

Peter Kistorz

Bewertungsmaßstäbe und Bezugsnormen bei der Notenvergabe unter der Lupe des Schulrechts – Was ist pädagogisch sinnvoll, was juristisch möglich?

1 Einleitung

Noten gehören zur Schule wie das Bier zum Grillabend – in beiden Fällen ist die Symbiose zwar kein unwiderrufliches Naturgesetz, gesellschaftlich aber derart stark im kollektiven Verständnis verankert, dass eine Trennung zumindest in absehbarer Zeit unmöglich erscheint.¹ Das war nicht immer so und nicht immer unumstritten: Ordinale Notenskalen bei der Leistungsbewertung von Schülerinnen und Schülern haben sich erst Ende des vorletzten Jahrhunderts in Preußen etabliert und wurden in den 1970er Jahren heftig diskutiert und von vielen zeitgenössischen Pädagoginnen und Pädagogen vehement in Frage gestellt.² Seit den 1980er Jahren ist diese Diskussion deutlich abgeebbt.³ Die Ziffernnoten von 1 (*sehr gut*) bis 6 (*ungenügend*) werden seitdem zwar als notwendig(es Übel) akzeptiert,⁴ aber nach wie vor in Bezug auf ihre teils unreflektierte Findung und Erteilung kritisiert. Diesbezüglich bedarf es unweigerlich einer Bezugsnorm – andernfalls ist eine ordinalskalierte Leistungsbewertung mit Ziffernnoten schlichtweg unmöglich.⁵ An welcher Bezugsnorm sich die Leistungsbewertung wiederum orientiert, ist abhängig von den Funktionen, die eine explizite Leistungsbewertung erfüllen soll.

Im Folgenden werden daher die gesellschaftlichen und pädagogischen Funktionen von Noten ebenso dargestellt (unter 2) wie die in Betracht kommenden Bezugsnormen bei der Zensurenfin-

1 Als Alternative zu den „klassischen“ Notenstufe von 1 (*sehr gut*) bis 6 (*ungenügend*), die in der Bundesrepublik bereits 1955 durch das sog. Düsseldorfer Abkommen (Abkommen zwischen den Ländern der Bundesrepublik zur Vereinheitlichung auf dem Gebiete des Schulwesens vom 17. Februar 1955) für alle Bundesländer verbindlich festgelegt worden sind, kämen beispielsweise verbale Beurteilungen in Betracht, die für gewöhnlich in den ersten beiden Schulstufen Anwendung finden: „So halten inzwischen alle Bundesländer die ersten beiden Schuljahre frei von Ziffernzeugnissen und nehmen stattdessen allgemeine Beurteilungen des Lernverhaltens der Schüler in Form eines verbalen Diagnoseberichts vor, mitunter ergänzt durch Ziffernzensuren mit verbalen Bemerkungen ‚angereichert‘“ (Jürgens, E./Sacher, W., *Leistungserziehung und Pädagogische Diagnostik in der Schule. Grundlagen und Anregungen für die Praxis*, Stuttgart 2008, S. 52).

2 Vgl. hierzu Winter, F., *Leistungsbewertung. Eine neue Lernkultur braucht einen anderen Umgang mit den Schülerleistungen*, Baltmannsweiler 6. Aufl. 2014, S. 39 ff. m. w. N. Ausgegangen war die Debatte in 1970er Jahren in erster Linie durch den von Karlheinz Ingenkamp 1971 herausgegebenen Band zur „Fragwürdigkeit der Zensurengebung“, in dem vor allem die mangelnde Vergleichbarkeit der erteilten Noten kritisiert wurde (Ingenkamp, K., *Die Fragwürdigkeit der Zensurenvergabe*, Weinheim 1971).

3 Vgl. Wengert, H. G., *Leistungsbeurteilung in der Schule*, in: Bovet, G./Huwendiek, V. (Hrsg.), *Leitfaden Schulpraxis. Pädagogik und Psychologie für den Lehrberuf*, Berlin 4. Aufl. 2004, S. 294.

4 Zu den Vor- und Nachteilen sog. Ziffernnoten vgl. zusammenfassend etwa Jürgens, E./Lissmann, U., *Pädagogische Diagnostik. Grundlagen und Methoden der Leistungsbeurteilung in der Schule*, Weinheim und Basel 2015, S. 40 ff. oder Becker, G. E., *Unterricht auswerten und beurteilen. Handlungsorientierte Didaktik. Teil III*, Weinheim und Basel 2007, S. 72 ff.

