

Werner Sacher

Leistungen entwickeln, überprüfen und beurteilen

**Bewährte und neue Wege für die
Primar- und Sekundarstufe**

6., überarbeitete und erweiterte Auflage

Sacher

**Leistungen entwickeln,
überprüfen und beurteilen**

Werner Sacher

Leistungen entwickeln, überprüfen und beurteilen

Bewährte und neue Wege
für die Primar- und Sekundarstufe

6., überarbeitete und erweiterte Auflage

Verlag Julius Klinkhardt
Bad Heilbrunn • 2014

k

Dieser Titel wurde in das Programm des Verlages mittels eines Peer-Review-Verfahrens aufgenommen. Für weitere Informationen siehe www.klinkhardt.de.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet abrufbar über <http://dnb.d-nb.de>.

2014.L. © by Julius Klinkhardt.

Das Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt.

Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Druck und Bindung: Friedrich Pustet, Regensburg.

Printed in Germany 2014.

Gedruckt auf chlorfrei gebleichtem alterungsbeständigem Papier.

ISBN 978-3-7815-1980-0

Inhalt

Vorwort zur 6. Auflage	11
-------------------------------------	-----------

1 Zum Sinn des Leistens, Prüfens und Beurteilens

1.1 Grundbegriffe	13
1.1.1 Leistung	13
1.1.2 Leistungsstreben	13
1.1.3 Leistungsanforderung	14
1.1.4 Leistungsorientierung	14
1.1.5 Leistungsprinzip	15
1.2 Zur Übertragbarkeit des Leistungsprinzips auf die Schule	17
1.3 Zur Geschichte der Prüfungen, Zeugnisse und Zensuren	19
1.4 Funktionen von Prüfungen, Zeugnissen und Zensuren	22
1.4.1 Selektion und Stigmatisierung	22
1.4.2 Sozialisation	24
1.4.3 Legitimation.....	27
1.4.4 Kontrolle.....	27
1.4.5 Prognose.....	27
1.4.6 Information und Rückmeldung.....	28
1.4.7 Disziplinierung.....	28
1.4.8 Lehr- und Lerndiagnose	29
1.4.9 Lern- und Leistungserziehung	30
1.5 Funktionale Überfrachtung schulischer Prüfungen und Leistungsbeurteilungen.....	30

2 Zur Messqualität von Prüfungen und Leistungsbeurteilungen

2.1 Die Beurteilung von Schulleistungen als Messvorgang	33
2.2 Anforderungen an Messungen	35
2.2.1 Objektivität.....	35
2.2.2 Reliabilität.....	36
2.2.3 Validität.....	36
2.2.4 Die Nebengütekriterien, Ökonomie, Nützlichkeit/ Zumutbarkeit und Akzeptanz.....	41
2.3 Forschungsergebnisse und Verbesserungsvorschläge zur Messqualität von Schulnoten	42
2.3.1 Zur Objektivität	42
2.3.2 Zur Reliabilität	43
2.3.3 Zur Validität.....	44

2.4	Das Objektivitäts-Validitäts-Dilemma der Leistungsmessung	46
2.5	Urteilsfehler bei der schulischen Leistungsbeurteilung	47
2.5.1	Ungleichmäßige Ausschöpfung des Beurteilungsspektrums	48
2.5.2	Interferenzen im Urteil (Voreingenommenheiten)	50
3	Die Planung und Anlage von Prüfungen	
3.1	Strategien der Leistungsüberprüfung.....	53
3.2	Das Konzipieren der Prüfung	54
3.2.1	Auswahl der Prüfungsinhalte	54
3.2.2	Auswahl der Prüfungsform	56
3.2.3	Auswahl der Aufgabenformen.....	59
3.2.4	Festsetzen der Anforderungsniveaus.....	66
3.2.5	Bestimmung des Aufgaben- und Prüfungsumfangs.....	67
3.2.6	Formulierung der Aufgaben.....	68
3.2.7	Die Reihenfolge der Aufgaben	70
3.2.8	Planung der Prüfungssituation.....	70
3.2.9	Ausarbeiten einer Musterlösung.....	72
3.3	Punkte- und Fehlerzuweisung: Schwierigkeit und Bedeutsamkeit	75
3.4	Festsetzen der Mindestkompetenz.....	79
3.5	Die Überprüfung der Messfehler.....	80
3.6	Die weiteren Schritte	80
4	Leistungsbeurteilung als Bewerten	
4.1	Die Komplexität schulischer Leistungsmessung	83
4.2	Benoten als abstraktes und verschleiertes Werten	83
4.3	Bezugsnormen der Leistungsbeurteilung	84
4.3.1	Die drei Bezugsnormen	84
4.3.2	Diskussion der Bezugsnormen	85
4.3.3	Angemessene Verwendung der Bezugsnormen: ein normenintegrierendes Anforderungsmodell	94
5	Benotungsmodelle und ihre Messfehler	
5.1	Basisbenotung und Ziffernbenotung	99
5.2	Merkmale eines Benotungsmodells	99
5.3	Arten von Benotungsskalen	101
5.4	Messfehler bei der Leistungsbeurteilung	104
5.4.1	Ursachen von Messfehlern	104
5.4.2	Der Stichprobenfehler	105
5.4.3	Indifferenzbereiche	107
5.4.4	Überprüfung der Messgenauigkeit einer Benotungsskala	113

5.4.5 Weitere praktische Konsequenzen	118
5.5 Probleme von Benotungsmodellen, die sich an der sozialen Norm orientieren	119
5.6 Probleme von Benotungsmodellen, die sich an der individuellen Norm orientieren	122
5.7 Festsetzen der Mindestkompetenz für ein kriteriales Benotungsmodell	123
6 Prüfungs- und Aufgabenanalyse	
6.1 Begründung und Analysegesichtspunkte	129
6.2 Praktische Hinweise zur Analyse ausgewählter Aspekte	131
6.2.1 Fehleranalyse	131
6.2.2 Schwierigkeit	132
6.2.3 Trennschärfe	134
6.2.4 Lernzielerreichung	137
6.2.5 Bedeutsamkeit der Aufgaben	138
6.2.6 Schwierigkeit und Bedeutsamkeit	140
6.2.7 Verzerrungen durch zu knappe Bearbeitungszeit	140
6.2.8 Ermüdungseffekte	142
6.2.9 Angemessenheit der Mindestkompetenz	143
6.3 Abschließende Hinweise	144
7 Beurteilung mündlicher und ganzheitlicher Leistungen und Bewertung von Diktaten	
7.1 Beurteilung ganzheitlicher Leistungen	147
7.1.1 Analyseeinheiten, Beurteilungskriterien, Ausprägungsgrade	147
7.1.2 Das Beispiel Aufsatzbeurteilung	148
7.2 Mündliche Leistungen und mündliche Prüfungen	153
7.2.1 Besonderheiten mündlichen Prüfens	153
7.2.2 Schwierigkeiten mündlichen Prüfens	154
7.2.3 Zur Organisation mündlicher Prüfungen	155
7.2.4 Beurteilung der Mitarbeit	157
7.3 Dikate und Fehlerbeurteilung	158
7.3.1 Unerträglich große Messfehler	158
7.3.2 Verschiedene Fehlerarten	160
7.3.3 Lückendikate	160
8 Verbalzeugnisse und Lernberichte	
8.1 Argumente für die Einführung von Verbalzeugnissen	163
8.2 Bedenken und Einwände gegen Verbalzeugnisse	164

8.3 Erfahrungen und Forschungsergebnisse	165
8.3.1 Messtheoretische Qualität.....	165
8.3.2 Umfang, Inhalt und Formulierung	166
8.3.3 Verständlichkeit.....	166
8.3.4 Akzeptanz.....	168
8.3.5 Wirkungen.....	169
8.4 Formen des Verbalzeugnisses.....	171
8.5 Vorschläge für die Gestaltung von Verbalzeugnissen.....	173
8.5.1 Kategorien des zu beschreibenden Verhaltens	173
8.5.2 Hinweise für die Formulierung.....	174
8.5.3 Trennscharfe Begriffe und Ausprägungsgrade.....	177
8.6 Erfassung von Entwicklungen.....	178
8.6.1 Formen und Techniken der Beobachtung	178
8.6.2 Vielfalt der Leistungs- und Entwicklungsdaten.....	181

