

Katrin Lintorf

Wie vorhersagbar sind Grundschulnoten?

Prädiktionskraft individueller
und kontextspezifischer Merkmale

LEISTUNG BILDUNG LEHREN SOZIALISATION JUGEND REFORM ERZIEHUNG
IDENTITÄT GESCHLECHT FAMILIE KULTUR SCHULE ARBEIT GEWALT LERNEN
SEXUALITÄT UNTERRICHT RELIGION ALTER EVALUATION GENERATION SOZIAL
STRUKTUR MEDIEN UMWELT KINDHEIT METHODEN PISA KRIMINALI
LITÄT FREIZEIT INSTITUTIONEN ELTERN UNGLEICHHEIT LEISTUNG
BILDUNG LEHREN SOZIALISATION JUGEND REFORM ERZIEHUNG IDENTITÄT
GESCHLECHT FAMILIE KULTUR SCHULE ARBEIT GEWALT LERNEN SEXUALITÄT
UNTERRICHT RELIGION ALTER EVALUATION GENERATION SOZIALSTRUKTUR
MEDIEN UMWELT KINDHEIT METHODEN PISA KRIMINALITÄT FREIZEIT
INSTITUTIONEN ELTERN UNGLEICHHEIT LEISTUNG BILDUNG LEHREN SOZIA
LISATION JUGEND REFORM ERZIEHUNG IDENTITÄT GESCHLECHT FAMILIE
KULTUR SCHULE ARBEIT GEWALT LERNEN SEXUALITÄT UNTERRICHT RELIGIO
ALTER EVALUATION GENERATION SOZIALSTRUKTUR MEDIEN UMWELT
KINDHEIT METHODEN PISA KRIMINALITÄT FREIZEIT INSTITUTIONEN ELTERN
UNGLEICHHEIT LEISTUNG BILDUNG LEHREN SOZIALISATION JUGEND REFORM
ERZIEHUNG IDENTITÄT GESCHLECHT FAMILIE KULTUR SCHULE ARBEIT PISA

Katrin Lintorf

Wie vorhersagbar sind Grundschulnoten?

Katrin Lintorf

Wie vorhersagbar sind Grundschulnoten?

Prädiktionskraft individueller
und kontextspezifischer Merkmale



VS VERLAG

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

Dissertation Technische Universität Dortmund, 2010

1. Auflage 2012

Alle Rechte vorbehalten

© VS Verlag für Sozialwissenschaften | Springer Fachmedien Wiesbaden GmbH 2012

Lektorat: Dorothee Koch | Britta Göhrisch-Radmacher

VS Verlag für Sozialwissenschaften ist eine Marke von Springer Fachmedien.

Springer Fachmedien ist Teil der Fachverlagsgruppe Springer Science+Business Media.

www.vs-verlag.de



Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Umschlaggestaltung: KünkelLopka Medienentwicklung, Heidelberg

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Printed in Germany

ISBN 978-3-531-18408-1

Inhaltsverzeichnis

Abbildungsverzeichnis	9
Tabellenverzeichnis	11
1 Einleitung	15
2 Zur individuellen und gesellschaftlichen Bedeutung von Schulnoten ..	19
2.1 Historische Entwicklung der Schülerbeurteilung.....	19
2.1.1 Zeugnisse.....	19
2.1.2 Schulnoten	22
2.1.3 Fazit	24
2.2 Aktuelle Rahmenbedingungen der Notengebung	25
2.2.1 Aktuelle Gesetzeslage.....	25
2.2.2 Funktionen von Schulnoten	29
2.2.3 Konsequenzen von Schulnoten.....	32
2.2.4 Fazit	35
3 Messtheoretische Güte von Schulnoten	37
3.1 Befundlage	37
3.1.1 Objektivität.....	38
3.1.2 Reliabilität	44
3.1.3 Validität	45
3.1.4 Fazit	50
3.2 Zum Verhältnis von Testleistung und Schulnote	51
3.2.1 Leistungstests zur Kriteriumsvalidierung von Schulnoten?.....	52
3.2.2 Was messen Schulleistungstests – eine Testkritik?	53
3.2.3 Sind Tests fairer als Schulnoten?.....	58
3.2.4 Fazit	59
3.3 Rekonzeptualisierung von Güte	60
3.3.1 Die Rolle der Konsequenzen bei der Validierung	60
3.3.2 Assessment im Klassenzimmer	63
3.3.3 Fazit	66

4	Korrelate von Schulnoten	67
4.1	Schulnoten als Lehrerurteil: Benotungspraxis	67
4.1.1	Einflussfaktoren von Schulnoten	68
4.1.2	Unterschiede in der Benotungspraxis: Lehrerunterschiede.....	77
4.1.3	Unterschiede in der Benotungspraxis: Klassen-/Subgruppen- unterschiede.....	81
4.1.4	Fazit	83
4.2	Befunde aus Korrelationsstudien	84
4.2.1	Prädiktoren von Schulnoten.....	85
4.2.2	Variabilität der Notengebung	89
4.2.3	Fazit	93
5	Synthese und Ableitung von Hypothesen	95
5.1	Individuelle Prädiktoren der Schulnoten.....	95
5.1.1	Schulleistung und leistungsnahe Prädiktoren	95
5.1.2	Leistungsferne Prädiktoren.....	97
5.2	Prädiktoren der Aggregatebene: Unterschiede des Notenniveaus	98
5.2.1	Bezugsgruppeneffekt und weitere Kompositionseffekte	98
5.2.2	Lehrermerkmale.....	99
5.3	Variabilität der Notengebung.....	100
5.3.1	Spezifische Gewichtung auf der Klassen- und Lehrerebene.....	100
5.3.2	Subgruppenunterschiede: Leistungsspezifische Gewichtung	101
6	Datengrundlage	103
6.1	Die Studie TIMSS 2007.....	103
6.1.1	Studiendesign in Deutschland.....	103
6.1.2	Der TIMSS-Test	104
6.2	Stichprobe.....	106
6.2.1	Fallausschlüsse für die vorliegende Untersuchung.....	106
6.2.2	Gewichtung.....	109
6.2.3	Beschreibung der finalen Analytestichprobe.....	110
6.3	Instrumente und erhobene Daten	112
6.3.1	Erhebung auf Schülerebene	113
6.3.2	Erhebung auf Lehrerebene.....	115
7	Methodische Vorgehensweise	119
7.1	Umgang mit fehlenden Werten	119
7.1.1	Auswahl der Imputationsmethode	119
7.1.2	Multiple Imputation	122
7.2	Mehrebenenanalyse	125

7.2.1	Besonderheiten von Mehrebenen Daten.....	126
7.2.2	Das Random-Coefficient Model (RCM) am Beispiel der linearen Regression.....	129
7.2.3	Parameterschätzung: Schätzmethoden und Stichprobengröße.....	131
7.2.4	Metrik der Variablen: Zentrierung.....	133
7.2.5	Modellgüte und -interpretation.....	142
7.2.6	Analysestrategie.....	147
8	Ergebnisse.....	149
8.1	Deskriptive Ergebnisse.....	149
8.1.1	Kennwertbeschreibung der Analysevariablen.....	150
8.1.2	Interkorrelationen.....	151
8.2	Nullmodelle: Varianzzerlegung für die Individualmerkmale.....	155
8.3	Mehrebenenmodelle mit mehreren festen Parametern.....	157
8.3.1	Aufbau des Individualmodells.....	157
8.3.2	Kommunalitätenanalyse.....	163
8.3.3	Exkurs: Geschlechterspezifische Notenunterschiede.....	165
8.3.4	Subgruppenspezifische Benotung.....	171
8.4	Random-Intercept-Modelle.....	177
8.5	Random-Slope-Modelle.....	182
8.5.1	Überprüfung der Random Slopes.....	183
8.5.2	Erklärung des Random Slopes.....	190
9	Zusammenfassung und Diskussion.....	203
9.1	Zusammenfassung und Interpretation der Befunde.....	203
9.1.1	Schulnoten-Korrelate.....	203
9.1.2	Individuelle und klassenspezifische Variabilität.....	211
9.1.3	Zur praktischen Signifikanz der Befunde.....	217
9.2	Implikationen für die Praxis.....	219
9.2.1	Noten als Selektionsinstrument.....	219
9.2.2	Der Lehrer – ein schlechter Diagnostiker oder guter Pädagoge?...	225
9.3	Ausblick für die weitere Forschung.....	230
9.3.1	Methodische Verbesserungsmöglichkeiten.....	230
9.3.2	Neue Richtungen für zukünftige Forschung.....	235
Literatur	239
Anhang	261