5 Hierzu allgemein Winter (Anm. 2), S. 64 oder Rheinberg, F., *Bezugsnormen und schulische Leistungsbeurteilung*, in: Weinert, F. E. (Hrsg.), *Leistungsmessungen in der Schule*, Weinheim und Basel 2. Aufl. 2002, S. 59.

dung (unter 3). Dabei wird die pädagogische Diskussion allerdings bewusst nur nachgezeichnet und nicht bewertet. Im Zentrum der vorliegenden Überlegungen steht vielmehr die Frage, wie diese pädagogischen Aspekte schulrechtlich zu bewerten sind und inwiefern geltendes Schulrecht den pädagogischen Beurteilungsspielraum von Lehrkräften beeinflussen darf oder gar muss. Das Hauptaugenmerk des vorliegenden Beitrags liegt dementsprechend auf der Untersuchung, inwiefern rechtliche Vorgaben die Notenfindung und -vergabe reglementieren und welche Bezugsnormen bei der Erteilung von Zensuren berücksichtigt werden dürfen (unter 4).

2 Funktionen der Notengebung

In der soziologischen und pädagogischen Literatur findet sich eine Vielzahl an Kategorisierungen, die die Funktionen ordinalskalierten Leistungsbewertungen an Schule in differenzierter Weise beschreiben.⁶ Im Folgenden werden die wichtigsten Funktionen kurz beschrieben und voneinander abgegrenzt; es handelt sich im Einzelnen um die Sozialisationsfunktion, die Rückmeldungs- und Orientierungsfunktion, die Motivations- und Disziplinierungsfunktion sowie die Evaluations- und Kontrollfunktion.⁷

2.1 Sozialisationsfunktion

Als moderne Industrie- und Dienstleistungsgesellschaft ist Deutschland eine Leistungsgesellschaft,⁸ in der das soziale Handeln vielfach durch Leistungsmotivation geprägt ist und in der sich der soziale Status der Gesellschaftsmitglieder nach den von ihnen erbrachten Leistungen bestimmt.⁹ In diese gesellschaftlichen Strukturen müssen junge Menschen hineinwachsen und zugleich zu eigenverantwortlichen, selbstreflexions- und handlungsfähigen Individuen erzogen, also nach soziologischer Terminologie sozialisiert werden. Schülerinnen und Schüler müssen demnach lernen und erfahren, dass sich Anstrengung und Leistung auszahlen und honoriert werden: im späteren „wahren Leben“ durch ökonomischen Wohlstand, beruflichen Erfolg und soziale Anerkennung, vorbereitend darauf während der Schulzeit durch „gute Noten“, für die sie wiederum durch das Lob, die Anerkennung und die Zuneigung der Eltern belohnt werden.¹⁰ Die Heranführung an Leis-

6 Grob unterschieden werden vor allem gesellschaftliche und pädagogische Funktionen, vgl. hierzu *Breidenstein, G.* et al., Das Projekt Leistungsbewertung in der Schulklasse, in: Zaborowski, K. U. et al., *Leistungsbewertung und Unterricht. Ethnographische Studien zur Bewertungspraxis in Gymnasium und Sekundarschule*, Wiesbaden 2011, S. 17: „Mit der ‚gesellschaftlichen Funktion‘ ist die Selektions- und Allokationsfunktion angesprochen, die die Schule für die Gesellschaft erfüllt; hier geht es um die Sortierung der nachwachsenden Generation nach Leistung (die ‚Selektionsfunktion‘ der Schule) und deren Verteilung auf die unterschiedlichen gesellschaftlichen und beruflichen Positionen (ihre ‚Allokationsfunktion‘). Mit der ‚pädagogischen‘ Funktion sind die Aufgaben der Rückmeldung, Beratung und Motivierung bezogen auf schulische Lernprozesse gemeint.“ Zur Überschneidung bzw. zum Vergleich dieser beiden Funktionsbereiche *Jürgens/Sacher* (Anm. 1), S. 54.

7 Vgl. hierzu und zur Kritik an den einzelnen Funktionen vor allem *Jürgens/Sacher* (Anm. 1), m. w. N.

8 Vgl. *Jürgens, E.*, *Leistung und Beurteilung in der Schule: Eine Einführung in Leistungs- und Bewertungsfragen aus pädagogischer Sicht*, Sankt Augustin 7. Aufl. 2010, S. 13 ff. m. w. N.

9 Vgl. *Rammstedt, O.*, *Leistungsgesellschaft*, in: Fuchs-Heinritz, W. et al. (Hrsg.), *Lexikon zur Soziologie*, Opladen 3. Aufl. 1994, S. 398.

10 Nach *Wengert* (Anm. 3), S. 296 lernen Kinder mit dem Schuleintritt daher neue Leistungsnormen kennen, „die sich von denen der Familie, der Spielgruppe und des Kindergartens deutlich unterscheiden: Nicht die gute Absicht, das Bravsein oder Sympathie und Liebe bestimmen die Noten, sondern einzig die Handlungsergebnisse.“

tungsprinzipien und -normen, also die Sozialisationswirkung von Zensuren, „gilt als eine der besonders wichtigen Funktionen der Notengebung und als Grund für ihre beharrliche Verteidigung“¹¹.