9 Die Beurteilung schulischer Leistungen aus rechtlicher Sicht (Stephan Rademacher)

9.1 Der Beurteilungsspielraum und seine Grenzen	183
9.2 Rechtliche Anforderungen an schulische Leistungsbeurteilungen	185
9.3 Fehler im Zusammenhang mit der Leistungserhebung	185
9.4 Fehler im Zusammenhang mit der Leistungsbewertung	188
9.4.1 Verfahrensfehler	188
9.4.2 Falsche Auslegung des anzuwendenden Rechts	188
9.4.3 Irrtum über die Bewertungsgrundlage	189
9.4.4 Verstoß gegen allgemein gültige Bewertungsmaßstäbe	189
9.4.5 Einfluss sachfremder Erwägungen	190
9.5 Erheblichkeit des Fehlers	190
9.6 Schulpraktische Probleme	191
9.6.1 Täuschungen	191
9.6.2 Gruppenarbeiten	193
9.7 Folgen einer rechtswidrigen Leistungsbeurteilung	194

10 Leistungsentwicklung und Leistungserziehung

10.1 Zwei Lebensideale und Leistungsverständnisse	195
10.2 Schulpraxis am Scheideweg	196
10.2.1 Anforderungen, Prüfungen und Beurteilungen im Dienste entfremdeter Arbeit	196
10.2.2 Anforderungen, Prüfungen und Beurteilungen im Dienste der Selbstverwirklichung	197
10.2.3 Fragwürdige Vermittlungsversuche	198

10.3 Leistungserziehung als pädagogische Aufgabe	198
10.3.1 Kritische Leistungsbereitschaft als Erziehungsziel	199
10.3.2 Selbstansprüche und Fremdanforderungen	199
10.3.3 Umfassendes Leistungsverständnis	200
10.3.4 Begrenzung der Leistungserziehung	202
10.4 Individuelle Förderung	205
10.4.1 Die Förderung nach individueller Förderung	205
10.4.2 Wege individueller Förderung	205
10.4.3 Voraussetzungen individueller Förderung	212
11 Wege einer neuen Diagnosekultur	
11.1 Der künstliche Gegensatz zwischen Selektion und Förderung	217
11.2 Korrekte Leistungsdiagnose als Grundlage von Selektion und Förderung	218
11.3 Kontraproduktive Prüfungskultur als Hindernis von Förderung	218
11.4 Förderdiagnose	222
11.4.1 Grundprinzipien der Förderdiagnose	222
11.4.2 Wiedergewinnung des Formenreichtums	224
11.4.3 Selbstbeurteilung der Schüler	228
11.4.4 Beurteilung von Leistungen in Prozessen	247
11.4.5 Dokumentation von Lernergebnissen und das Portfoliokonzept	248
11.4.6 Beurteilung von kooperativ erbrachten Leistungen.....	254
11.4.7 Leistungsbeurteilung bei selbst gesteuertem Lernen.....	257
11.4.8 Leistungsbeurteilung im differenzierenden Unterricht.....	258
11.4.9 Beurteilungsgespräche.....	259
11.5 Beurteilungskonzepte	264
Literatur	267
Anhang	
Anhang 1 Bestimmung von Indifferenzbereichen	279
Anhang 2 Bestimmung der Bedeutsamkeit von Prüfungsaufgaben	284
Sachregister	285

Vorwort zur 6. Auflage

In den zwanzig Jahren nach dem Erscheinen der ersten Auflage dieses Buches hat sich die Schulleistungsdagnostik gravierend verändert. Durch Impulse der nationalen und internationalen Schulleistungsuntersuchungen (TIMSS, PISA, IGLU, DESI, VERA usw.) kam es zu bedeutenden technischen Verbesserungen: Die Orientierung an Kompetenzmodellen, Beurteilungsrastern und Bildungsstandards erfreut sich zunehmender Verbreitung. Dabei wird allerdings oft übersehen, dass die in den nationalen und internationalen Schulleistungsuntersuchungen angewandten Verfahren der Testpsychologie im Schulalltag nur eingeschränkt brauchbar sind. Jedenfalls aber haben Willkür und Subjektivität in der Leistungsdiagnose merklich an Boden verloren. Zugleich hat sich die Erkenntnis durchgesetzt, dass eine lediglich technisch perfektionierte Schulleistungsdiagnose noch zu wenig für die individuelle Förderung der Schülerinnen und Schüler austrägt: In der Konsequenz werden neue Wege einer speziellen Förderdiagnose beschritten, die sich manchmal sogar als Gegenentwurf gegen eine traditionelle, für die Förderung irrelevante oder sogar kontraproduktive Leistungsdiagnose sieht. Aber die neuen Verfahren haben auch methodische Schwächen und bedürfen sehr wohl eines Unterbaus bewährter Diagnosetechniken.

Überhaupt ist der oft aufgemachte Gegensatz zwischen traditioneller, der Selektion verpflichteter Leistungsdiagnose und neuer, der Förderung dienender Diagnose ein falscher: Selektion im Laufe eines Bildungsganges muss immer zugleich Bedingungsselektion und Personenselektion sein. Indem sie als Bedingungsselektion für die Lernenden eine möglichst optimale Lernumgebung organisiert, arbeitet sie auch unmittelbar ihrer Förderung zu. Sie kann das aber nur leisten auf der Grundlage einer korrekten Diagnose, die den üblichen messtheoretischen Anforderungen genügt. Deshalb werden in diesem Buch nach wie vor die Grundlagen und Verfahren bewährter traditioneller Diagnosetechniken behandelt. Sie enthalten – richtig eingesetzt – durchaus auch ein Förderpotential. Darüber hinaus geben wir aber auch eine Einführung in die Ansätze einer speziellen Förderdiagnose. Dabei wird deutlich, dass Probleme der Überprüfung und Beurteilung von Schulleistungen nicht losgelöst von der Unterrichtsgestaltung erörtert werden können.

Die Impulse der nationalen und internationalen Schulleistungsuntersuchungen haben nicht nur zu einer technischen Verbesserung der Schulleistungsdignose geführt, sondern auch zur Vernachlässigung der bildungspolitischen und soziologischen Zusammenhänge, in welchen sie steht. Es gibt einen verbreiteten Trend, Fragen der Schulleistungsdignose mit kühler sozialtechnologischer Rationalität zu behandeln und dabei die Gefahren zu übersehen, die durch ökonomische und neoliberale Sichtweisen und Praktiken heraufbeschworen werden. Diesem Trend möchte die neue Auflage entgegenwirken, indem sie weiterhin in einiger Aus-

führlichkeit auch Fragen der Leistungsgesellschaft und der Leistungserziehung im Lichte der Vision einer humanen und gerechten Gesellschaft diskutiert.

Die Forschungsbasis, auf welche sich das Buch bezieht, wird häufiger als etwas veraltet erscheinen. Es war leider nicht immer möglich, sich auf aktuellere Studien zu beziehen. Mindestens teilweise ist dies auch dem Umstand geschuldet, dass es einerseits in der sich beständig wandelnden geschichtlich-gesellschaftlichen Welt keine bleibenden Erkenntnisse und keine zeitlos gültigen Gesetzmäßigkeiten gibt, dass andererseits aber natürlich nicht die Ressourcen für beständige Nachuntersuchungen verfügbar sind.

Neumarkt, im Frühjahr 2014

Werner Sacher

1| Zum Sinn des Leistens, Prüfens und Beurteilens

1.1 Grundbegriffe

1.1.1 Leistung

Die sprachlichen Wurzeln von „Leistung“ – idg. „leis-“ = „Fußspur, Furche“ und got. „laistjan“ = „einer Spur nachfolgen“ (Trübner 1943, S.438ff) – deuten auf ein normatives Moment des Leistungsbegriffs: Die Spur ist vorgegeben, an ihr hat man sich zu orientieren. In diesem Sinne sind auch Leistungen von Schülern Ausdruck des Grades, in dem sie Maßstäben und Anforderungen zu entsprechen vermögen. *Leistung ist der Vollzug und das Ergebnis einer Tätigkeit, die mit Anstrengung verbunden, auf die Erlangung eines Zieles gerichtet und auf Gütemaßstäbe und Anforderungen bezogen ist.* (Klafki 1996, S.228; Schröder 2000, S.287)

In der Physik ist der Gütemaßstab der Leistung die benötigte Zeit: Leistung wird dort bekanntlich definiert als Arbeit pro Zeiteinheit. Auch die Leistung von Menschen kann an der Zeit gemessen werden, innerhalb deren sie erbracht wird. Daneben gibt es aber auch ganz andere Gütemaßstäbe, z. B. den Grad der Übereinstimmung einer Tätigkeit mit einem Ideal oder Vorbild, ihre Exaktheit und Präzision, ihre Kreativität, ihre Umweltverträglichkeit, entstehende Kosten usw.