Abbildungsverzeichnis

Abbildung 2.1: Erläuterung der Notenstufen bei Schulzeugnissen und Einzelergebnissen in staatlichen Prüfungszeugnissen (KMK, 1968, S. 1)	26
Abbildung 3.1: Progressive Matrix der Validität (Messick, 1989a, 1989b, 1990)	61
Abbildung 7.1: Regression innerhalb der Gruppen (innerhalb der Ellipsen), Regression zwischen den Gruppen (fett) und Gesamtregression (gestrichelt) der Mathematikleistung auf den sozioökonomischen Status (nach Raudenbush & Bryk, 2002, S. 137)	128
Abbildung 7.2: Graphische Veranschaulichung einer Zentrierung am Gesamtmittelwert für ein Modell mit variabler Konstante (links) und ein Modell mit variabler Konstante und variabler Steigung (rechts)	134
Abbildung 7.3: Graphische Veranschaulichung einer Zentrierung am Gruppemittelwert für ein Modell mit variabler Konstante (oben) und ein Modell mit variabler Konstante und variabler Steigung (unten)	136
Abbildung 7.4: Regression innerhalb der Gruppen (innerhalb der Ellipsen), Regression zwischen den Gruppen (fett) und Gesamtregression (gestrichelt) der Mathematikleistung auf den sozioökonomischen Status (nach Raudenbush & Bryk, 2002, S. 140)	139
Abbildung 7.5: Illustration der Homogenisierung der Steigungsparameter bei einer Zentrierung am Gesamtmittelwert: OLS-Schätzer, links vs. EB-Schätzer, rechts (nach Raudenbush & Bryk, 2002, S. 146)	141
Abbildung 8.1: Kommunalitätenanalyse für die Varianzaufklärung auf der Individualebene bei der Vorhersage der Mathematik- und Sachkundenote: Spezifische Varianzanteile	164
Abbildung 8.2: 50%- und 95%-Konfidenzintervall (KI) um den jeweils geteilten Steigungsparameter (fester Effekt)	189

Abbildung 9.1: Verteilung der Klassenmittelwerte in Mathematik und in den Naturwissenschaften (mit minimaler und maximaler Schülerleistung in der betreffenden Klasse)	216
--	-----

Tabellenverzeichnis

Tabelle 4.1:	Ergebnis der explorativen Faktorenanalyse von Frary, Cross und Weber (1993, S. 27).....	70
Tabelle 4.2:	Die Bedeutung verschiedener Schülermerkmale bei der Notenvergabe in Mathematik bzw. im Sprachunterricht (nach McMillan, Myran & Workman, 2002, S. 207).....	71
Tabelle 4.3:	Berücksichtigte Faktoren bei der Benotung (nach Cizek, Fitzgerald & Rachor, 1995, S. 167).....	72
Tabelle 4.4:	Relative Wichtigkeit verschiedener Merkmale bei der Benotung: Mittelwerte und Standardabweichungen der Ratings von Lehrern und Schülern (Zeidner, 1992, S. 232)	74
Tabelle 4.5:	Facetten und zugeordnete structs in den Studien von Randall und Engelhard (2009a, 2009b, 2010a, 2010b).....	75
Tabelle 6.1:	Demographische Merkmale der Schülerstichprobe (N = 4238).....	111
Tabelle 6.2:	Demographische Merkmale der Lehrerstichprobe (je N = 222).....	112
Tabelle 7.1:	Algebraische Veranschaulichung einer Zentrierung am Gesamtmittelwert (Ersetzung von x_{ij} durch $x_{ij} - \bar{X} + \bar{X}$).....	135
Tabelle 7.2:	Algebraische Veranschaulichung einer Zentrierung am Gruppenmittelwert (Ersetzung von x_{ij} durch $x_{ij} - \bar{X}_j + \bar{X}_j$).....	137
Tabelle 7.3:	Nachweis eines Kompositionseffektes mittels Zentrierung am Gruppenmittelwert vs. am Gesamtmittelwert	139
Tabelle 7.4:	Effekt der Zentrierung auf den Achsenabstand.....	140
Tabelle 8.1:	Interkorrelationen der Individualmerkmale (N = 4328).....	152
Tabelle 8.2:	Interkorrelationen der Aggregatmerkmale und der genuinen Ebene-2-Merkmale – Mathematik- und Sachkundedatensatz (N = 222)	154
Tabelle 8.3:	Klassenspezifische Korrelationen der Schulnoten mit den Individualprädiktoren (N = 222).....	155
Tabelle 8.4:	Nullmodelle – Varianzkomponentenzerlegung und Intraklassenkorrelation (unstandardisierte Lösung).....	156

Tabelle 8.5:	Aufbau des Individualmodells zur Vorhersage der Mathematiknote (unstandardisierte Regressionskoeffizienten b).....	159
Tabelle 8.6:	Aufbau des Individualmodells zur Vorhersage der Sachkundenote (unstandardisierte Regressionskoeffizienten b).....	162
Tabelle 8.7:	Modelle zur Vorhersage der in den Individualmodellen identifizierten gruppenzentrierten Notenprädiktoren durch das Schülergeschlecht (unstandardisierte Regressionskoeffizienten b).....	166
Tabelle 8.8:	Mediationsanalyse bei der Mathematiknote (unstandardisierte Regressionskoeffizienten b)	168
Tabelle 8.9:	Mediationsanalyse bei der Sachkundenote (unstandardisierte Regressionskoeffizienten b)	170
Tabelle 8.10:	Moderationsanalyse für eine leistungsgruppenspezifische Benotung in Mathematik (unstandardisierte Regressionskoeffizienten b).....	173
Tabelle 8.11:	Moderationsanalyse für eine leistungsgruppenspezifische Benotung in Sachkunde (unstandardisierte Regressionskoeffizienten b).....	175
Tabelle 8.12:	Aggregatmodell des Random-Intercept-Modells für die Mathematiknote (unstandardisierte Regressionskoeffizienten b).....	179
Tabelle 8.13:	Aggregatmodell des Random-Intercept-Modells für die Sachkundenote (unstandardisierte Regressionskoeffizienten b).....	181
Tabelle 8.14:	Finales Random-Slope-Modell für die Mathematik- und Sachkundenote (unstandardisierte Regressionskoeffizienten b).....	186
Tabelle 8.15:	Aggregatmodell zur Erklärung des zufälligen Steigungsparameters der Mathematikleistung im unteren Leistungsquartil (unstandardisierte Regressionskoeffizienten b).....	193
Tabelle 8.16:	Aggregatmodell zur Erklärung des zufälligen Steigungsparameters der Lernmotivation (unstandardisierte Regressionskoeffizienten b)	194
Tabelle 8.17:	Aggregatmodell zur Erklärung des zufälligen Steigungsparameters der Naturwissenschaftsleistung (unstandardisierte Regressionskoeffizienten b)	196

Tabelle 8.18:	Aggregatmodell zur Erklärung des zufälligen Steigungsparameters der Gewissenhaftigkeit (unstandardisierte Regressionskoeffizienten b)	199
Tabelle 8.19:	Finalmodelle zur Erklärung der Mathematik- und Sachkundenote	200

1 Einleitung

Die kritische Betrachtung von Schulnoten hat eine lange Tradition. Einen Höhepunkt findet diese Kritik im letzten Jahrhundert in zwei Wellen der empirischen Auseinandersetzung mit Schulnoten (Ingenkamp, 1993b, 1995). Zentrales Ergebnis dieser Forschung sind vielfältige Belege dafür, dass Schulnoten messtheoretischen Gütekriterien nicht genügen, und die damit verbundene Schlussfolgerung, dass Schulnoten kein aussagekräftiges Maß für Schulleistung sind. Hinzu kommen unerwünschte Nebenwirkungen wie etwaige Selbstkonzeptprobleme sowie ihre relative Informationsarmut, die sich in fehlenden Aussagen über bereichsspezifische Stärken und Schwächen äußert.