2.2 Rückmeldungs- und Orientierungsfunktion

Um Schülerinnen und Schüler erfolgreich auf ein Leben in der Leistungsgesellschaft vorbereiten zu können, müssen sie Kenntnis über ihre Leistungen, also ihren Erfolg bzw. Misserfolg erlangen. „In ihrer Rückmeldefunktion sollen die Noten [daher] Auskunft geben über die Leistungen bei einzelnen Arbeiten und über den Leistungsstand in den Fächern.“¹²

Diese Rückmeldungsfunktion ist eng verknüpft mit der Orientierungsfunktion von Noten. Hiermit wird „der Sachverhalt beschrieben, dass sich Schüler anhand ihrer Noten ein Bild darüber machen, wie sie bezüglich ihrer Leistungen in der Schule stehen und welche Fähigkeiten sie besitzen“¹³. Diese Orientierung über den eigenen Leistungsstand kann nur über einen Vergleich erfolgen; Lernende müssen sich (im wahrsten Sinne des Wortes) mit ihren Mitschülerinnen und Mitschülern (oder mit Freundinnen und Freunden, die eine andere Klasse oder Schule besuchen) messen können. Das setzt voraus, dass sie die Zensuren der anderen kennen, um über die relative Platzierung ihrer Leistung in der Klasse einschätzen zu können, wie erfolgreich oder auch erfolglos sie waren.¹⁴

2.3 Motivations- und Disziplinierungsfunktion

Eng verbunden mit der Sozialisations- und der Rückmeldungs- sowie Orientierungsfunktion von Ziffernzensuren ist deren Motivations- und Disziplinierungsfunktion. Schülerinnen und Schüler sollen zu guten Leistungen angespornt, motiviert werden. Insofern sollen „gute“ Zensuren als Be-

11 Winter (Anm. 2), S. 56; ebenso Hofmeyer, W., „Allgemein anerkannte Bewertungsgrundsätze“ als schulrechtliche Bewertungskriterien, Berlin 1988, S. 25, nach dem „der Leistungsaspekt in der Schule im Vergleich zur eigentlichen pädagogischen Aufgabe deutlich im Vordergrund“ steht. So war die Sozialisationsfunktion auch ein Grund für die Einführung von Ziffernnoten: Im 19. Jahrhundert gelangte die preußische Führung zu der Erkenntnis, „dass die alte Führungsschicht des Adels nicht mehr ausreichte, um die Leitungspositionen eines modernen Staates adäquat zu besetzen. In dieser Situation konnte sich das Leistungsprinzip als demokratische Leitvorstellung des liberalen Bürgertums allmählich durchsetzen“ (Wengert [Anm. 3], S. 295). Nicht die Herkunft sollte den Zugang zu gesellschaftlichen Positionen und politischen Ämtern bestimmen, sondern die individuelle Leistung, die im schulischen Bereich in der Einführung der ordinalskalierten Zensuren ihren Spiegel fand.

12 Winter (Anm. 2), S. 45. Diese Rückmeldungsfunktion überschneidet sich deutlich mit dem zum Teil gesondert als *Beurteilungsfunktion* beschriebenen Sinn von Noten: Durch die Rückmeldung der Zensur sollen die Schülerinnen und Schüler über ihren Leistungsstand informiert werden, gleichzeitig erhalten die Eltern einen aussagekräftigen Bericht über die Leistungen ihrer Kinder; vgl. Jürgens/Sacher (Anm. 1), S. 56.

13 Winter (Anm. 2), S. 55.

14 Vgl. Winter (Anm. 2), S. 45. An dieser Stelle ist zu betonen, dass es Lehrkräften aus datenschutzrechtlichen Gründen selbstverständlich verwehrt ist, Noten von Schülerinnen und Schülern coram publico bekannt zu geben; vgl. Oeynhaus, M./Birnbau, C., Schulrecht Nordrhein-Westfalen. Handbuch für die Praxis mit Datenschutzrecht und Lehrerdienstrecht, Stuttgart et al. 2. Aufl. 2005, Rdnr. 439. Zumindest bei schriftlichen Lernerfolgskontrollen kann dieses Hindernis jedoch dadurch überwunden werden, dass die Lehrkraft nicht nur den Notendurchschnitt, sondern auch die Notenverteilung an die Tafel schreibt; hierzu eher kritisch Wengert (Anm. 3), S. 311, der hierin die Gefahr des „Anprangern[s] der oft ohnehin schon bekannten ‚Schlusslichter‘“ sieht. Zudem zeigt die Erfahrung, dass die Mitglieder einer Klasse über die Noten ihrer Mitschülerinnen und Mitschüler ohnehin bestens informiert sind – so ist es beispielsweise üblich, dass die Lernenden die erhaltenen Noten unmittelbar nach der Rückgabe einer Klausur oder einer Klassenarbeit untereinander austauschen.