1.1.2 Leistungsstreben

Leistungsstreben ist das Bestreben, etwas zu leisten. Alle Menschen streben nach Anerkennung durch ihre Umgebung, zumindest durch relevante Bezugspersonen. Ein wichtiger Weg, solche Anerkennung zu erlangen, führt über Leistungen, die Beifall finden. Und in dem Maße, wie Menschen diesen erhalten, wächst ihr Selbstbewusstsein und ihre Selbstachtung.

Leistungsstreben zeigt sich schon beim kleinen Kind in aller Deutlichkeit, bleibt – wenn es nicht durch schwere Frustrationen beeinträchtigt wird – im Leben des Erwachsenen bestimmend und dauert letztendlich fort bis zum Tode. Auch alte und kranke Menschen möchten noch zu irgendetwas nütze sein, und das Gefühl, etwas zu leisten, trägt ganz wesentlich zu ihrem Wohlbefinden, ja vielfach sogar zu ihrer Genesung bzw. zu ihrer Gesundheit bei.

1.1.3 Leistungsanforderung

Leistungsanforderungen nennen wir nachdrücklich artikulierte Erwartungen hinsichtlich zu erbringender Leistungen. Solche Erwartungen und die ihnen zu Grunde liegenden Maßstäbe können von außen herangetragen oder von den Leistenden selbst formuliert werden, d. h. Leistungsanforderungen können Fremd- oder Selbstanforderungen sein. Man muss sich bei der Unterscheidung von Fremd- und Selbstanforderungen allerdings im Klaren darüber sein, dass auch Selbstanforderungen in der Regel auf internalisierte Fremdanforderungen zurückgehen. Zugespitzt könnte man sagen, dass die Schule vielfach Kinder und Jugendliche dazu bringt und dazu bringen muss, „dass sie das gerne tun (aus eigener Initiative), was sie tun müssen.“ (Fend 1971, S.103)

Doch wenn auch insofern die Unterscheidung von Selbst- und Fremdanforderungen nicht ganz trennscharf ist, so hilft sie doch, einen verbreiteten Fehlschluss zu vermeiden: Dass das Leistungsstreben zum Menschsein gehört und jeder Mensch etwas leisten will, legitimiert keineswegs beliebige Fremdanforderungen, auch nicht willkürliche Leistungsanforderungen an Schüler. Gerade, wenn es ein Grundbedürfnis des Menschen ist, etwas zu leisten, sollte es solcher Anstöße von außen eigentlich gar nicht bedürfen – außer man will sich seines Leistungsstrebens bemächtigen und damit Kontrolle über ihn erlangen.

Die Legitimation von Leistungsanforderungen, welche Erzieher und pädagogische Institutionen an dafür nicht von sich aus motivierte junge Menschen herantragen, stellt das zentrale Problem der Leistungserziehung dar. Der häufig bemühte Vorgriff auf die Zukunft junger Menschen und ihr angebliches Wohl ist nicht unproblematisch. Diese Zukunft ist ja, genau besehen, nur unter großer Unsicherheit prognostizierbar. Deshalb ist ein solcher Vorgriff zumindest immer dann suspekt, wenn die erhobenen Leistungsanforderungen jungen Menschen nicht verständlich zu machen sind und nur mit Druck- und Zwangsmaßnahmen durchgesetzt werden können.

1.1.4 Leistungsorientierung

Leistungsorientierung ist die Ausrichtung des Wertsystems eines Individuums, einer Gruppe oder einer Gesellschaft auf den führenden Wert Leistung. Das Leistungsmotiv wird dann für das Leben eines Einzelnen oder eines Sozialverbandes dominant. Selbst wenn dabei Druck und Zwang fehlen, ist eine allzu ausschließliche Leistungsorientierung nicht unbedenklich: Kein Mensch kann nämlich auf die Dauer in dem Bewusstsein leben, nur so viel wert zu sein, wie er leistet. Jeder braucht auch Lebensverhältnisse, in denen er sich ohne Vorbedingungen, einfach um seiner selbst willen, angenommen und geschätzt wissen darf. So muss ein Kind, um gesund aufzuwachsen, der Zuneigung und Wertschätzung seiner Eltern auch dann sicher sein können, wenn es Misserfolge in der Schule erleidet, auch wenn

diese von ihm mitverschuldet sind. Alte und Kranke brauchen das Bewusstsein, auch dann noch angenommen und geachtet zu sein, wenn sie für ihre Umgebung keine Leistungen mehr erbringen. Das gilt ebenso für alle anderen Menschen, die auf Grund von Umständen, die sie nicht zu verantworten haben, keine Leistung erbringen.

Leistungsorientierung erfordert das kompensatorische Prinzip der Solidarität bzw. – etwas pathetisch ausgedrückt – der Prinzip der barmherzigen Liebe, das tief in der christlich-abendländischen Tradition verwurzelt ist. Barmherzige Liebe ist aber nicht nur ein Korrektiv zur Leistungsorientierung und Leistungsanforderung: Sie ist zugleich der Nährboden, auf dem Leistung sich überhaupt erst gesund entwickeln und gedeihen kann.

1.1.5 Leistungsprinzip

Von der Leistungsorientierung müssen wir das Leistungsprinzip unterscheiden. *Das Leistungsprinzip ist ein gesellschaftliches Verteilungsprinzip, welches Chancen und Positionen nach Leistung zuweist – nicht etwa nach Abstammung, Bekanntheit, weltanschaulicher Nähe usw.* Das Leistungsprinzip in seiner modernen Ausprägung entstand im Zusammenhang beginnender Demokratisierung und Industrialisierung im 18. und 19. Jahrhundert. (Nipkow 1978, S.7ff.; Furck 1975) In der Ständegesellschaft des Mittelalters war für jedes Mitglied durch die Geburt bestimmt, welche gesellschaftliche Position es erlangen konnte. Seine Leistung spielte bei der Zuteilung von Lebenschancen kaum eine Rolle. Je mehr aber die Privilegien der Geburt an Bedeutung verloren, desto mehr wurden neue Verteilungsprinzipien benötigt. Vor allem bedurfte die moderne Industriegesellschaft eines funktionalen und zweckrationalen Verteilungsprinzips für Berufs- und Lebenschancen, Besitz, Einfluss und Macht, das zugleich höchste Effektivität und Mobilität ermöglichte. Sie fand dieses im Leistungsprinzip.

Das Leistungsprinzip als neues Verteilungsprinzip trug auch dem Gleichheitspostulat der Französischen Revolution Rechnung und hatte zunächst eine emanzipatorische Funktion, vor allem für das Bürgertum: Wer etwas leistete, konnte nun seinen Weg machen, auch wenn er nicht von adeliger Abkunft war.

Ein Markstein in dieser gesellschaftlichen Neuorientierung war die Neuregelung der Stipendienvergabe nach Leistungsgesichtspunkten, vor allem in den theologischen Fakultäten, durch die preußische „Instruktion“ vom 23.12.1788: Stipendien durften nun nur noch an Studenten vergeben werden, die zuvor im Abiturientenexamen das Zeugnis der Reife erworben hatten. (Furck 1975, S.39) Ein anderer Markstein war die preußische Heeresreform von 1806 bis 1814, welche auch Bürgerlichen die Offizierslaufbahn und den Zugang zu höheren Beamtenpositionen eröffnete, die bis dahin dem Adel vorbehalten waren.

Im 19. Jahrhundert wurde das Leistungsprinzip jedoch zunehmend gegen das nachdrängende Proletariat gewendet. Dasselbe Bürgertum, das dem Leistungsprinzip seinen Aufstieg verdankte, verteidigte nun damit seine Privilegien und kontrollierte den sozialen Aufstieg der Unterschicht, indem es zusammen mit der Oberschicht die Kriterien für Leistung so definierte, wie es der Sicherung der eigenen sozialen Position und der ihres Nachwuchses dienlich war, z. B. durch die starke Betonung des sprachlichen Anteils und die weit gehende Vernachlässigung praktischer Elemente in der höheren Bildung. Das ursprünglich emanzipatorische Leistungsprinzip wurde auf diese Weise zum Repressionsinstrument. Vermutlich überhaupt nur, weil Bürgertum und Oberschicht den Bedarf an Nachwuchs für gehobene Positionen nicht vollständig decken konnten, nahmen stets auch einige Angehörige der Unterschicht am Aufstieg teil.