Trotz der fundamentalen Notenkritik werden Schulnoten immer noch als Maß der schulischen Leistungsbeurteilung eingesetzt. Oelkers (2002) schreibt: „26 Jahre später gibt es immer noch Noten, ist keineswegs, wie Ingenkamp seinerzeit forderte, ein ‚besseres diagnostisches Instrumentarium‘ entwickelt worden“. Das „hauptsächliche Instrument zur Beschreibung von Leistung ist immer noch das Notenschema. Das gilt mit Zunahme des Alters des Schülers umso mehr: Je höher die Schulstufe, desto strikter wird das Notenschema verwendet“ (S. 1). So beschert der fortwährende Einsatz von Zensuren zur schulischen Leistungsbewertung dem Thema eine dauerhafte Aktualität. Zumindest in der Öffentlichkeit herrscht immer noch reges Interesse am Thema der Notengebung, da davon Schüler¹ und ihre Eltern täglich betroffen sind (s. den Beitrag von Esser, 2006 in Focus-Schule). Aus bildungsbiographischer Sicht sind gute Schulnoten heute wichtiger denn je, regeln sie doch den Zugang zu weiterführenden Schulen und auch zu den tertiären Bildungswegen. Mit dem Wissen um die Wirkung verschiedener Schulformen als differentielle Entwicklungsmilieus (Baumert, Stanat & Watermann, 2006) und dem Wissen um die geringe Durchlässigkeit des deutschen Schulsystems (Bellenberg, 1999) kommt dem Thema Schulnoten somit nicht nur eine hohe Aktualität, sondern auch eine kritische Relevanz zu.

Dies sollte Schulnoten auch für eine neuerliche Betrachtung durch die Wissenschaft interessant machen. Tatsächlich jedoch scheint die Wissenschaft das Interesse an Schulnoten verloren zu haben. Aktuelle Literatur zur Notengebung

1 Zur Verbesserung der Lesbarkeit wird in dieser Arbeit auf zusätzliche weibliche Personenbezeichnungen verzichtet. Die männliche Form ist durchweg als geschlechtsneutral zu verstehen.

beschränkt sich meist auf Ratgeber oder Lehrbücher, die die gängige Forschung rezipieren (z.B. Meyenberg, 1992; Jäger, 2004; Jürgens & Sacher, 2008). Nur sehr wenige neue Studien beschäftigen sich mit dem Thema (z.B. Treutlein, Roos & Schöler, 2008; Hochweber, 2010). Zumeist wird die Notengebung am Rande anderer Forschungsfragen thematisiert wie in der Forschung zum Grundschulübergang (Bos et al., 2004; Maaz et al., 2008) oder zu Verbalzeugnissen (Valtin, 2002). Eine neuerliche Erforschung der Schulnoten ist aber aus verschiedenen Gründen sinnvoll. *Erstens* sind viele der etablierten und heute noch zitierten Befunde inzwischen mehrere Jahrzehnte alt. Auch wenn ihre weitere Gültigkeit sehr wahrscheinlich ist, bedarf dies einer Replikation. *Zweitens* untersuchen viele Notenstudien ganze Schulklassen. Insbesondere die älteren Studien vernachlässigen bei der Auswertung aber die hierarchische Struktur der Daten, was die darauf gegründeten Schlussfolgerungen entwerten kann. Methodische Weiterentwicklungen im Bereich der Mehrebenenanalyse und deren Implementation in gängiger Software erlauben heute eine verhältnismäßig einfache Berücksichtigung der Datenklumpung und bieten somit eine methodisch adäquatere Herangehensweise an Fragen zur Benotung. *Drittens* ist die Frage nach einer differentiellen Benotung bisher kaum untersucht. Die meisten Studien konzentrieren sich auf den Zusammenhang zwischen Schulnoten und Testleistung. Nur in seltenen Fällen wird dabei geprüft, ob dieser Zusammenhang variabel ist. Für andere Korrelate der Notengebung wird dies gar nicht geprüft. So wäre es denkbar, dass sich die Noten verschiedener Schülergruppen unterschiedlich zusammensetzen. Dabei können solche Gruppen nicht nur durch Merkmale wie z.B. das Geschlecht definiert sein, sondern auch durch die Zugehörigkeit zu einem bestimmten Kontext wie z.B. der Schulklasse. Auch Fragestellungen aus diesem Bereich kommt die Weiterentwicklung der Mehrebenenanalyse zugute, da sie eine methodisch adäquate Abschätzung von Kontexteffekten erlaubt.

Die vorliegende Arbeit beschäftigt sich daher mit dem Thema Notengebung. Anhand einer aktuellen Schülerstichprobe sollen Korrelate von Schulnoten identifiziert werden. Deren Zusammenhänge werden anschließend auf gruppen- bzw. kontextspezifische Variabilität geprüft. Die Einbettung des Themas wird aus verschiedenen Perspektiven gesehen: *Kapitel 2* dieser Arbeit beleuchtet die individuelle und gesellschaftliche Bedeutung der Schulnoten. Vor dem Hintergrund der historischen Entwicklung der Schülerbeurteilung und unter den aktuellen Rahmenbedingungen der Notengebung wird beschrieben, welche Funktionen Schulnoten haben und welche praktischen Probleme mit ihnen verbunden sind.

Kapitel 3 fasst dann die Befundlage zur messtheoretischen Güte von Schulnoten zusammen. Dargestellt werden Befunde zur mangelnden Objektivität, Reliabilität und Validität von Schulnoten. Gleichzeitig wird die Interpretation der mangelnden Güte von Schulnoten durch einen kritischen Exkurs zum Ver-

hältnis von Testleistung und Schulnote relativiert. Auch die nachfolgend dargestellten Ansätze aus der US-amerikanischen Forschungsrichtung zum *classroom assessment* legen nahe, dass dem Lehrerurteil in Form von Schulnoten eine gewisse Validität zugeschrieben werden darf.

Kapitel 4 trägt anschließend Befunde zu den Korrelaten von Schulnoten zusammen. In der ersten Hälfte des Kapitels werden dafür Studien zur Notenpraxis berichtet. In deren Zentrum stehen immer die Personen, die die Noten vergeben: die Lehrer. Unter Rückbezug auf das vorausgehende Kapitel, das Lehrern auf das Thema Notengebung eine eigene Sichtweise zuschreibt, werden in dieser Kapitelhälfte Einflussfaktoren der Notengebung gesammelt, die für Lehrer wichtig bzw. relevant sind. Die zweite Hälfte des Kapitels reichert dies dann durch Befunde aus Korrelationsstudien an. In *Kapitel 5* folgt abschließend eine Zusammenfassung des empirischen Hintergrunds und die Ableitung von Hypothesen.

In Kapitel 6 und 7 erfolgt eine Beschreibung der methodischen Herangehensweisen bzw. Grundlagen dieser Arbeit. *Kapitel 6* beginnt mit einer kurzen Beschreibung der TIMS-Studie, deren Deutschland-Daten die Datengrundlage der vorliegenden Arbeit bilden. Anschließend werden die für die Analysen verwendete Stichprobe sowie die genutzten Daten beschrieben. *Kapitel 7* geht dann sehr ausführlich auf den Umgang mit fehlenden Daten sowie die Grundlagen der Mehrebenenanalyse ein. Die ausführliche Darstellung wird bewusst gewählt, da es sich bei beiden Methoden um Methoden handelt, die noch nicht sehr weit verbreitet sind. Multiple Imputation und Mehrebenenanalyse sind zwar gängige Methoden im Bereich des *large-scale assessments*, in anderen Forschungsbereichen werden sie dagegen nur äußerst selten angewendet. Durch die ausführliche Beschreibung der methodischen Hintergründe soll die gewählte Vorgehensweise für einen möglichst breiten Leserkreis transparent werden.

Kapitel 8 stellt dann die Ergebnisse der zur Hypothesenprüfung durchgeführten Analysen vor. Den Abschluss der Arbeit bildet *Kapitel 9*. In diesem Kapitel werden die Befunde vor dem Hintergrund der formulierten Hypothesen und Fragstellungen zusammengefasst und in Hinsicht auf Implikationen für die Praxis diskutiert. Abschließend gibt das Kapitel einen Ausblick auf weitere Forschungsfragen, die sich an diese Arbeit anschließen könnten.

2 Zur individuellen und gesellschaftlichen Bedeutung von Schulnoten

Das folgende Kapitel beschreibt die Bedeutung von Noten für das Individuum und die Gesellschaft. Dafür geht es zuerst auf die historische Entwicklung der Schülerbeurteilung ein (Unterkapitel 2.1) und dann auf die aktuellen Rahmenbedingungen der Notengebung (Unterkapitel 2.2).