lohnung empfunden werden „und als ‚Anreiz‘ gelten, nicht in den bisherigen Anstrengungen nachzulassen, um weiterhin so erfolgreich wie bisher sein zu können. Schlechte Noten sollen demgegenüber dazu auffordern, zukünftig bessere Leistungen im Unterricht zu zeigen.“¹⁵

Die Kehrseite der Motivations- bildet die Disziplinierungsfunktion: Leistungsunwillige Schülerinnen und Schüler sollen „schlechte“ Noten als Makel oder „Bestrafung“ empfinden und so dazu angehalten werden, sie durch mehr Leistungs- und Einsatzbereitschaft zu vermeiden.¹⁶

2.4 Berechtigungs- und Selektionsfunktion

Das gesamte deutsche Bildungssystem, aber auch der deutsche Arbeitsmarkt ist geprägt von seinem Berechtigungswesen: Zensuren (und mit ihnen Zeugnisse, Zertifikate und formelle Abschlüsse) bestimmen die spätere Schul- und Berufslaufbahn von Schülerinnen und Schülern – allein diese Kriterien entscheiden über die zu besuchende Schulform, öffnen die Tür zu einem Studium an einer Hochschule und ermöglichen die Ausbildung und die Tätigkeit in einem bestimmten Beruf.¹⁷ „Wegen der überragenden Bedeutung des Prüfungswesens im deutschen Bildungssystem nimmt die Berechtigungsfunktion der Zensur im gesamten Funktionsspektrum einen besonderen Stellenwert ein.“¹⁸

Die Selektionsfunktion von Noten schränkt die (hoch)schulischen und beruflichen Möglichkeiten der mit Ziffernnoten bewerteten Schülerinnen und Schüler noch weiter ein: Gibt es für einen Studien-, Ausbildungs- oder Arbeitsplatz mehr Bewerberinnen und Bewerber (mit der entsprechend erforderlichen formellen Qualifikation und dem Nachweis eines von allen geforderten Bildungszertifikats) als Plätze bzw. Stellen, muss die Hochschule bzw. der Ausbildungsträger oder der Arbeitgeber (aus Kapazitätsgründen und/oder wirtschaftlichen Erwägungen) nach bestimmten Kriterien selektieren. Diese Selektion kann mehr oder weniger formalisiert erfolgen; in jedem

15 Jürgens/Sacher (Anm. 1), S. 56.

16 Vgl. Jürgens/Sacher (Anm. 1), S. 57.

17 So auch Jürgens/Sacher (Anm. 1), S. 55, nach denen „Berechtigungen, die in bestimmten Phasen der Schullaufbahn erworben werden, über den weiteren Bildungs- und Lebensweg entscheiden. Eine zweite oder dritte Chance gibt es in der Regel nicht oder nur über aufwendige Umwege.“ So ist bei schlechten Grundschulnoten bereits der Übergang zum Gymnasium verschlossen, so dass die allgemeine Hochschulreife – bei aller Durchlässigkeit der länderspezifischen Schulsysteme zum Trotz – nur verhältnismäßig schwer erlangt werden kann. Ohne Abitur ist Schulabgängerinnen und -abgängern wiederum die Aufnahme eines Studiums an einer Universität verwehrt, weswegen ihnen mangels eines entsprechenden Hochschulabschlusses die Ausübung bestimmter Berufe (z. B. eine Tätigkeit im anwaltlichen oder ärztlichen Bereich oder in einer Lehrtätigkeit an einer staatlichen Schule) verwehrt bleibt. Das Verfassungsrecht hat dieses Berechtigungsprinzip übrigens nachvollzogen: Zwar gewährt Art. 12 Abs. 1 GG allen Deutschen das Recht, Beruf, Arbeitsplatz und Ausbildungsstätte frei zu wählen, doch gelten hierbei spätestens seit einem Urteil des Bundesverfassungsgerichts vom 11. Juni 1958 (Az. 1 BvR 596/56 [sog. Apotheken-Urteil]) sog. subjektive Zulassungsvoraussetzungen als grundsätzlich zulässige Eingriffe in den Schutzbereich des Art. 12 GG. Hierzu gehören beispielsweise Voraussetzungen, die den Zugang zu bestimmten Ausbildungen oder Berufen vom Vorliegen bestimmter formeller Qualifikationen abhängig machen, wie etwa durch Prüfungen nachgewiesene spezifische Kenntnisse, Fähigkeiten und Fertigkeiten; vgl. etwa Gubelt, in: von Münch, I./Kunig, P., Grundgesetz-Kommentar. Band 1, München 2. Aufl. 2012, Rdnr. 53. Zur Bedeutung der Grundrechte aus Art. 12 GG für Schülerinnen und Schüler vgl. insgesamt Hofmeyer (Anm. 1), S. 32 ff.: „Bedeutung erlangt diese Norm für das Schulwesen, weil die schulische Ausbildung der Berufswahl vorgelagert ist und Schule über die Leistungsbeurteilung einen direkten Einfluss auf die in Art. 12 GG kodifizierte Rechtsposition des einzelnen nimmt“ (a. a. O., S. 32).