Das Leistungsprinzip kann offenbar lediglich dann eine gerechte Verteilungspraxis begründen, wenn Chancengleichheit bereits gegeben ist. Somit ist es nur bedingt als ein Instrument zum Abbau von Chancenungleichheit anzusehen. Bei bestehender Chancenungleichheit wird es leicht zum strategischen Instrument von Gruppen und Schichten, die über die Macht verfügen zu definieren, was als hohe Leistung gelten soll. Insgesamt scheint das Leistungsprinzip „sozial gerechte und funktional angemessene Zuordnungen und Verteilungen mehr ideologisch vorzutauschen als tatsächlich zu gewährleisten.“ (Nipkow 1978, S.10)

Die gesellschaftliche Umdeutung des Leistungsprinzips verdarb schließlich auch die mit dem Aufstieg des Bürgertums verknüpfte Bildungsidee: Statt des zu Beginn des 19. Jahrhunderts propagierten neuhumanistischen Ideals einer allseitig gebildeten Persönlichkeit beherrschte bald eine arg schematisierte und sehr äußerliche Vorstellung von Allgemeinbildung die Schule. Es kam zur Fixierung fester Jahrespensen, zum Vorherrschen überprüfbarer Fertigkeiten und Fähigkeiten, zur Abwertung beruflicher Bildung und nichtklassischer Bildungsinhalte.

Mit der wieder statisch gewordenen gesellschaftlichen Ordnung korrespondierte die Verwendung eines statischen Begabungsbegriffs: Unterschiede im Schulerfolg wurden als Ausdruck natürlicher Ungleichheit der Bevölkerungsschichten gedeutet. Diesem Begabungsbegriff entsprechend, legte die Schule die Hauptbetonung auf abschließende Leistungskontrollen. Sie versäumte es fast vollständig, diagnostische Leistungsfeststellungen durchzuführen, wie sie erforderlich gewesen wären, um kompensatorische und fördernde Maßnahmen für die weniger Leistungsfähigen zu ergreifen. Es galt ja als gesichert, dass es bei diesen nicht viel zu fördern gab. Und dahinter stand unausgesprochen, dass die Förderung von Unterschichtkindern gesellschaftlich einfach unerwünscht war, da sie ja die Privilegien des Bürgertums und der Oberschicht gefährdet hätte. Für die in der gegenwärtigen pädagogischen Diskussion erhobene Forderung nach einer Abkehr von der traditionellen Prüfungs- und Beurteilungskultur, welche einseitig auf Lernergebnisse fixiert ist und Lern- und Leistungsprozesse zu wenig beachtet, sind

viele gute Gründe geltend zu machen. (Vgl. dazu im Einzelnen Kapitel 11) Ein gewichtiger ist auch dieser, dass die traditionelle Prüfungs- und Beurteilungskultur in einer heute nicht mehr haltbaren Begabungstheorie verankert und einem überholten Gesellschaftsmodell verpflichtet ist. Dieser Einwand kann auch gegen eine allzu strikte Outputorientierung der Leistungsbeurteilung von Schülern vorgebracht werden, wie sie neuerdings im Gefolge internationaler Vergleichsstudien vermehrt praktiziert wird.

Die knappe Skizze der neuzeitlichen Geschichte des Leistungsprinzips offenbart seine enorme Ideologiefälligkeit: Als Leistung gilt immer, was mächtige gesellschaftliche Gruppen als solche definieren. Schon das lässt fraglich erscheinen, ob die Verteilung von Bildungs- und Berufschancen nach dem Leistungsprinzip ohne weiteres als gerecht gelten darf.

1.2 Zur Übertragbarkeit des Leistungsprinzips auf die Schule

Selbst Lehrkräfte halten es z. T. für ganz „natürlich“, dass die Schule sich inmitten einer Leistungsgesellschaft als Leistungsschule versteht. So einfach liegen die Dinge freilich nicht. Denn einer Übertragung des Leistungsprinzips von der gesamtgesellschaftlichen Wirklichkeit und von der Wirtschafts- und Arbeitswelt auf die Schule steht eine ganze Reihe von Schwierigkeiten entgegen.

Zum einen ist es nur eine Halbwahrheit, dass wir in einer Leistungsgesellschaft leben. Die begehrten Positionen werden nämlich auch in unserer gegenwärtigen Gesellschaft keineswegs nur nach der erbrachten Leistung zugewiesen, sondern auch nach anderen Verteilungsprinzipien:

- Neben dem Leistungsprinzip gibt es immer noch *Vorrechte der Geburt*, die zwar andere Formen angenommen haben, aber nie aufhörten zu bestehen. Jencks u. a. (1973) zeigten, dass in den westlichen Industriegesellschaften etwa die Hälfte aller schichtspezifischen Vorteile und Handicaps von den Eltern auf ihre Kinder übergehen. Die vom PISA-Konsortium wiederholt festgestellte große soziale Ungleichheit der Bildungschancen in Deutschland zeigt, dass diese Verhältnisse gerade bei uns in besonderem Maße ungebrochen fort dauern.
- Mit dem Leistungsprinzip konkurriert ferner das *Ancienitäts-Prinzip*. D. h. Positionen werden oft nach dem Alter oder nach der Dauer der Zugehörigkeit zu einer Institution oder zu einem Betrieb vergeben.
- Ein weiteres Verteilungsprinzip ist das *Ideologie-Prinzip*: Häufig ist für den Berufs- und Lebenserfolg die Zugehörigkeit zu einer Weltanschauungsgruppe oder zu einer Partei und die loyale Ergebenheit ihr gegenüber entscheidend.

- Das Ideologie-Prinzip geht mit fließenden Grenzen über in das *Bekanntheits- und Beliebtheitsprinzip*: Oft ist es ausschlaggebend, Entscheidungsträger in formellen und informellen Gruppen persönlich zu kennen oder doch wenigstens Leute zu kennen, die sie persönlich kennen. Und natürlich sollte dieses Kennen nach Möglichkeit einschließen, dass der Bewerber es verstand und versteht, sich sozial angenehm zu machen.
- Schließlich wird unsere gesellschaftliche Verteilungspraxis aber auch durch das *Sozialprinzip* bestimmt: Unser soziales Netz gewährleistet eine halbwegs erträgliche menschliche Existenz selbst dann, wenn Menschen aus Gründen, die sie nicht selbst zu verantworten haben, keine Leistung erbringen, wie es bei Alten, Kranken und Arbeitslosen der Fall ist.

Das Leistungsprinzip ist also nur eines von sechs in unserer Gesellschaft herrschenden Verteilungsprinzipien. Wir sollten sie deshalb zutreffender als eine stark leistungsorientierte Gesellschaft bezeichnen und nicht als Leistungsgesellschaft. An einer solchen Sprachregelung ist pädagogisch viel gelegen: Lehrkräfte können leicht Enttäuschungen bei jungen Menschen vorprogrammieren, wenn sie ihnen im Widerspruch zur faktischen Pluralität der Verteilungsprinzipien den Eindruck vermitteln, es käme in unserer Gesellschaft ausschließlich auf Leistung an.