2.1 Historische Entwicklung der Schülerbeurteilung

Die Beschäftigung mit dem Thema der Schülerbeurteilung – und somit auch mit dem Thema der Notengebung – kommt nicht ohne eine Betrachtung ihrer historischen Entwicklung aus. Viele der damit verbundenen Aspekte, wie z.B. die Funktionen, Verfahren und auch die Probleme der Schülerbeurteilung, lassen sich nur aus einer historischen Perspektive verstehen (Schaub, 1983). Einen umfassenden Überblick über die Entwicklung der Zensur und der Schülerbeurteilung in Deutschland bietet Dohse (1967). Unter Auswertung einer Vielzahl historischer Quellen beschreibt er die geschichtliche Entstehung des Zeugnisses und der Zensur über gut 500 Jahre. In seinen Ausführungen wird deutlich, dass sich die Zensur als wesentlicher Bestandteil des Schulzeugnisses entwickelte. Aufgrund der engen parallelen Entwicklung von Schulnoten und Zeugnissen wird nachfolgend nicht nur die Entstehung der Schulnoten (Abschnitt 2.1.2), sondern auch die der Zeugnisse (Abschnitt 2.1.1) beschrieben.

2.1.1 Zeugnisse

Zeugnisse sind keine „naturegebenen Bestandteile der Schule“ wie Breitschuh (1979, S. 36) festhält, sondern haben sich unter bestimmten historischen Bedingungen entwickelt. Verschiedene Autoren führen erste Belege für die Existenz von Zeugnissen und Zensurengebung aus dem 16. Jahrhundert an (Kleinert, 1951; Dohse, 1967; Breitschuh, 1979).

Benefizienzzeugnisse markieren für Dohse (1967) den Beginn der Zeugnis-schreibung. Obwohl in ihnen auch Leistungsbewertungen vorgenommen werden,

sind sie doch in erster Linie Sittenzeugnisse, die den Fleiß und die Führung eines Schülers beurteilen (Dohse, 1967, S. 13). Insofern haben sie den Charakter eines Empfehlungsschreibens: Sie wenden sich mit betont positiven Formulierungen an unbekannte Dritte „zum Zweck der Erlangung eines Beneficiums“ (Dohse, 1967, S. 13). Benefizienzzeugnisse werden auf Bitten ausgestellt und dienen dazu, Schülern mittelloser Eltern die Bewerbung um Stipendien oder ähnliches zu ermöglichen. Und obwohl diese Form des Zeugnisses kein Pflichtdokument ist, hat es dennoch eine hohe Verbindlichkeit. Es ist eng an vorausgehende Prüfungen gebunden und erhält durch die schriftliche Fixierung sowie durch Unterschrift und Siegel einen amtlichen Charakter (Dohse, 1967).

Das Benefizienzzeugnis ist die Grundlage für weitere Zeugnisformen, die aus ihm hervorgehen und sich parallel mit ihm weiterentwickeln. Diese sind das Reifezeugnis, das Abgangszeugnis und das periodische Zeugnis.

Das *Reifezeugnis* soll die aufgrund einer Prüfung festgestellte erforderliche geistige Reife für den Besuch einer Hochschule belegen und dadurch weitere Prüfungen für die Aufnahme an einer Hochschule ersparen (Dohse, 1967). Im Vergleich zum Benefizienzzeugnis hat es eher den Anspruch der Allgemeingültigkeit als den der Sozialfürsorge (Ziegenspeck, 1999). Lange Zeit bleibt jedoch auch das Reifezeugnis unverbindlich. Dies ändert sich mit dem im 18. Jahrhundert zunehmenden Interesse des Staates an qualifizierten Hochschulabgängern (z.B. zur Deckung des Bedarfs an Beamten und der zukünftigen Führungsschicht) und dem daraus folgenden Bedarf der Auslese. Zudem entwickelt die Schule verschiedene Eigeninteressen: „Hebung ihres Ansehens“, „Stärkung des Schulbesuchs“ und „Wiederaufrichtung der ‚zerfallenen Schulzucht‘“ (Dohse, 1967, S. 15). Aber selbst mit der gesetzlichen Einführung der Reifeprüfung erlangt das darauf basierende Reifezeugnis noch keine allgemeine Bedeutung, da die Scheu vor dem Eingriff in die „bürgerliche Freiheit“ und die damit verbundene Bedeutung von „Lernfreiheit und Elternrecht“ noch zu groß sind (Dohse, 1967, S. 16). Nur langsam setzt sich dann doch das Leistungsprinzip als Ablösung des Abstammungsprinzips der Ständegesellschaft durch und wird maßgebend für den sozialen Aufstieg (Dohse, 1967). Erst in der Mitte des 19. Jahrhunderts erhält das Reifezeugnis den allgemeinen Berechtigungscharakter, den es heute noch besitzt (Ziegenspeck, 1999). Dabei wandelt sich der Zweck des Reifezeugnisses von einem Ausweis allgemeiner Reife für das Studium der Wissenschaft hin zu einem differenzierten Ausweis „der erfolgreichen Aneignung des durch Lehrpläne der verschiedenartigen Anstalten erforderlichen Wissens- und Bildungstoffes“, „wobei gleichzeitig die sittliche Reife des abgehenden Schülers für die nun anhebende freiere Studien- und Lebensführung durch die bisherigen Lehrer zu bescheinigen ist“ (Roloff, 1913, Sp. 17, zit. nach Dohse, 1967, S. 17)

In die Zeit der Entwicklung des Reifezeugnisses fallen auch kritische Stimmen zur Prüfbarkeit von Bildung und Reife. Ziegenspeck (1999) fasst die von

Schwartz (1910) dokumentierte Diskussion des preußischen Oberschulkollegiums zum Problem einer Prüfung der allgemeinen Bildung für die Zulassung zum Universitätsstudium zusammen. Dabei zeigen sich bei der über 300 Jahre alten Diskussion erstaunliche Parallelen zu aktuellen Diskussionen: Schon dort heißt es, dass Prüfungen die Gefahr mit sich bringen „echte Bildungsarbeit der Schule zu ersticken“ und Lernen auf das „Einpauken von Prüfungswissen“ zu beschränken (Ziegenspeck 1999, S. 69). Echte Bildung sei nicht prüfbar. Abgesehen davon seien nicht reine Kenntnisse, sondern deren sinnvolle Anwendung relevant. Weiterhin wird die Notwendigkeit der Auslese der Freiheit zur Bildung gegenübergestellt und das Erfordernis landeseinheitlicher Beurteilungsmaßstäbe diskutiert.

Als eine weitere Grundform des Zeugnisses entwickelt sich gemeinsam mit dem Benefizien- und dem Reifezeugnis das *Abgangszeugnis*. Wie das Benefizien- und das Reifezeugnis entsteht es zunächst in der höheren Schule und wird bei einem Schulwechsel bzw. beim Schulabbruch ausgestellt. Erst bei Einführung der allgemeinen Schulpflicht werden auch in der niederen Schule Abgangszeugnisse ausgestellt. Sie dienen dann als Bescheinigung über die Erfüllung der Schulpflicht. Die dadurch erfüllte Kontrollfunktion schafft gleichzeitig einen Anreiz für den Schulbesuch, denn das Abgangszeugnis wird vielfach beim Berufsbeginn eingefordert (Dohse, 1967). Das Abgangszeugnis hat schon von Beginn an im Vergleich zum *ordnungsgemäßen Abschlusszeugnis* eine geringere Bedeutung (Dohse, 1967). So dokumentiert das Abschlusszeugnis das ordnungsgemäße Durchlaufen sämtlicher Volksschulklassen und soll damit eine breite Grundbildung entsprechend dem Fächerkanon der Volksschule nachweisen. Das Abgangszeugnis hingegen bescheinigt nur die Erfüllung der allgemeinen Schulpflicht, ohne dass alle Schulklassen durchlaufen wurden (Ziegenspeck, 1999).