18 Jürgens/Sacher (Anm. 1), S. 55.

Fall wird sie sich an Schulnoten und den hinter der formellen Qualifikation liegenden schulischen Leistungsbewertungen orientieren.¹⁹

2.5 Evaluations- und Kontrollfunktion

So wie Schulnoten den Schülerinnen und Schülern Informationen darüber liefern sollen, auf welchem Stand des Lernprozesses sie sich zu einem bestimmten Zeitpunkt befinden (siehe Rückmeldungs- und Orientierungsfunktion), sind sie auch wichtige Indikatoren für die Lehrkräfte selbst: Sie geben ihnen Aufschluss über den Leistungsstand der von ihnen unterrichteten Klasse und erlauben insofern Rückschlüsse auf die Qualität und den Erfolg des eigenen unterrichtlichen Handelns.²⁰ Entspricht etwa das Ergebnis einer Klausur nicht den Erwartungen der Lehrkraft, kann sie neben den Leistungen ihrer Schülerinnen und Schüler auch die Reliabilität und Validität ihrer Aufgabenstellung oder die Adäquanz der Elemente ihrer Unterrichtsgestaltung in Frage stellen.²¹

Hinzu kommt ein rechtlicher Aspekt: Die Zensurenvergabe ist de facto der einzige Bereich, in dem Schülerinnen und Schüler die „Lehrleistung“ ihrer Lehrerinnen und Lehrer juristisch überprüfen lassen können, da beides miteinander korrespondiert: „Zum Schutz der Schüler kann eine juristische Klärung darüber herbeigeführt werden, ob [diesbezüglich; d.V.] Verfahrensmängel oder Rechtsverletzungen seitens der Schule aufgetreten sind“²² – und dies trotz des Bewertungsvorbehalts der beurteilenden bzw. prüfenden Lehrkraft.²³

19 So ist beispielsweise die Vergabe von Studienplätzen in zulassungsbeschränkten Studiengängen derart stark formalisiert, dass für gewöhnlich die Note des Schulabschlusses, der zum Hochschulzugang berechtigt, die entscheidende Rolle bei der Vergabe eines Studienplatzes spielt. Bei der Besetzung von Ausbildungs- und Arbeitsplätzen sind die Personalauswahlverfahren in aller Regel deutlich ausdifferenzierter; die Entscheidung, wer zu einem Auswahlgespräch eingeladen wird, wird gleichwohl häufig nach den Zeugnisnoten der Bewerberinnen und Bewerber getroffen.

20 Vgl. Winter (Anm. 2), S. 60. Neben dieser internen Evaluation gilt entsprechendes auch für die Evaluation durch Dritte, wie etwa die Schulleiterin bzw. den Schulleiter oder die Schulaufsicht (externe Evaluation).

21 Hierzu etwa Kistorz, P./Schlosser, D., Praxis des Unterrichtens, Brake 2014, S. 129 f. bzw. S. 61 ff.

22 Jürgens/Sacher (Anm. 1), S. 55. Justiziabel sind neben den genannten Verfahrensfehlern insbesondere auch inhaltliche Bewertungsfehler; vgl. hierzu Hoegg, G., SchulRecht! Aus der Praxis – für die Praxis, Weinheim und Basel 3. Aufl. 2006, S. 88 ff. Zurückzuführen ist dies vor allem auf zwei Urteile des Bundesverfassungsgerichts aus dem Jahr 1991 (BVerfG vom 17. April 1991 [Az. 1 BvR 419/81 und 213/83] bzw. BVerfG vom 17. April 1991 [Az. BvR 1529/84 und 138/87]). Danach ergibt sich insbesondere für berufsbezogene Prüfungen der allgemeine Bewertungsgrundsatz, dass eine vertretbare und mit gewichtigen Argumenten folgerichtig begründete Lösung einer Schülerin bzw. eines Schülers nicht als falsch bewertet werden darf.