Moderne Industriegesellschaften haben außerdem zunehmend Schwierigkeiten, das Leistungsprinzip anzuwenden und darauf eine gerechte Verteilungspraxis zu gründen:

- *Die Leistungen im Arbeitsleben und öffentlichen Leben werden immer komplexer.* Der Beitrag, den der Einzelne dazu liefert, ist immer weniger zu überschauen. Er ist eingebunden in eine vielschichtige und verzweigte Arbeitsorganisation mit zahlreichen sachlichen und personellen Komponenten. Es ist oft einfach unmöglich zu entscheiden, wo Effekte eines einzelnen Arbeitsbeitrags und wo Systemeffekte vorliegen. Die Lehrarbeit ist dafür ein gutes Beispiel: Ihr Ergebnis wird außer von der Kompetenz und vom Einsatz der Lehrkraft offensichtlich auch von der Zusammensetzung der Klasse, von der Arbeit anderer Kollegen, von der Ausstattung der Schule, vom Lehrplan und vielen anderen Faktoren beeinflusst.
- *In unserer pluralistischen Gesellschaft gibt es kaum noch eindeutige und allgemein anerkannte Ansichten darüber, was gute Leistung eigentlich ausmacht.* Wie sollte man sonst z. B. erklären, dass Bankmanager, Profi-Fußballer und Schlagersänger mehr verdienen als der Bundeskanzler oder der Bundesbankpräsident? Zwar mag weithin Konsens bestehen, dass Eigeninitiative und kritisches Mitdenken ebenso zu guter Arbeitsleistung gehören wie fehlerfreies, rasches Arbeiten, pünktliche und gewissenhafte Befolgung von Anweisungen, praktische Anerkennung des Betriebszwecks, Wahrung des Arbeitsfriedens und Anerkennung der Leitungsstruktur des Betriebs. Im Einzelnen kann diesen unterschiedlichen Leistungsaspekten in verschiedenen Betrieben aber höchst unterschiedliche Bedeutung beigelegt werden.

- Multinationale Konzerne agieren auf den globalen Märkten nach Strategien, die einseitig auf Aktienkurse und Profitmaximierung ausgerichtet sind und oft die Leistungen ganzer Belegschaften ignorieren. Selbst berufliche Höchstleistung sichert schon lange nicht mehr den beruflichen Aufstieg, oft nicht einmal mehr den Erhalt des Arbeitsplatzes.

Einer Übertragung des Leistungsprinzips von der Wirtschaft und vom Arbeitsleben auf die Schule stehen außerdem pädagogische Bedenken entgegen:

- Selbst wenn wir eine reine Leistungsgesellschaft mit einem einheitlichen Leistungsverständnis hätten, könnte es angemessen sein, jungen Menschen während der Schulzeit erst einmal zu helfen, ichstarke Persönlichkeiten mit Selbstvertrauen und Lebensmut zu entwickeln, bevor sie nachher dem oft schon fast unmenschlichen Leistungsdruck im Arbeitsleben ausgesetzt sind. Vielleicht können sie ihn ja gar nicht aushalten, ohne ernsten Schaden zu nehmen, wenn sie nicht zuvor eine solche Schonzeit genossen haben. *Seit wann hätte Erziehung sich bloß als Anpassung an gesellschaftliche Verhältnisse zu verstehen?*
- Der schulische Wettbewerb um Bildungs- und Lebenschancen ist in einem erheblichen Maße verzerrt. *Nicht jeder Schüler hat dieselbe Chance, sein Leistungsoptimum zu zeigen.* Allein mit dem Forträumen von Hindernissen, die einem freien Leistungswettbewerb im Wege stehen, ist es nicht getan. Auch nach der Herstellung einer solchen formalen Chancengleichheit existieren immer noch Barrieren des sozialen Milieus, die es Angehörigen mancher Bevölkerungsschichten und -gruppen erschweren oder sogar verunmöglichen, ihren Fähigkeiten entsprechende Leistungen zu erbringen. Zu den bekannten schichtspezifischen Bevorzugungen und Benachteiligungen kommt ein typisches Vermeidungsverhalten von Eltern und Kindern aus der Unterschicht, auf welches die französischen Soziologen Bourdieu und Passeron (1971) bereits in den Siebzigerjahren hinwiesen: Eltern aus unterprivilegierten Schichten und Bevölkerungsgruppen resignieren häufig schon im Vorfeld der Entscheidung. Der Übertritt ins Gymnasium z. B. wird oft gar nicht erst gewagt, wenn die Grundschulnoten nicht mit hoher Sicherheit den erfolgreichen Abschluss einer höheren Bildung erwarten lassen.

1.3 Zur Geschichte der Prüfungen, Zeugnisse und Zensuren¹

Periodische Leistungsbeurteilungen und Prüfungen gehörten nicht schon immer und nicht überall zur Schule. Obgleich sich Vorläufer schon bei den alten Ägyptern und in der Antike nachweisen lassen, kam die Schule in ihrer fünftau-

¹ Vgl. dazu von Hohenzollern/ Liedtke 1991; Fischer-Elfert 1991; Rösger 1991; Buchinger 1991; Keck 1991; Breitschuh 1991; Ziegenspeck 1973; Ziegenspeck 1999.

sendjähriĝen Geschichte viereinhalbtausend Jahre weit gehend ohne formalisierte Notensysteme und Zeugnisse aus. Die zunehmende Verbreitung von Zeugnissen und Prüfungen war aber nicht nur ein Indiz für wachsenden Druck in der Schule, sondern auch Zeichen eines zunehmenden Interesses für die Schüler und eines erstarkenden Verantwortungsgeföhls für ihre Lernprozesse.

Die wichtigsten Vorläufer der heutigen Zeugnisse sind die seit dem 16. Jahrhundert verbreiteten *Benefizienzeugnisse*. In ihnen wurde dem Empfänger testiert, dass er im Hinblick auf seine charakterlichen Eigenschaften und seine Leistungen würdig sei, an einem sog. Freitisch teilzunehmen (d. h. kostenlose Verpflegung zu erhalten) und mit einem Stipendium unterstützt zu werden. Diese Zeugnisse waren also im Grunde Empfehlungsschreiben für bedürftige Schüler. Sie nahmen unter ihnen eine Auslese im Hinblick auf ihre Förderungswürdigkeit vor. Kinder wohlhabender Eltern brauchten natürlich keine solchen Zeugnisse und waren folglich auch der Selektion durch sie nicht unterworfen.

Zu einer ersten Perfektion gelangte das Prüfungs- und Beurteilungswesen im 16. und 17. Jahrhundert an den jesuitischen Schulen. An den staatlichen höheren Schulen etablierte es sich auf breiter Front erst im 18. und 19. Jahrhundert.

Die Verordnungen und die entsprechende Praxis schwankten lange zwischen freien Formulierungen, die unseren Verbalzeugnissen ähnlich waren, und den heutigen Noten vergleichbaren Schematismen.² Beispiele für frei formulierte Zeugnisse finden sich noch weit bis ins 19. Jahrhundert herauf – auch an Gymnasien.

Auch die *Abschlussprüfung für das Gymnasium*, deren Bestehen Zulassungsbedingung für ein Universitätsstudium war, wurde erst Ende des 18./Anfang des 19. Jahrhunderts eingeföhrt. Bis dahin besuchte man das Gymnasium, solange man sich Gewinn davon versprach, und wechselte zur Universität über, wenn es einem sinnvoll schien. Auch von dieser ging man häufig ohne förmliches Examen ab, um ein Amt zu übernehmen.

Selbst das Nichtbestehen der in Preußen 1788 eingeföhrtten Abiturientenprüfung schloss zunächst noch nicht vom Universitätsstudium aus. Mit dem Reifezeugnis wurde nur der Nachweis der Würdigkeit erbracht, ein Stipendium zu erhalten. Damit verrät das Reifezeugnis seine Herkunft vom Benefizienzeugnis. Es erfüllte hinsichtlich der Studienberechtigung vorerst ebenfalls nur eine Selektionsfunktion für die unteren Schichten der Bevölkerung. Die eigentliche Intention des Staates bei der Einföhrtung des Abiturientenexamens galt auch gar nicht der Kontrolle des Zugangs zu den Universitäten. Es sollte vielmehr das Niveau des Beamtennachwuchses sichergestellt werden. (Breitschuh 1991) Die begehrten Stellen

2 Das bayerische Regulativ über „Schul-Censuren“ von 1809 verbot in tabellarischer Form verfasste Schülerzensuren und verpflichtete die Lehrer wieder zur Verwendung qualitativer Prädikate, nachdem vorher in Bestimmungen von 1793 und 1805 durchaus schon Schematismen vorgeschrieben worden waren. (Buchinger 1991, S.95ff.)

im Staatsdienst wurden nämlich nur noch an Kandidaten vergeben, welche die Abiturprüfung mit guten Ergebnissen bestanden hatten.