Neben den bisher beschriebenen Zeugnissen, die alle beim Abgang von einer Schule ausgestellt wurden, existieren auch *periodische Zeugnisse*. Sie werden in regelmäßigen Abständen mehrmals im Laufe eines Schuljahres erteilt. Wie für die anderen Zeugnisse finden sich auch für die periodischen Zeugnisse erste Belege aus dem 16. Jahrhundert. Insbesondere bei den Jesuiten scheint es periodische Zeugnisse gegeben zu haben (Dohse, 1967). Dem damaligen Menschenbild des Ordens liegt die Vorstellung von einem intrinsisch motivierten Menschen fern. „Genaueste Verplanung der Arbeitszeit, fortwährende Repetitionen und regelmäßige Examina [sind] daher wesentliche Kennzeichen des jesuitischen Schulbetriebs“ (Breitschuh, 1979, S. 37). Demgemäß verwenden die Jesuiten das periodische Zeugnis, um Wettbewerb zu schaffen und den Ehrgeiz anzuregen. Dadurch hat das Zeugnis weniger die Funktion eines nach außen wirkenden Qualifikationsnachweises, sondern dient mehr schulinternen pädagogischen Zielen (Dohse, 1967; Breitschuh, 1979). Periodische Zeugnisse finden sich aber auch andernorts und werden dann als Grundlage für die Ausstellung eines Benefizienzeugnisses verwendet. Auch in der Elementarschule werden periodische Zeugnisse erteilt. Im

Gegensatz zur Anreizfunktion in der höheren Schule dienen sie mit Einführung der allgemeinen Schulpflicht eher der Kontrolle des Schulbesuchs (Dohse, 1967).

Zusammenfassend ist festzuhalten, dass Zeugnisse keine genuinen Bestandteile des Schulwesens sind, sondern sich als Folge des Berechtigungswesens entwickelt haben. Damit sollen sie in erste Linie Schulleistung ausweisen. In viele Zeugnisformen wie z.B. in das Benefizienzeugnis oder das Reifezeugnis fließt aber auch eine Bewertung von Sekundärtugenden ein.

2.1.2 Schulnoten

Bei der Frage nach dem Ursprung der Zensur stellt Dohse (1967) fest, dass sie schon in den ältesten Zeugnissen, den sogenannten Benefizienzeugnissen, ein wesentlicher Bestandteil ist. Als Beleg führt er ein Beispiel an, in dem zwar keine Ziffernzensuren verwendet werden, wohl aber verbale Kategorien zur Klassifikation von Leistung (Dohse, 1967, s. S. 45). Dies entspricht auch der ursprünglichen Wortbedeutung des Begriffs zensieren. Er leitet sich vom lateinischen Wort *consere* – beurteilen, prüfen, schätzen – ab und bezieht sich auf die Tätigkeiten des *Censors* im alten Rom (Sost, 1926; Kleinert, 1951; Dohse, 1967). Der *Censor*, ein für fünf Jahre gewählter Staatsbeamter, hatte während der Ausübung des Amtes der *consura* sowohl die Oberaufsicht in Polizeianglegenheiten als auch über das Vermögen und die Sitten der Bürger. Er hatte die Aufgabe den Bürgern gemäß ihrem Vermögen, ihrer Geburt und ihrer sittlichen Führung einen Stand und bestimmte bürgerliche Rechte zuzuweisen (Dohse, 1967). Später dann wurde der Begriff *consere* allgemein für das Begutachten und Beurteilen eines Menschen und seiner Handlungsweisen verwendet. „Daraus entstand dann das ‚Zensieren‘, also die Handlungsweise, die schulisches Verhalten und Schulleistungen feststellt und einordnet“ (Ziegenspeck, 1999, S. 65).

Dohse (1967) unterscheidet drei Grundformen der Zeugniszensur: die mündliche Zensur, die symbolische Zensur und die sich daraus entwickelnde Leistungseinstufung durch die Zahl. Er führt einen Beleg für die *mündliche Zensur* aus dem 18. Jahrhundert an und beschreibt sie als Instrument, das nicht als Rückmeldung für den Schüler, im Sinne von Lob oder Tadel, gedacht ist, sondern das sich hauptsächlich an Dritte richtet. Für die *symbolische Zensur* dagegen finden sich noch frühere Belege. Im Vergleich zu Zensurenkarten, Fleißbillets, Meritentafeln, Orden etc., die allesamt nur sehr kurzlebig waren, ist Lokation als die älteste Form der symbolischen Zensur zu nennen. Bei ihrer Anwendung wird die Sitzordnung der Schüler nach ihrem Leistungsstand bestimmt. Kleinert (1951) erläutert, dass solche Rangfolgen im Allgemeinen auf Basis der Gedächtnisleistungen der Schüler erstellt wurden, da „die möglichst getreue Wiedergabe des gelernten Stoffes als die

beste Leistung“ gilt (S. 919). Weiterhin bewertet er diese Rangreihe als Vorläufer der Zahlwertung einer Schulleistung.

Bei den genannten Grundformen der Zensur finden sich verschiedene Stufungsformen, die Dohse (1967) als Reihenzensur, polare Zensur und Zonenzensur bezeichnet. Die *Reihenzensur* kann sowohl eine normorientierte als auch eine ipsative Bewertung darstellen. Im Fall der Lokation beispielsweise werden Ordnungszahlen mit absteigendem Wert vergeben, d.h. die Beurteilung erfolgt in Relation zur Schülergruppe und es werden so viele Zahlen vergeben, wie Schüler vorhanden sind. Bei den *Meritentafeln* dagegen werden zwar ebenfalls Zahlen vergeben, ihre Höhe orientiert sich aber an einem früheren Leistungsstand des Schülers.

Im Gegensatz zur Reihenzensur, bei der lediglich eine Reihenfolge hergestellt wird, erfolgt bei der *polaren Zensur* eine „Einstufung von Leistungen und Verhaltensweisen nach ihrem positiven oder negativen Wertcharakter“ (Dohse, 1967, S. 47f). Infolgedessen treten Extremgruppen von Schülern hervor, was insbesondere dort genutzt wird, wo das Ehr- und Schamgefühl der Schüler angesprochen werden soll, wie z.B. bei den Jesuiten (s. auch Abschnitt 2.1.1). Zwischen den beiden Extremgruppen liegt allerdings keine scharfe Trennlinie, sondern eher ein breites Mittelfeld von Schülern. Dies begründet die Entwicklung der Zonenzensur.

Die *Zonenzensur* ist eine Klassifikation der Schüler in annähernd gleich große Gruppen. In den Anfängen bedeutet dies zunächst eine Dreiteilung der Schülergruppe. Damit scheint die Zonenzensur eine Mischform aus der normorientierten Reihenzensur und der polaren Zensur zu sein. Im Vergleich zur polaren Zensur und der mit ihr beabsichtigten Motivierung steht bei der Zonenzensur die Auslese- und Grenzfunktion im Vordergrund (Dohse, 1967). Die schriftliche Zonenzensur ist daher vermutlich die Urzensur, die zur Zeugniszensur und somit zu einem wesentlichen Bestandteil des Schulzeugnisses wird. Grundlage für die Festlegung der Zeugniszensur sind Prüfungszensuren sowie einzelne, während des Unterrichts vergebene Zensuren. Diese werden oftmals in sogenannten Klassenkatalogen, den Vorläufern des Klassenbuchs, festgehalten. Um hier eine gewisse Übersichtlichkeit und Vergleichbarkeit zu wahren, sind kurze aber aussagekräftige Bewertungen notwendig. Als kürzeste Formulierung bietet sich daher die Zifferzensur an (Dohse, 1967).

Mit der Entwicklung der *Zifferzensur* variiert auch die Zahl ihrer Stufen und deren verbaler Beschreibung. Frühe Zensurenskalen haben meist nur wenige Stufen, die aber mit der Zeit immer weiter ausdifferenziert werden. Eine sechsstufige Notenskala, die für alle Schulen verbindlich ist, wird erstmals 1938 eingeführt (Dohse, 1967). Nach dem Zweiten Weltkrieg ist im Jahr 1954 ein Beschluss der KMK notwendig, um angesichts des Kulturföderalismus wieder zu einer einheitlichen Notenskala zu gelangen. Gemäß dem Beschluss ist bei der Leistungsbewer-

tung der Leistungsdurchschnitt der Klassenstufe als Referenz heranzuziehen (Ziegenspeck, 1999). Ziegenspeck (1999) kritisiert, dass dabei nicht genau definiert wird, was unter dem Begriff der Durchschnittsleistung zu verstehen ist und auf welchen Rahmen sich dieser Begriff beziehen soll. Ein späterer Beschluss der KMK (1968) räumt dies aus. Er verfügt, dass die Leistungsbewertung sich nach den Anforderungen des Unterrichts richten soll (s. dazu auch Abschnitt 2.2.1).