23 Von den genannten Ausnahmen abgesehen räumt die Rechtsprechung mit Rücksicht auf das bei Leistungsbewertungen zu treffende höchstpersönliche Fachurteil Lehrkräften einen weiten Beurteilungsspielraum ein. „Dieser wird mit der Erwägung begründet, ... es widerspräche dem das Prüfungsverfahren prägenden Grundsatz der Chancengleichheit (Art. 3 Abs. 1 GG), wenn einzelne Kandidaten bei einer rechtlichen Überprüfung die Chance einer vom Vergleichsrahmen unabhängigen Bewertung erhielten“ (Avenarius, H./Füssel, H.-P., Schulrecht. Ein Handbuch für Praxis, Rechtsprechung und Wissenschaft, Köln und Kronach 8. Aufl. 2010, S. 443 m. w. N. der Rechtsprechung). Zu den Kriterien der rechtlichen Überprüfung des sog. Beurteilungsspielraums vgl. auch Böhm, T., Grundkurs Schulrecht III. Zentrale Fragen zur Leistungsbeurteilung und zum Prüfungsrecht, Köln und Kronach 2008, S. 3 ff. sowie Hofmeyer (Anm. 11), S. 43 ff.

3 Bezugsnormen bei der Leistungsbewertung

Eine Leistungsbewertung mit Ziffernzensuren verfügt nur dann über Aussagekraft, wenn die erteilte Note an einer Bezugsnorm orientiert ist.²⁴ Möglich ist diesbezüglich dreierlei: die Orientierung an einer Referenzgruppe (sog. kollektive Bezugsnorm oder Sozialnorm), am eigenen Lernfortschritt (sog. individuelle Bezugsnorm oder Fortschrittsnorm) oder an sachlichen Erwägungen (sog. kriteriumsorientierte Bezugsnorm oder Sachnorm).²⁵ Je nachdem, an welcher Bezugsnorm die vergebene Zensur orientiert ist, kann die einer Schülerin bzw. einem Schüler erteilte Note – wie noch zu zeigen sein wird – sehr unterschiedlich ausfallen.²⁶ Es ist daher bereits mehrfach der Vorschlag gemacht worden, Lernenden regelmäßig drei Leistungsnoten zu geben, um „damit den relativen Leistungsstand in der Klasse, den Grad der Kriteriumserfüllung und ... [den] individuellen Fortschritt auszudrücken“²⁷. Im Schulalltag konnte sich eine solche Praxis bislang indes nicht durchsetzen; hier herrscht nach wie vor eine Kombination unterschiedlicher Bezugsnormorientierungen vor.²⁸

3.1 Kollektive Bezugsnorm

Eine ordinalskalierte Leistungsbewertung unter Berücksichtigung einer kollektiven Bezugsnorm folgt dem Grundsatz: *Gut ist, wer über dem Durchschnitt liegt, schlecht, wer darunter liegt*.²⁹ Die individuellen Leistungen der Schülerinnen und Schüler werden also in Relation gesetzt zu den Leistungen der Mitglieder einer Referenzgruppe. Da die beurteilende Lehrkraft nur die jeweilige Klasse als Referenzgruppe kennt (und nicht etwa sämtliche Schülerinnen und Schüler einer bestimmten Jahrgangsstufe an einer bestimmten Schulform oder eines bestimmten Bildungsgangs), orientiert sich ihre Bewertung an den durchschnittlichen Ergebnissen aller Mitschülerinnen und Mitschüler der jeweiligen Lerngruppe.³⁰ Diese Art der Leistungsbewertung wird mithin besonders der Sozialisations-, der Motivations- und Disziplinierungs- sowie der (Rückmeldungs- und) Orientierungsfunktion von Noten gerecht: Schülerinnen und Schüler erfahren, dass überdurch-

24 Hierzu allgemeiner *Jürgens* (Anm. 8), S. 26, der zu Recht ausführt, es gebe keine Leistung per se: „[D]ie Entscheidung, ob ein Verhalten – Prozess oder Ergebnis – als Leistung bezeichnet wird, bedarf der Definition. Daraus resultiert: Um eine Bestimmung einer Leistung vornehmen zu können, ist die Existenz einer Norm erforderlich“.

25 Zu den Vor- und Nachteilen der einzelnen Bezugsnormen vgl. vor allem *Rheinberg* (Anm. 5).

26 Zur mathematisch-statistischen Berechnung von Ziffernzensuren nach den drei genannten Bezugsnormen ausführlich *Klauer, K. J.*, Zensierungsmodelle und ihre Konsequenzen für die Notengebung, in: Jäger, R. S. et al. (Hrsg.), *Test und Trends 7. Jahrbuch der pädagogischen Diagnostik*, Weinheim 1989, S. 40–68.