Dass in den 30iger Jahren des 19. Jahrhunderts (in Preußen 1834) dann schließlich doch Kandidaten ohne Reifezeugnis in den meisten deutschen Staaten der Zugang zur Universität verwehrt wurde, hatte politische Gründe: Nach den revolutionären Unruhen des Vormärz befürchteten die deutschen Fürstenhäuser einen Umsturz und versuchten auf diesem Wege letztlich, politisch missliebige Elemente von den Universitäten fern zu halten und die sich dort sammelnde Opposition auszutrocknen. Es ging ihnen weniger um die Sicherung ausreichender wissenschaftlicher Vorkenntnisse für das Studium als um die Kontrolle des Verhaltens und der Gesinnung der Studienanwärter. (Breitschuh 1991)

Im Zusammenhang dieser Entwicklung kam es auch zur Einführung eines *Zeugnisses der mittleren Reife*. Mit ihm war die Berechtigung zum einjährigen Militärdienst und der Zugang zur Offizierslaufbahn verbunden.

Die Durchsetzung der allgemeinen Schulpflicht Ende des 18./Anfang des 19. Jahrhunderts in den deutschen Ländern³ war eng verbunden mit der Ausfertigung von *Entlassungszeugnissen* (sog. Schulentlassscheinen), aus welchen die ordnungsgemäße Erfüllung der Schulpflicht hervorging. Ohne Schulentlassschein durfte niemand als Dienstbote oder Lehrling beschäftigt werden, und auch die Genehmigung, ein Haus zu erwerben oder zu heiraten, hing davon ab.

Dieser kurze Blick in die Geschichte der Zensuren, Prüfungen und Zeugnisse zeigt, dass sie schon früh der Auslese und der Sozialisation dienten, wobei diese Funktionen zunächst stärker bei den unteren Bevölkerungsschichten griffen als bei den einflussreicheren, also bereits im Ansatz schichtspezifisch verzerrt waren. Die Ablösung freier Formulierungen in den Zeugnissen durch Ziffernnoten geht im Großen und Ganzen konform mit der Ausdehnung der Selektionsfunktion auf alle Bevölkerungsschichten. Offensichtlich diente die Einführung von Ziffernnoten vor allem der Vereinfachung der Selektion. Daneben entsprang sie wohl auch dem Bemühen um größere Gerechtigkeit bei der Selektion, das freilich angesichts der nicht zu beseitigenden Ungenauigkeit und Unsicherheit schulischer Leistungsbeurteilungen⁴ nicht von Erfolg gekrönt sein konnte.

3 Versuche ihrer Einführung hatte es schon früher gegeben, in Preußen z. B. 1717 und in Bayern 1770.

4 Von Hohenzollern/ Liedtke 1991, S.31f. Vgl. auch unten 5.4.

1.4 Funktionen von Prüfungen, Zeugnissen und Zensuren

Prüfungen, Zensuren und Zeugnisse hatten im Laufe der Geschichte ganz unterschiedliche Funktionen. Manche wuchsen ihnen erst in der jüngsten Vergangenheit zu, ohne dass sie altehrwürdige andere verdrängten.

1.4.1 Selektion und Stigmatisierung

Schulleistungsüberprüfungen und -beurteilungen dienen der Auslese befähigter Anwärter auf höhere Bildungslaufbahnen, begehrte Abschlüsse und angesehene berufliche und gesellschaftliche Positionen.

Angesichts der Knappheit der begehrten Positionen ist es zwar unvermeidlich, dass bei ihrer Zuteilung und auch schon bei der Zulassung zu den Zugangswegen eine strenge Auslese erfolgt. Aber die Art, wie sie über Prüfungen, Zeugnisse und Noten abgewickelt wird, ist in mehrfacher Hinsicht problematisch:

Bei der herrschenden Selektionspraxis ist unsicher, inwieweit tatsächlich nach Leistung ausgelesen wird. Noten erlauben nämlich – wie wir noch sehen werden (vgl. 5.4) – lediglich sehr unsichere Rückschlüsse auf Schülerleistungen.

Die Selektion ist schichtspezifisch verzerrt. Schon der kurze Blick in die Geschichte zeigte, dass zunächst nicht alle Bevölkerungsschichten in gleicher Weise von der Selektionsfunktion der Prüfungen, Zeugnisse und Zensuren betroffen waren. Erst im 19. und 20. Jahrhundert erfasste die Selektionsfunktion zunehmend auch die Mittel- und Oberschicht. Es lässt sich belegen, dass ihre Verallgemeinerung noch nicht abgeschlossen ist und die Selektionsfunktion von Zeugnissen und Zensuren auch in unserer Zeit die verschiedenen Bevölkerungsschichten mit ungleicher Schärfe trifft⁵: Die PISA-Studie zeigte, dass in Deutschland Kinder der Oberschicht eine 4,28 mal so hohe Chance haben, ein Gymnasium zu besuchen, wie Arbeiterkinder. Wenn man nur Kinder gleicher kognitiver Grundfähigkeiten und gleicher Lesekompetenz berücksichtigt, bleibt diese Chance immer noch 2,96 mal so hoch. (Deutsches PISA-Konsortium 2001, S.357)

In Deutschland setzt die schulische Selektion zu sehr auf Prognose und zu wenig auf Bewährung. Leistungsprognosen sind insbesondere bei jüngeren Kindern äußerst problematisch, deren Leistungsbild oft noch unklar und in beständiger Veränderung und Entwicklung begriffen ist.

Die Selektion beginnt in unserem Bildungssystem viel zu früh. Nur in Deutschland und Österreich haben Kinder eine lediglich vierjährige gemeinsame Schulzeit. In Liechtenstein beträgt diese immerhin schon fünf Jahre, in weiteren 38 Staaten jedoch sechs und mehr Jahre! (Schmitt 2001, S.17)

Der internationale Leistungsvergleich der PISA-Untersuchung zeigte, wie problematisch ein so früher Selektionszeitpunkt ist: Je früher ausgelesen wird, umso

5 Zur aktuelleren Situation in Deutschland vgl. u. a. Bildungsbericht Deutschland 2006, S.178; Bildungsbericht Deutschland 2012, S.96, S.254, S.258; Baumann u. a. 2012, S.19.

stärker hängt das Einschlagen einer höheren Bildungslaufbahn von der Schichtzugehörigkeit ab, in PISA-Terminologie: umso höher ist der soziale Gradient. Deutschland nimmt mit einem Gradienten von 0,60 international eine traurige Spitzenstellung ein.

Dabei wäre es ein Irrtum zu glauben, dass eine frühe Selektion der Elitebildung zugute komme: Schon die IEEA-Studie der UNESCO zeigte, dass spät selektierende Bildungssysteme mit ihrer Begünstigung der Breitenförderung zugleich auch die besten Voraussetzungen für die Förderung von Hochbegabungen bieten. (Svenson 1962; Reble 1981, S.14ff.)

Auch an den PISA-Ergebnissen lässt sich die Problematik einer zu frühen Selektion verdeutlichen: Leider liegen nur für 22 Länder (darunter Deutschland) Daten zum frühesten Selektionszeitpunkt vor (OECD 2002a, S.195; OECD 2002b, S.56). Aber die Schüler derjenigen 15, welche erst im 14. Lebensjahr oder später selektieren, erzielten im PISA-Lesetest eine durchschnittliche Punktezahl von 503, die Schüler jener Länder hingegen, welche schon vor dem 14. Lebensjahr selektieren, nur eine solche von 476. Der Unterschied lässt sich statistisch absichern.⁶ Auch die Unterschiede hinsichtlich des Anteils der Risikoschüler, die nur das unterste Leseniveau I oder noch nicht einmal dieses erreichten, sind bestürzend: Hier stehen durchschnittlich 16,1% Risikoschülern in den später selektierenden Ländern 25,1% in den früher selektierenden gegenüber. Dieser Unterschied ist ebenfalls statistisch absicherbar.⁷ Dagegen werden Vertreter der deutschen Elitebildungsideologie einwenden, dass durch die in Deutschland übliche frühe Selektion die bestmögliche Förderung der Begabten gewährleistet sei. Wenn dies zuträfe, müssten aber die früher selektierenden Länder einen höheren Anteil an Spitzenschülern haben, welche das höchste Leseniveau V erreichen. Es zeigt sich jedoch, dass 8,7% Spitzenschülern in den später selektierenden Ländern lediglich 6,3% in den früher selektierenden gegenüberstehen. Dieser Unterschied ist zwar nicht statistisch absicherbar⁸, doch die Erwartung eines höheren Anteils von Spitzenschülern in den früher selektierenden Ländern lässt sich erst recht nicht belegen. Der beobachtbare Trend stützt eher gegenteilige Vermutungen.