Insgesamt ist festzuhalten, dass Noten sich als Beurteilungsinstrument entwickelt haben. Wie bei den Zeugnissen steht dabei die Beurteilung von Schulleistungen weitgehend im Vordergrund. Daneben ist aber auch eine Verhaltensbeurteilung Bestandteil verschiedener Notenformen.

2.1.3 Fazit

Die Übersicht über die geschichtliche Entwicklung der Notengebung zeigt, dass Schulnoten und Schulzeugnisse keine originären Bestandteile des Schulwesens sind. Sie entwickeln sich zunächst in Form von Benefizienzzeugnissen bzw. verbalen Kategorisierungen. Diese frühen, noch unverbindlichen Formen der Schülerbeurteilung haben einen starken sozialfürsorglichen Charakter. Dementsprechend stehen im Zentrum der Beurteilung Sekundärtugenden. Auch die Zeugnis- und Benotungspraktiken des Jesuitenordens zielen auf diese Sekundärtugenden ab. Hier richtet sich die Schülerbeurteilung aber nicht an Dritte, sondern an die Schüler: Mithilfe von periodischen Zeugnissen und polaren Zensuren soll deren Ehrgeiz angefacht werden.

Die verbindliche Einführung von Noten und Zeugnissen ist eng mit einer zunehmenden Institutionalisierung und Bürokratisierung des Schul- und Berufswesens verbunden. Zu nennen sind in diesem Zusammenhang die Bedeutung von Zeugnissen für den Übergang zur Hochschule und ins Berufsleben (Dohse, 1967) sowie die Einführung des Jahrgangsklassensystems mit der jahrgangsweisen Einschulung, der Versetzung, dem verbindlichen Fächerkanon usw. (Ingenkamp, 1995, S. 49). So wird beispielsweise die Ausstellung des Reifezeugnisses verpflichtend, als der Staat ein verstärktes Interesse an qualifizierten Hochschulabgängern entwickelt. Als Ursache dafür sieht Ingenkamp (1995) die Tatsache an, dass der benötigte Nachwuchs nicht mehr aus einer standesgemäßen Klasse gedeckt werden kann und andere Gesellschaftsklassen an Einfluss gewinnen. Somit treten Zeugnis und Schulnote in den Dienst des Berechtigungswesens und mit ihnen löst das Leistungsprinzip das Ständeprinzip ab.

2.2 Aktuelle Rahmenbedingungen der Notengebung

Nach dem Blick in die historische Entwicklung der Schülerbeurteilung folgt nun eine Betrachtung der aktuellen Rahmenbedingungen der Notengebung. Dazu zählt in erster Linie die geltende Gesetzeslage zur Notengebung. Abschnitt 2.2.1 stellt einen für alle Bundesländer verbindlichen Beschluss der Kultusministerkonferenz dar und geht exemplarisch auf verschiedene Ländergesetze und -verordnungen ein. Die darauf folgenden Abschnitte berücksichtigen, dass neben der Legislative, die den formalen Rahmen der Notengebung ausmacht, auch noch ein praxisbezogener Rahmen besteht. So werden der Notengebung in der Literatur viele verschiedene und zum Teil gegensätzliche Funktionen zugedacht (Abschnitt 2.2.2). Auch sind Schulnoten mit unterschiedlichen individuellen Konsequenzen verbunden (Abschnitt 2.2.3).

2.2.1 Aktuelle Gesetzeslage

Der bereits in Abschnitt 2.1.2 genannte Beschluss der KMK aus dem Jahr 1968 regelt auch heute noch die rechtsverbindliche Definition der Schulnoten. Bis zum Jahr 1968 sind die Schulnoten mit Bezug auf die Durchschnittsleistung definiert, wobei eine genauere Beschreibung dessen, was unter der Durchschnittsleistung zu verstehen ist, fehlt (Ziegenspeck, 1999). Mit dem Beschluss aus dem Jahr 1968 wird der Bezug zur Durchschnittsleistung fallen gelassen und durch den Bezug zu Anforderungen ersetzt. Laut Beschluss bezieht sich der Begriff der Anforderungen „auf den Umfang sowie auf die selbständige und richtige Anwendung der Kenntnis und auf die Art der Darstellung“ (KMK, 1968, S. 1). Gleichzeitig wird darauf hingewiesen, dass diese Anforderungen je nach Schulgattung, Fach und Alter der Schüler variieren können (vgl. Abbildung 2.1).

Diese Beschreibung der Schulnoten, die zwar als Verbesserung gegenüber der durchschnittsbezogenen Definition bewertet wird, wird dennoch auch kritisiert. Ziegenspeck (1999) kann in keinem Bundesland amtliche Bestimmungen ausmachen, die den Begriff der Anforderungen durch genaue Lernziele konkretisieren. Seiner Einschätzung nach sind Lehrer aufgrund des fehlenden Bezugspunkts nach wie vor auf die Orientierung am klassen- oder schulinternen Bezugsrahmen angewiesen.

Abbildung 2.1: Erläuterung der Notenstufen bei Schulzeugnissen und Einzelergebnissen in staatlichen Prüfungszeugnissen (KMK, 1968, S. 1)

Den Noten werden die folgenden Definitionen zugrunde gelegt:

1. sehr gut (1)
Die Note „sehr gut“ soll erteilt werden, wenn die Leistung den Anforderungen in besonderem Maße entspricht.
2. gut (2)
Die Note „gut“ soll erteilt werden, wenn die Leistung den Anforderungen voll entspricht.
3. befriedigend (3)
Die Note „befriedigend“ soll erteilt werden, wenn die Leistung im allgemeinen den Anforderungen entspricht.
4. ausreichend (4)
Die Note „ausreichend“ soll erteilt werden, wenn die Leistung zwar Mängel aufweist, aber im ganzen den Anforderungen noch entspricht.
5. mangelhaft (5)
Die Note „mangelhaft“ soll erteilt werden, wenn die Leistung den Anforderungen nicht entspricht, jedoch erkennen läßt, daß die notwendigen Grundkenntnisse vorhanden sind und die Mängel in absehbarer Zeit behoben werden können.
6. ungenügend (6)
Die Note „ungenügend“ soll erteilt werden, wenn die Leistung den Anforderungen nicht entspricht und selbst die Grundkenntnisse so lückenhaft sind, daß die Mängel in absehbarer Zeit nicht behoben werden können.

Bei der Bewertung von Schülerleistungen ist der Eigenart der Schulgattung und des Faches sowie dem Alter des Schülers Rechnung zu tragen. Der Begriff „Anforderungen“ in den Definitionen bezieht sich auf den Umfang sowie auf die selbständige und richtige Anwendung der Kenntnis und auf die Art der Darstellung.

Auch für die aktuelle Gesetzeslage scheint diese Einschätzung immer noch gültig zu sein. Eine genauere Beschreibung der Notenstufen fehlt. Meist greift eine bundeslandspezifische Verordnung oder ein Gesetz die Definition der KMK (1968) wörtlich auf (z.B. Berlin: § 58 SchulG², Baden-Württemberg: § 5 NVO³). Hinzu kommen oft auch verbindliche Notendefinitionen für das Sozial- sowie

2 *Schulgesetz für das Land Berlin*. Vom 26. Januar 2004 (GVBl. S. 26), das zuletzt durch Gesetz vom 25. Januar 2010 (GVBl. S. 14) sowie durch Artikel 1 des Gesetzes vom 25. Januar 2010 (GVBl. S. 22) geändert worden ist. Verfügbar unter: http://www.berlin.de/imperia/md/content/sen-bildung/rechtsvorschriften/schulgesetz_25012010_ueberarb.pdf?start&ts=1266916205&file=schulgesetz_25012010_ueberarb.pdf [06.03.2010].

3 *Verordnung des Kultusministeriums über die Notenbildung* [Notenbildungsverordnung]. Vom 5. Mai 1983. Letzte berücksichtigte Änderung: §§ 3 und 9 geändert durch Artikel 9 Verordnung vom 11. November 2009 (GBl. S. 693, 709). Verfügbar unter: <http://www.landesrecht-bw.de/jportal/?quelle=jlink&query=NotBildV+BW&psml=bsbawueprod.psml&max=true&aiz=true> [12.08.2011].

das Lern- und Arbeitsverhalten, die sogenannten „Kopfnoten“. Das Bundesland Bayern sieht dafür beispielsweise eine vierstufige Skala mit den Attributen „sehr gut“, „gut“, „befriedigend“ und „nicht befriedigend“ vor und regelt, dass die Vergabe dieser Noten erläutert werden muss (§ 50 VSO⁴). Auch in Baden-Württemberg gibt es eine vierstufige Notenskala zur Bewertung des Arbeits- und Sozialverhaltens. Im Gegensatz zur bayrischen Ausgabe enthält die entsprechende Verordnung aber eine Kurzbeschreibung der vier Notestufen (§ 6 NVO). Je nach Bundesland kann die Stufung dieser Noten variieren. Auch die Merkmale, auf die sie sich beziehen, werden unterschiedlich benannt (Bohl, 2003).