27 *Winter* (Anm. 2), S. 49 unter Hinweis auf *Klauer* (Anm. 26), S. 44. Hierzu auch *Hofmeyer* (Anm. 11), S. 213, der ausführt, „daß es für eine Leistungsmessung drei Noten geben kann, was aus pädagogischer Sicht ein Maximum an Aussagekraft bedeutete“; hierzu bildet er folgendes Beispiel: „Eine Leistungsbeschreibung könnte dann wie folgt lauten: Die Leistung des Schülers X ist unterdurchschnittlich (gruppenorientiert), entspricht dem Lernziel (kriteriumsorientiert), ist hervorragend für diesen Schüler (individuell), was ebenso mit den Noten ‚mangelhaft‘, ‚ausreichend‘ und ‚sehr gut‘ ausgedrückt werden könnte“ (a. a. O., S. 227).

28 Vgl. *Rheinberg, F.*, Bezugsnormorientierung, in: Arnold, K.-H. et al. (Hrsg.), *Handbuch Unterricht*, Bad Heilbrunn 2006, S. 647.

29 Ähnlich *Rheinberg* (Anm. 5), S. 61.

30 Hierzu beispielhaft *Jürgens/Sacher* (Anm. 1), S. 65: „In einer Klasse wird eine Klausur geschrieben, bei der maximal 30 Punkte zu erreichen sind. Erreicht ein Schüler 28 Punkte in der Klausur, kann man diese mit dem durchschnittlichen Punktwert der Lerngruppe vergleichen. ... Liegt das Resultat über dem durchschnittlichen Punktwert, bekommt der Schüler eine Eins oder Zwei; erreicht er nicht den durchschnittlichen Punktwert [sic!] fällt die Zensur schlechter als Drei aus; und ist das Resultat dem durchschnittlichen Punktwert gleich, bekommt er eine Drei.“

schnittliche Anstrengungen und Leistungen auch zu überdurchschnittlichen Ergebnissen bzw. Erfolgen führen; sie sollen – bezogen auf die soziale Gruppe ihrer Klasse – zu entsprechenden Anstrengungen motiviert werden und so eine Orientierung über ihr Leistungsvermögen im Vergleich mit ihren Klassenkameradinnen und -kameraden erhalten.

Die kollektive Bezugsnorm gilt als die am häufigsten angewendete Bezugsnorm;³¹ ihre Umsetzung erfolgt (bewusst oder unbewusst) regelmäßig orientiert an der Gauß'schen Normalverteilung.³² Dies hat ihren Grund nicht zuletzt darin, dass Lehrerinnen und Lehrer bei einer Normalverteilung von Leistungsnoten die geringste Kritik von Lernenden, Eltern und der Schulleitung zu befürchten haben³³ und in ihrer Umsetzung die beste Möglichkeit sehen, der (aus ihrer Sicht) manchmal unangenehmen oder gar entlarvenden Evaluation und Kontrolle ihres eigenen Unterrichts zu entgehen. Das bedeutet häufig auch, dass Bewertungsraster im Nachhinein so geändert werden, dass das Ergebnis einer Leistungskontrolle der Normalverteilung entspricht.³⁴ Besondere Bedeutung erlangt das unter Berücksichtigung einer schulrechtlich festgelegten Quote, bei der eine schriftliche Arbeit wiederholt werden muss, wenn ein bestimmter Anteil der Schülerinnen und Schüler nicht mindestens die Note 4 (*ausreichend*) erhalten und die Schulleitung nach erfolgter Prüfung keine Ausnahmeentscheidung getroffen hat.³⁵ Hier ist nicht selten zu beobachten, dass betroffene Lehrerinnen und Lehrer ihre Leistungsbewertung bzw. das zugrundeliegende Be-

31 Vgl. *Rheinberg* (Anm. 5), S. 64.

32 Vgl. etwa *Paradies, L. et al.*, Leistungsmessung und -bewertung, Berlin 2005, S. 32. In Anlehnung an die Gauß'sche Normalverteilungskurve erhalten die 10% der besten Schülerinnen und Schüler die Note *sehr gut*, die 10% schlechtesten ein *mangelhaft* bzw. *ungenügend*. Den mittleren 34% wird die Note *befriedigend* erteilt und den jeweils 23% besseren bzw. schlechteren Schülerinnen und Schülern die Note *gut* bzw. *ausreichend*; vgl. *Wengert* (Anm. 3), S. 298. Nach *Klawer* (Anm. 26), S. 42 kann eine Klausur oder Klassenarbeit, die auf diese Weise bewertet wird, nie „gut“ oder „schlecht“ ausfallen, da die Noten von Arbeit zu Arbeit relativ stabil sind und die Rangordnung der Schülerinnen und Schüler in aller Regel weitgehend erhalten bleibt.

