebeneinflüsse 98  
 Normalbegabte 176f, 179, 251  
 Normskala 35f, 53ff  
 Normtabellen 73, 112  
 Normvariante 43, 259f  
 Notenschutz 49, 56, 74, 260  
 Notenwert 216, 218

**O**

Oberrealschulen 65, 78, 79  
 Oberschicht-Eltern 260  
 Oberstufenreform 233  
 Objektivität 34f, 102, 105  
 ökologische Nische 32, 177, 220  
 Orientierungsstufe 67f  
 Orthographie 45, 47, 70, 75

**P**

Parallelform 144, 147f, 156, 199f, 249, 252  
 Philologenverband 265  
 PISA-Studie 13, 71, 254  
 Primärfaktoren 31  
 problemgerecht 231  
 Prokrustesbett 90  
 Prozentwert 102, 104  
 PSB 28, 50f, 173, 241, 244, 268  
 Pubertät 25, 90, 216  
 Punktabzug 73ff

**R**

Rahmenvereinbarung 65, 66  
 Rationalisierungsprozess 177  
 Rauschen 116f, 146, 273  
 reasoning 28, 100, 157  
 Rechenaufgaben 29, 35, 99, 100  
 Rechentests 262  
 Rechteckverteilung 121, 166f, 250, 253, 272  
 Regressionsgerade 192f, 241, 253f, 272  
 relative Messwerte 135  
 Reliabilität 34f, 148  
 reproduktiv 229  
 Risiko-Kinder 71  
 Rohwertverteilung 36

**S**

Satzergänzung 99, 230  
 Satzkonstruktion 232  
 Schaltplan 233  
 Scheinlösung 142  
 schöpferisch 129, 228  
 Schriftsteller 45ff, 81, 88

Schulerfolg 24, 50, 71, 199, 205, 255  
 Schulleistungstests 11, 263, 268  
 schulrelevante Aufgabengruppen 236  
 Schulselektion 209  
 Schulstreit 76, 78  
 Schulunterricht 34, 112, 222, 266  
 Schulversager 71, 217  
 Schwachsinnige 176  
 Schwierigkeitsgrad 34, 105, 144, 148f, 153f  
 Selbständigkeit 13, 99, 229ff  
 Selektion 133, 141, 225, 250, 253  
 Signifikanzschranken 191, 195, 222, 224  
 Simplifikation 255  
 Sonderfälle 135  
 Sonderschüler 241  
 sozial schädlich 260  
 soziale Kompetenz 108  
 soziale Randgruppen 46  
 soziale Sensibilität 82  
 sozialer Status 42, 257  
 Sozialschichten 70  
 Speicher 51, 229, 273  
 Sperrklausel 74  
 Spezialisten 122  
 Sprachbegabte 67, 124, 152  
 Sprachgefühl 99, 194  
 sprachliche Differenzierungen 88, 151, 233  
 sprachlos 232f  
 Sprachnoten 215, 218f, 221, 223, 225  
 Sprachprozesse 267  
 sprachunabhängig 267  
 statistische Sicherheit 120, 222, 272  
 stigmatisieren 105  
 Stress 73, 128  
 Stresssituationen 176  
 Strukturprofile 124, 262  
 Summenkurve 186, 272

**T**

Täuschung 194  
 Team 32, 41, 110  
 Testergebnis 35, 52, 98, 115, 248  
 Testhälfte 35, 148  
 Testkonstrukteur 147, 150, 155  
 Testskalen 262  
 Toleranz 109, 127  
 Tradition 67, 76, 109, 228  
 Trennschärfe 135  
 Tüchtigkeit 60  
 Typeneinteilung 81, 255

**U**

Übereinstimmung 35, 187, 255

überfordert 177  
 Überleben 23, 41, 97, 229  
 Überprüfung 16, 28, 157f, 231, 263  
 Überschneidungen 135  
 Umweltfaktoren 25, 233  
 undeutlich 107, 119  
 Unfähigkeit 95, 97, 107, 167, 231  
 Ungerechtigkeiten 128  
 Untertestleistung 146, 180  
 Unverständnis 58, 256  
 Urteilsbildung 74, 99, 135, 221, 230  
 Urteilstheorien 263

**V**

Validität 17, 21, 24, 30, 34f  
 Varimax-Rotation 262  
 VDI 265  
 verballogisch 158  
 verbesserungsbedürftig 150, 228, 261  
 Vergleichsuntersuchungen 264  
 Vertraulichkeit 134  
 Virtuosität 231, 264  
 Vorannahmen 216  
 Vorschulzeit 233  
 Vorurteil 203

**W**

Wahrscheinlichkeitspapier 166, 168, 227, 272  
 Werte  
     absolute 115  
     relative 149, 185  
 Wiederholer 63, 71  
 Wirklichkeitssinn 99, 230  
 Wortauswahl 99, 112, 168, 189, 248  
 Wortwahl 74, 232

**Z**

Zahlenmaterial 141, 143  
 Zahlenreihen 88, 99, 100  
 zeitliches Nacheinander 189  
 Zivilisation 84, 269  
 ZNS 189  
 zufällige Messabweichung 98, 112f, 271  
 zufällige Schwankung 146, 167, 216  
 zwei Kulturen 15, 81, 84, 207, 229, 255, 269  
 Zweifaktorenlösung 262  
 Zweifaktoretheorie 30