Die Selektionsentscheidungen sind meist viel zu endgültig. Gerade bei jungen Menschen ist es völlig absurd anzunehmen, dass sich ihre Leistungsfähigkeit nicht noch infolge von Entwicklungsschüben und Umwelteinflüssen grundlegend ändern könnte. Es bedürfte eines Systems fairer Möglichkeiten, im zweiten und dritten Anlauf eine Revision einmal getroffener Selektionsentscheidungen zu bewirken. Zwar gibt es zweite und dritte Bildungswege. Aber der Einsatz und Aufwand, den ihr Beschreiten bisher erfordert, ist immer noch um ein Vielfaches höher als jener, der für das Absolvieren des ersten Bildungsweges erbracht werden muss.

6 Der Mittelwertsvergleich ergibt eine Irrtumswahrscheinlichkeit von lediglich $p = 0,031$.

7 Der Mittelwertsvergleich ergibt eine Irrtumswahrscheinlichkeit von lediglich $p = 0,019$.

8 Der Mittelwertsvergleich ergibt eine Irrtumswahrscheinlichkeit von $p = 0,214$.

Mit der Selektion ist nicht nur die Zuweisung von Berechtigungen, sondern vielfach auch die Stigmatisierung der Abgewiesenen verbunden. Die Schüler, welche bei der Auslese auf der Strecke bleiben, werden als mehr oder weniger minderwertig abgestempelt und – je nach Konjunkturlage in wechselnder Deutlichkeit – als ein Heer von Versagern ausgegrenzt, das in gesellschaftlichen und wirtschaftlichen Krisen bevorzugt als Manövriermasse benutzt wird. Die Folgen einer so frühzeitigen Stigmatisierung für die Persönlichkeitsentwicklung sind katastrophal.

1.4.2 Sozialisation

Meistens verdeckter, aber nicht weniger wirksam ist die Sozialisationsfunktion von Zensuren und Zeugnissen:

- Durch Prüfungen, Noten und Zeugnisse wird die nachwachsende Generation in die Leistungsorientierung unserer Gesellschaft eingeübt, wobei leider oft ein verengtes und überspitztes Leistungsverständnis zu Grunde liegt.
- Dass Ausleseentscheidungen sich nicht nur an den Schulleistungen, sondern auch an den Gesinnungen und Einstellungen junger Menschen orientieren, gehört keineswegs der Geschichte an. Auch heute noch ist der Schulerfolg zum Teil von angepasstem Verhalten abhängig.
- Die Notenbürokratie macht die jungen Menschen mit einer formal-bürokratischen Behandlung vertraut, die von individuellen Besonderheiten weit gehend abstrahiert und nach oft recht äußerlichen und für die Leistung belanglosen Kriterien wie Lebensalter, Geschlecht, Wohngegend und Schulart kategorisiert. In gewisser Weise ist die Notenbürokratie ein heimlicher Lehrplan zur Einübung in die bürokratische Gesellschaft.
- Zensuren und Zeugnisse sorgen für die Akzeptanz eines Zertifikatenunwesens, in dem es primär darauf ankommt, Leistungen in amtlichen Dokumenten testiert zu bekommen oder irgendwann einmal testiert bekommen zu haben, und in dem der Nachweis tatsächlicher Kompetenz nachgeordnete Bedeutung hat.
- Sie leisten ferner einen entscheidenden Beitrag zur „Abkühlung“, zum sog. „cooling out“⁹

Auf dieses „Cooling out“ soll etwas ausführlicher eingegangen werden:

Jede leistungsorientierte demokratische Gesellschaft ist mit dem Problem konfrontiert, dass sie zunächst möglichst alle ihre Mitglieder im Höchstmaß motivieren muss, begehrte Positionen anzustreben und dafür Leistungen zu erbringen, dass sie dann aber bei den in diesem Wettbewerb Unterliegenden die angeheizte Motivation wieder „abkühlen“, ihnen eine realistische Selbsteinschätzung vermitteln und sie mit ihrem Misserfolg aussöhnen muss.

⁹ Das Theorem des „cooling out“ geht zurück auf den amerikanischen Soziologen Erving Goffman (1952) und wurde von Burton R. Clark (1973) und Rudolf Schmid (1977) weiterentwickelt.

Die anfängliche Motivierung aller, zum Wettbewerb um die begehrten Positionen anzutreten, ist nötig, damit die Gesellschaft sich bei der Selektion auf eine möglichst breite Basis stützen kann und wirklich die Besten von allen in die ranghöchsten Positionen gelangen. Davon hängt schließlich die Konkurrenzfähigkeit auf dem Weltmarkt ab. Bei den Erfolglosen muss darauf aber eine Abkühlung der Motivation, eben das „cooling out“, folgen, damit sie ihr allzu positives Selbstkonzept korrigieren, ihre Grenzen erkennen und ihre Aufstiegserwartungen entsprechend anpassen. Andernfalls würden Neidgefühle überhandnehmen, und es wären letztes Endes der soziale Friede und der gesellschaftliche Zusammenhalt gefährdet.

Für diese Abkühlung benutzt eine Gesellschaft gewöhnlich folgende Strategie:

- Den weniger Erfolgreichen werden alternative Ziele angeboten, deren soziales Ansehen nicht allzu weit unter dem der ursprünglichen Ziele liegt. Wer z. B. die allgemeine Hochschulreife nicht erwirbt, kann das Fachabitur machen und an einer Fachhochschule Diplomingenieur F. H. werden statt Diplomingenieur Univ. an einer Technischen Universität, wenn auch zunächst mit einem geringerem Einkommen und mit einer weniger verantwortungsvollen Tätigkeit.
- In schrittweisem „Degagement“ wird die Fixierung der weniger Erfolgreichen auf die höheren Ziele gelockert. Dies geschieht z. B. durch permanente Anstöße zu einer Revision der ursprünglich zuversichtlichen Selbstbeurteilung. Wenn wir schwächere Schüler durch häufige Leistungsüberprüfungen und -bewertungen beständig mit ihrer geringen Leistungsfähigkeit konfrontieren, dann hat das genau diese Wirkung des „Degagements“: Der Schüler, dem wir immerzu die Nase auf seine schlechten Leistungen stoßen, denkt am Ende, was er denken soll: „Ach Gott, wie bin ich dumm.“ Damit das funktioniert, muss er sich permanent mit den Besseren vergleichen. Einen solchen Vergleich regen Ziffernoten sehr viel stärker an als Verbalzeugnisse; denn sie weisen den Rangplatz in der Gruppe so unmissverständlich zu, dass auch der Letzte begreift, wo er steht.
- Die Ablehnung der weniger Leistungsfähigen bei der Bewerbung um höhere Positionen wird nach Möglichkeit einer Objektivierung unterworfen. Man kann eine Ablehnung eher akzeptieren, wenn man glaubt, dass sie nicht persönlich gemeint ist und dass sie nicht auf willkürlichen Entscheidungen beruht. Um wenigstens den Anschein der Objektivität zu erzeugen, ist es günstig, Entscheidungen auf wissenschaftliche oder angeblich wissenschaftliche Grundlagen zu stellen und möglichst exakt zu begründen. In diesem Zusammenhang kommt der Scheinobjektivität und Scheinexaktheit der Ziffernbenotung enorme Bedeutung zu: Kritische Rückfragen von Eltern und Schülern werden gewöhnlich schon im Keim erstickt, wenn eine Lehrkraft vorrechnen kann, dass sich eine umstrittene Zeugnisnote durch Mittelung einer Reihe von Einzelnoten exakt – möglichst auf mehrere Nachkommastellen – so ergibt, wie sie erteilt wur-

de.¹⁰ Auch die immer noch zu findende (problematische!) Benotung nach der Normalverteilung¹¹ umgibt sich gerne mit der Gloriole eines wissenschaftlichen Verfahrens.