Während explizite Regelungen der Notengebung in der Ländergesetzgebung selten sind und kaum über die Definition der KMK (1968) hinausgehen, lassen sich aber aus manchen Gesetzen Regelungen zur Leistungsbeurteilung sowie aus Zeugnisregelungen notenrelevante Informationen entnehmen. In Sachsen-Anhalt beispielsweise ist die Leistungsbewertung in der Grundschule für verschiedene Bewertungsbereiche sehr ausführlich durch einen Erlass⁵ geregelt. So wird für Klassenarbeiten eine Zuordnung der prozentual erreichten Leistung zu einer Note vorgenommen und es werden Notensenkungen bei starken Abweichungen vom normgerechten Schreiben verfügt. Auch die Erstellung der Zeugnisnoten wird geregelt. So sollen für Halbjahresnoten alle vom Schüler erbrachten Leistungsnachweise herangezogen werden, wobei jede Klassenarbeit zu 1/5 in die Halbjahresnote eingeht. Die Schuljahresnote dagegen soll die Zusammenfassung zweier Halbjahresnoten sein, bei der zusätzlich die Notentendenz sowie die Leistungsentwicklung und das Lernverhalten des Schülers im Verlauf des Schuljahres berücksichtigt werden.

Im Gegensatz zu solch genauen Vorgaben über die Leistungsbeurteilung enthalten aber auch viele Länderregelungen den Hinweis auf die pädagogische Freiheit bzw. Verantwortung des Lehrers (z.B. Nordrhein-Westfalen: § 57 SchulG⁶, Sachsen-Anhalt: § 30 SchulG⁷). Auch Rheinland-Pfalz betont die be-

4 *Schulordnung für die Grund- und Hauptschulen (Volksschulen) in Bayern* [Volksschulordnung]. Vom 11. September 2008 (GVBl. S. 684, BayRS 2232-2-UK), zuletzt geändert durch Verordnung vom 31. März 2010 (GVBl. S. 185). Verfügbar unter: http://by.juris.de/by/gesamt/VoSchulo_BY_2008.htm [23.05.2010].

5 *Leistungsbewertung in der Grundschule*. RdErl. des MK vom 2. Juli 2003-31-83200 (SVBl. LSA S. 205) einschließlich 1. Änderung RdErl. des MK vom 30. Juni 2004-31-83200 (SVBl. LSA S. 156) sowie Wieder-Inkraftsetzung durch RdErl. des MK vom 12. Oktober 2005-31-83200 (SVBl. LSA S. 389), 2. Änderung RdErl. des MK vom 30. Juni 2008. Verfügbar unter: <http://www.mk-intern.bildung-lsa.de/Bildung/er-leistungsbewertunggrundschulen.pdf> [12.08.2011].

6 *Schulgesetz für das Land Nordrhein-Westfalen*. Vom 15. Februar 2005 (GV. NRW. S. 102) zuletzt geändert durch Gesetz vom 17. Dezember 2009 (GV. NRW. S. 863). Verfügbar unter: http://www.schulministerium.nrw.de/BP/Schulrecht/Gesetze/SchulG_Info/Schulgesetz.pdf [11.03.2010].

sondere pädagogische Verantwortung der Lehrer bei der Leistungsfeststellung und -beurteilung (§ 25 SchulG⁸). Diese scheint sich in einem größeren Spielraum bei der Leistungsbeurteilung niederzuschlagen. So sollen „Form und Anzahl der Leistungsfeststellungen [...] von pädagogischen Gesichtspunkten bestimmt“ werden; außerdem darf die „Anzahl der Leistungsfeststellungen [...] bei einzelnen Schülerinnen und Schülern verschieden sein“ (§ 33 GrundschulO⁹). Daneben sollen bei der Leistungsbewertung nicht nur der Grad des Erreichens von Lernanforderungen beurteilt werden, sondern auch der individuelle Lernfortschritt des Schülers, seine Leistungsbereitschaft und die Lerngruppe, in der die Leistung erbracht wird (§ 34 GrundschulO). Hinzu kommt, dass die Zusammenfassung von Einzelleistungen zu Zeugnisnoten nicht notwendigerweise nach Durchschnittsbildung erfolgen muss, sofern dies „durch den Schwierigkeitsgrad oder den Umfang der überprüften Einzelleistung begründet ist“ (§ 42 GrundschulO). Trotz ihrer pädagogischen Freiheit sind die Lehrer aber dem Schulleiter Rechenschaft schuldig. Dieser kann beispielsweise in Abstimmung mit der Fachkonferenz Notenurteile des Lehrers umgehen oder die Wiederholung einer Klassenarbeit verfügen, wenn mehr als ein Drittel der Schüler die Note ausreichend erhalten hat (§ 34 GrundschulO).

Grundsätzlich ist festzuhalten, dass die deutsche Gesetzgebung Schulnoten durch die Anforderungen des Unterrichts bestimmt sieht. Die Mehrheit der Bundesländer versteht unter der Erfüllung der Anforderungen dabei die Erfüllung bestimmter Lern- bzw. Leistungsanforderungen. Verhaltensbeurteilungen sind für gewöhnlich durch gesonderte Noten vorzunehmen. Allerdings sind viele Gesetzestexte und Verordnungen wenig konkret formuliert und betonen zum Teil sogar die pädagogische Verantwortung des Lehrers, sodass Lehrkräfte einen großen Spielraum bei der Notenvergabe haben.

-
- 7 *Schulgesetz des Landes Sachsen-Anhalt*. In der Fassung der Bekanntmachung vom 11. August 2005. Letzte berücksichtigte Änderung: § 71 geändert durch § 30 Absatz 3 des Gesetzes vom 16. Dezember 2009 (GVBl. LSA S. 684, 689). Verfügbar unter: <http://www.landesrecht.sachsen-anhalt.de/jportal/?quelle=jlink&query=SchulG+ST&pml=bssahprod.psm1&max=true&aiz=true> [11.03.2010].
 - 8 *Schulgesetz*. Vom 30. März 2004, zuletzt geändert durch Gesetz vom 22. Dezember 2009 (GVBl. S. 418). Verfügbar unter: http://rlp.juris.de/rlp/gesamt/SchulG_RP_2004.htm [07.03.2010].
 - 9 *Schulordnung für öffentliche Grundschulen*. Vom 10. Oktober 2008. Verfügbar unter: http://grundschule.bildung-rp.de/fileadmin/user_upload/grundschule.bildung-rp.de/Downloads/Amtliches/Neue_Grundschulordnung_08/Schulordnung_fuer_die_oeffentlichen_Grundschulen_vom_10_10_2008.pdf [12.08.2011].

2.2.2 Funktionen von Schulnoten

In der Literatur existieren viele Publikationen, die verschiedene Funktionen von Schulnoten auflisten und klassifizieren (z.B. Zielinski, 1974). Diese vielfältigen Funktionen der Schulnoten – oder Schülerbeurteilung im Allgemeinen – lassen sich grob zwei Oberbegriffen zuordnen. Je nach Autor werden sie unterschiedlich betitelt. So sprechen Tillmann und Vollstädt (2000) oder Tent (2006) von pädagogischen und gesellschaftlichen Funktionen der Schulnoten. Shepard (2000a) dagegen teilt Schülerbeurteilungen in solche mit summativer und solche mit formativer Funktion ein.

2.2.2.1 Gesellschaftliche Funktionen

Zu den gesellschaftlichen Funktionen zählen alle Funktionen, bei denen die Schülerbeurteilung die Grundlage für gesellschaftsrelevante Entscheidungen ist. Die vordringlichste gesellschaftliche Funktion der Schulnoten ist die *Selektions- oder Allokationsfunktion*. Schulnoten berechtigen oder verweigern den Aufstieg in die nächst höhere Klassenstufe, den Besuch einer bestimmten Schulform des Sekundarschulwesens und den Zugang zur Hochschule bzw. zu einem bestimmten Berufszweig sowie zu gesellschaftlich anerkannten Positionen (Weiss, 1989; Sacher, 1994). In diesem Zusammenhang wird Schulnoten eine Prognosefunktion zugeschrieben. Unter der Prämisse, dass Schulleistungen konstant sind, gelten sie als Maß, das die Eignung eines Schülers für zukünftige Anforderungen beschreibt (Rieder, 1990). Hintergrund der Prognosefunktion ist die Idee der leistungsgerechten Auslese: der Auslese nach dem Leistungsprinzip. Demnach sollen Schulnoten ermöglichen, die vom Schulsystem „geforderten Zuweisungen und Laufbahntscheidungen nach einem als ‚gerecht‘ geltenden Kriterium vorzunehmen“ (Tillmann & Vollstädt, 2000, S. 38).

Das Leistungsprinzip steht auch mit der *Sozialisations- bzw. Qualifizierungsfunktion* von Schulnoten in Verbindung. Leistungsbewertungen sollen dazu beitragen bei den Schülern „das Kriterium der individuellen Leistung als gerechten Maßstab zur Verteilung unterschiedlicher Lebenschancen zu internalisieren“ (Tillmann & Vollstädt, 2000, S. 38). Auf diese Weise sollen die Leistungsorientierung unserer Gesellschaft vermittelt und Akzeptanz für Leistungszertifikate geschaffen werden (Sacher, 1994), um so den Erhalt der bestehenden gesellschaftlichen Ordnung zu sichern. Gleichzeitig soll die Identifikation mit diesen Zielen auch einen Anreiz bieten, eigene gesellschaftliche Ziele anzustreben und die dafür erforderlichen Leistungen zu erbringen bzw. geeignete Qualifizierungsmaßnahmen zu ergreifen (Sacher, 1994; Tillmann & Vollstädt, 2000).

Im weitesten Sinne zählt zu den gesellschaftlichen Funktionen auch die *Berichtsfunktion*. Während der deutsche Sprachraum damit meist eine Rückmeldung an die Eltern und Schüler meint und sie den pädagogischen Funktionen zuordnet (Tillmann & Vollstädt, 2000; aber s. auch Rieder, 1990), versteht die englische Literatur darunter eine Rechenschaftslegung gegenüber der Öffentlichkeit (Resnick & Resnick, 1996), die der Qualitätssicherung dient. Unter dem Begriff *accountability* kennzeichnet sie ein Rechenschaftsprinzip, demgemäß alle an der schulischen Bildung und Erziehung beteiligten Personen der Öffentlichkeit für ihre Handlungen Rechenschaft schuldig sind. Im Vergleich dazu zielt der ebenso gebräuchliche Begriff *monitoring* stärker auf eine Fortschrittskontrolle ab (Wheeler & Haertel, 1993). Beide Begriffe beziehen sich allerdings eher auf die Leistungsbeurteilung im Allgemeinen und weniger auf Schulnoten.

2.2.2.2 Pädagogische Funktionen

Zu den pädagogischen Funktionen zählen Funktionen, bei denen die Schülerbeurteilung die Grundlage für Entscheidungen bietet, mit denen dem Lernenden die Möglichkeit zum Lernfortschritt und zur Optimierung des Lehr-Lern-Prozesses geboten wird. Ziel jeder formativen Art von Evaluation ist in erster Linie eine Prozesssteuerung. Dies spiegelt sich auch in der *Rückmelde- bzw. Berichtsfunktion* von Schulnoten wider. Die Rückmeldefunktion der Schulnoten lässt sich am besten anhand der Personen beschreiben, denen die Rückmeldung dient. Üblicherweise werden drei Personengruppen unterschieden (Weiss, 1989; Ziegenspeck, 1999). Die Rückmeldung an die Schüler ist der zentrale Aspekt der Rückmeldefunktion. Aufgrund kontinuierlicher Rückmeldungen sollen sie allmählich dazu befähigt werden, das eigene Lernen richtig einzuschätzen und zu steuern (Tillmann & Vollstädt, 2000). Neben der Rückmeldefunktion für die Schüler haben die Schulnoten auch eine Berichtsfunktion gegenüber den Eltern. Noten sollen hier den Lernstand bzw. -fortschritt dokumentieren, damit die Eltern den Lernprozess unterstützend begleiten können (Tillmann & Vollstädt, 2000). Außerdem bilden Noten die Informationsgrundlage für Beratungsgespräche zwischen Schülern, Eltern und Lehrern (Ziegenspeck, 1999; Rieder, 1990). Nicht zuletzt bieten Schulnoten auch für die Lehrer eine Rückmeldung. Letztlich bringen sie nicht nur Informationen über den Lernprozess, sondern eben auch über den Lehrprozess hervor und geben somit dem Lehrer die Gelegenheit seinen Unterricht zu modifizieren (Ziegenspeck, 1999; Tillmann & Vollstädt, 2000).

Neben der bisher geschilderten Rückmeldefunktion von Schulnoten, die die kognitive Wirkung der Rückmeldung betont, wird der Leistungsrückmeldung in Form von Schulnoten auch eine motivierende Funktion zugeschrieben. Durch Schulnoten soll der Schüler zu weiterem Lernen motiviert werden – auch wenn

er nur ein geringes Interesse an den Lehrzielen hat (Rieder, 1990). Dabei wird dieser Motivierungsaspekt je nach Autor als *Anreizfunktion* oder als *Disziplinierungsfunktion* aufgefasst. Die Anreizfunktion von Schulnoten ist gegeben, sobald Schulnoten die extrinsische Motivation ansprechen. Nach der Ansicht von Weiss (1989) verlagert sich die ursprünglich intrinsische Lernmotivation von Kindern durch zunehmende Beschulung und Leistungsbeurteilung nahezu zwangsläufig zur extrinsischen Lernmotivation. Ziel ist dann nicht mehr das Lernen um des Lernens willen, sondern die Erlangung von Lob aufgrund guter Noten und die Vermeidung von Tadel aufgrund schlechter Noten. Die Disziplinierungsfunktion kann dabei als extreme Form der Anreizfunktion verstanden werden, durch die besonders leistungsunwillige Schüler bestraft werden sollen, mit dem Ziel sie zum erwünschten Leistungsverhalten zu veranlassen (Zielinski, 1974).

2.2.2.3 Die funktionale Überfrachtung der Schulpraxis

Der oben dargestellte Überblick über die der Schülerbeurteilung zugeordneten Funktionen verdeutlicht, dass diese sehr vielfältig sind. Diese Vielfalt geht so weit, dass Sacher (1994) von einer „funktionalen Überfrachtung der Schulpraxis“ spricht (S. 21) und anzweifelt, dass man angesichts einer solchen Fülle allen Funktionen in gleicher Weise gerecht werden kann. Zudem merkt er an, dass manche Funktionen sogar gegensätzliche Anforderungen an die Praxis stellen. Und auch aus solchen Divergenzen ergibt sich für ihn die Schlussfolgerung, dass man wenigstens einem Teil der genannten Funktionen nur sehr unzureichend nachkommen kann.

Letztlich ist auch die Perspektive der bei der Notengebung beteiligten Personen nicht zu vernachlässigen. Die oben zusammengestellten Funktionen der Notengebung haben keine empirische Grundlage, sondern sind intuitive und normative Funktionszuschreibungen, die von Lehrern, aber auch von Eltern nicht unbedingt geteilt werden müssen. Tatsächlich stehen für Lehrer wie für Eltern die pädagogischen Funktionen im Vordergrund: Aus der Sichtweise von Haupt- und Grundschullehrkräften sowie Lehramtskandidaten sind die Funktion der Rückmeldung an die Schüler aber auch an ihre Eltern sowie die Motivierungsfunktion am wichtigsten. Eine Disziplinierungsfunktion wird den Noten nur höchst selten zugeschrieben, während die Selektionsfunktion immerhin noch eine mittlere Wichtigkeit hat (Krampen, 1984). Diese Ansichten werden von Eltern weitgehend geteilt (Valtin & Schmude, 2002).

Die starke Favorisierung der pädagogischen Funktionen durch die Lehrer steht aber in einem Kontrast zur Betonung der gesellschaftlichen Funktionen in der Gesetzeslage zur Notengebung. Gleichzeitig lässt die Gesetzeslage dem Lehrer viele Freiräume bei der Schülerbeurteilung. Es ist daher zu vermuten, dass Lehrer diese nutzen, um manchen Funktionen der Benotung mehr und anderen weniger Gewicht zu verleihen.