- Für die weniger Erfolgreichen werden sogen. „Trostspender“, z. B. in der Gestalt von Beratern, zur Verfügung gestellt. Sie verweisen auf alternative Ziele und Laufbahnen und helfen, Misserfolge zu verarbeiten. Die in unserem Schulsystem in den letzten Jahren auf- und ausgebauten psychologischen Dienste erfüllen neben mancherlei begrüßenswerten Aufgaben auch diese, die Erfolglosen mit ihrem Misserfolg zu versöhnen und ihn durch Umlenkung auf andere Laufbahnen weniger fühlbar zu machen.
- Schließlich ist man um Vermeidung aller eindimensionalen Modelle von Leistung, Aufstieg und Karriere bemüht. Jeder Fähigkeit wird an ihrem Ort ein eigener Wert zugeschrieben. So konnte ein früherer bayerischer Kultusminister sagen, den verschiedenen weiterführenden Schulen und ihren Abschlüssen komme trotz fehlender Gleichartigkeit dennoch Gleichwertigkeit zu – nach dem Motto „Alle sind gleich, nur manche sind gleicher!“

Besondere Beachtung verdient der Umstand, dass es für das Funktionieren der Abkühlung durch Prüfungen, Zeugnisse und Noten nicht einmal erforderlich ist, dass diese den Leistungsstand richtig wiedergeben.

Damit Abkühlung erreicht wird, genügt es, Leistungen sehr oft zu überprüfen und zu beurteilen (möglichst in Form von Ziffernnoten) und dies mit den Anspruch wissenschaftlicher Exaktheit und Objektivität zu verbinden. Dann wird ein junger Mensch schließlich auch an die Gerechtigkeit der Auslese und des gesamten Schul- und Gesellschaftssystems glauben – selbst wenn ihm in Wirklichkeit keine Gerechtigkeit widerfährt.

Ist Abkühlung also eine manipulative Befriedungsstrategie, die wir ablehnen müssen? Ganz so einfach liegen die Dinge nicht. Das von Goffman und seinen Interpreten beschriebene Problem existiert, und es muss irgendwie gelöst werden. Insofern sind auch die Beiträge der Schule zum „cooling out“ zunächst wertfrei zu sehen. Problematisch werden sie erst dadurch, dass die Selektion, bei welcher ein Teil der Schüler abgewiesen wird, zugleich stigmatisiert, zu früh erfolgt, schichtspezifisch verzerrt und viel zu endgültig ist. Eine weit gehende Entschärfung der Problematik wäre zu erreichen, wenn man die Funktionen des Prüfens und Beurteilens dem Leitziel der Förderung unterordnen und frühzeitig die Selbstbeurteilung der Schüler anregen und kultivieren würde. (Vgl. unten 11.4.3.)

¹⁰ Zur Unzulässigkeit solcher Berechnungen vgl. 2.1.

¹¹ Siehe dazu die Einwände in 2.1.

Wenn Ziffernnoten in unserem Schulsystem so beharrlich verteidigt werden, dann geschieht dies wohl vor allem um ihrer Sozialisationsfunktion willen. Die Sozialisationsfunktion der Schule würde nämlich durch die Abschaffung der Ziffernnoten stark beeinträchtigt.

1.4.3 Legitimation

Leistungserhebungen und -beurteilungen dienen auch der Legitimation bildungspolitischer, administrativer und unterrichtlicher Entscheidungen.

Ergebnisse sorgfältiger empirischer Begleituntersuchungen von Schulprojekten sind im Allgemeinen zu vielschichtig, um von einer breiten Öffentlichkeit rezipiert zu werden. Auch Bildungspolitiker brauchen in einer Demokratie eine Massenbasis. Sie greifen deshalb gewöhnlich zu recht einfachen Argumenten, um ihre Entscheidungen zu legitimieren. Schulnoten haben in diesem Zusammenhang wegen ihrer allgemeinen Verständlichkeit große Bedeutung. Wenn man z. B. auf besonders gute Abiturnotendurchschnitte im eigenen Bundesland verweisen kann, darf man sich der Sympathien des Wählers sicher sein.

Auch Schulen und Lehrer pflegen sich über Noten zu legitimieren: Wenn diese im Großen und Ganzen nicht allzu gut sind, dann wurde offenkundig Leistung abgefordert. Wenn sich außerdem schlechte Noten nicht allzu sehr häufen, dann hat man den Schülern anscheinend auch etwas beigebracht und keine überzogenen Anforderungen gestellt.

1.4.4 Kontrolle

Prüfungen, Zeugnissen und Noten dienen auch der Kontrolle der Lehrkräfte, der Lehrpläne, ganzer Schulen, Schularten und Schulsysteme.

Im Einzelnen ist das Verhältnis zwischen Notenniveau und unterstellter pädagogischer Qualität von Lehrplänen, Schule und Unterricht allerdings widersprüchlich: Zwar gelten im Allgemeinen gute Noten als Indiz für erfolgreiches Lehren und Lernen. Zugleich aber wird ein gewisser Anteil an schlechten Noten als Ausweis dafür genommen, dass etwas gefordert wurde. Dementsprechend muss eine Lehrkraft sowohl peinliche Untersuchungen über sich ergehen lassen, wenn der Notendurchschnitt in einer Prüfung sehr schlecht ist, als auch dann, wenn die Noten überdurchschnittlich gut ausfallen. Am unbedenklichsten ist es, keine irgendwie auffälligen Noten zu vergeben.

1.4.5 Prognose

Die meisten „Abnehmer“ von Zeugnissen und Noten, auch die betroffenen Schüler selbst, interpretieren diese prognostisch, d. h. sie leiten Erwartungen hinsichtlich des weiteren Lernfortschritts und der künftigen Leistung daraus ab. Soweit Prüfungsergebnisse und Noten in amtlichen Statistiken verarbeitet werden, die-

nen sie auch als gesellschaftliche und wirtschaftliche Planungsdaten. Man versucht, daraus Aufschluss über das verfügbare Potenzial an Begabungen und Nachwuchskräften in verschiedenen Branchen zu gewinnen und Konsequenzen für die Bildungs- und Beschäftigungspolitik zu ziehen.

Bei alledem ist es von nachgeordneter Bedeutung, ob Lehrer ihren Zeugnissen und Noten eine prognostische Bedeutung beimessen oder nicht, d. h. ob sie solche Voraussagen überhaupt machen wollen. Ihre Zeugnisse und Noten werden ihnen gewissermaßen aus der Hand genommen und auch ohne ihr Zutun prognostisch interpretiert.

1.4.6 Information und Rückmeldung

Prüfungsergebnisse, Noten und Zeugnisse haben auch die Funktion, Schüler, Eltern, sonstige Erziehungsberechtigte, potenzielle Arbeitgeber, betriebliche Ausbilder und Lehrkräfte aufnehmender Klassen und Schularten über den erreichten Lernstand und die gemachten Lernfortschritte zu informieren.

Dadurch, dass Zeugnisse und Noten Schülern eine Rückmeldung über ihren Leistungsstand geben, erfahren diese, wie sie innerhalb des Leistungskontinuums der Klasse liegen, in welchem Ausmaß sie den Anforderungen der Schule entsprechen, ob sie ihre Lernanstrengungen intensivieren müssen oder ob sie davon ausgehen dürfen, dass der bisherige Arbeitseinsatz genügt.

Eine solche Rückmeldung ist im Allgemeinen wirksamer, wenn sie sich nicht auf die Information über die bisher erzielte Leistung beschränkt, sondern auch Hinweise zur weiteren Gestaltung des Lern- und Bildungswegs und ermutigende Bemerkungen enthält. Rückmeldungen über Misserfolge müssen psychologisch viel einfühlsamer gestaltet werden als solche über Erfolge. Rückgemeldete Erfolge sind meist wirklich Anreiz zu weiteren Lernanstrengungen. Unsensibel rückgemeldete Misserfolge aber führen leicht dazu, dass Schüler ihre Anstrengungen mehr und mehr aufgeben und schließlich resignieren. Vor allem bei den Leistungsschwächeren kommt es oft zu verheerenden Auswirkungen auf die Lern- und Leistungsmotivation und auf die weitere Persönlichkeitsentwicklung.

1.4.7 Disziplinierung

Die Rückmeldungen, welche Schüler über ihre Leistungen erhalten, haben oft auch einen disziplinierenden Effekt: Sie können z. B. eine realistischere Selbsteinschätzung und vermehrte Anstrengungsbereitschaft bewirken, wo sich Leichtsinn und Oberflächlichkeit breit zu machen drohen. Insoweit ist es noch durchaus legitim, einer Klasse oder einzelnen Schülern anhand des Ergebnisses einer Prüfung zu demonstrieren, dass zusätzliche Anstrengung, bessere Aufmerksamkeit, höhere Sorgfalt oder eine andere Verbesserung des Lernverhaltens vonnöten ist, um den Anforderungen zu genügen. Die Grenze zur Manipulation wird aber überschrit-