33 *Winter* (Anm. 2), S. 61. So auch *Wengert* (Anm. 3), S. 298, nach dem nach weit verbreiteter Auffassung „erst eine stärkere Abweichung von der Normalverteilung erklärungsbedürftig [ist], wenn etwa eine Klassenarbeit ‚zu schlecht‘ oder ‚zu gut‘ (!) ausgefallen ist“. In diesem Zusammenhang gibt er zu bedenken, Lehrer könnten bzw. sollten „mit einem Unterrichtsergebnis wenig zufrieden sein ..., wenn die Lernleistungen ‚normal‘ verteilt sind und es immer einen gewissen Anteil an Versagern gibt“ (a. a. O., S. 298 f.).

34 *Paradies et al.* (Anm. 32), S. 16. In der Nomenklatur von Lehrkräften werden dann beispielsweise Klausuren oder Klassenarbeiten „herauf-“ oder „herabgesetzt“.

35 Wegen der entsprechenden Quote werden derartige Richtlinien häufig auch *Drittelerauslass* genannt; vgl. exemplarisch *Hessen*: „Ist mehr als ein Drittel der abgelieferten schriftlichen Arbeiten mit den Noten mangelhaft oder ungenügend ... bewertet worden, ist die Arbeit einmal zu wiederholen, sofern nicht die Schulleiterin oder der Schulleiter nach Beratung mit der Fachlehrerin oder dem Fachlehrer entscheidet, dass die Arbeit zu werten sei. Die Arbeit ist zu wiederholen, wenn mehr als die Hälfte mit den Noten mangelhaft oder ungenügend ... bewertet wurde“ (§ 34 Abs. 1 Verordnung zu Gestaltung des Schulverhältnisses), *Niedersachsen*: „Zeigt sich bei der Korrektur und Bewertung, dass mehr als 30% der Arbeiten einer Klasse oder Lerngruppe mit ‚mangelhaft‘ oder ‚ungenügend‘ bewertet werden müssen, so wird die Arbeit nicht gewertet. Von dieser Vorschrift darf mit Zustimmung der Schulleiterin oder des Schulleiters abgewichen werden“ (Punkt 8 RdErl. Schriftliche Arbeiten in den allgemein bildenden Schulen), *Schleswig-Holstein*: „Wenn ein Drittel oder mehr der Klassenarbeiten einer Klasse mit weniger als vier Punkten bewertet werden soll, ist die Genehmigung der Schulleiterin bzw. des Schulleiters erforderlich. Dazu müssen die unterrichtende Lehrkraft und die Klassensprecherin oder der Klassensprecher gehört werden“ (Punkt II.7 RdErl. Zahl und Umfang der Klassenarbeiten in der gymnasialen Oberstufe). Ähnliche Regelungen finden sich darüber hinaus in § 17 Abs. 6 SekI-VO BE, Punkt 8 Abs. 5 VV-Leistungsbewertung BB, Punkt 4.4 Erl. Schriftliche Arbeiten im Unterricht der allgemeinbildenden Schulen in den Jahrgangsstufen 5 bis 10 HB, Punkt 5 Abs. 2 Richtlinie für schriftliche Lernerfolgskontrollen in allgemeinbildenden Schulen (Klassen 3 bis 10) HH, § 7 Abs. 9 LeistBewVO MV, § 53 Abs. 5 ÜSchO RP, Punkt 1.3 Klassenarbeitenerlass SL, Punkt 4.1.17 RdErl. Leistungsbewertung und Beurteilung an berufsbildenden Schulen ST.

wertungsraster in der Hinsicht korrigieren, dass eine höhere Quote der Schülerinnen und Schüler das Ziel des Bestehens der schriftlichen Arbeit erreicht hat.³⁶

3.2 Individuelle Bezugsnorm

Bei der Anwendung einer individuellen Bezugsnorm bei der Zensurenvergabe gilt der Grundsatz: *Gut ist, wer im Lernprozess Fortschritte gemacht hat, schlecht, wessen Leistungen rückläufig sind.* Die individuellen Leistungen einer Schülerin bzw. eines Schülers werden also nicht mit den durchschnittlichen Ergebnissen der gesamten Klasse verglichen, sondern mit den eigenen Leistungen bei einer zurückliegenden Lernerfolgskontrolle; erfasst wird also der persönliche Leistungszuwachs bzw. Lernfortschritt.³⁷

Begründet mit der Ermangelung eines intersubjektiven Vergleichsmaßstabs wird häufig die Auffassung vertreten, eine dieser Bezugsnorm folgende Leistungsbewertung könne nur in Form einer mündlichen Rückmeldung oder einer verbalen Beurteilung erfolgen und sei dem Grunde nach einer ordinalskalierten Zensurenvergabe unzugänglich.³⁸ Dem ist indes nicht so:³⁹ Schülerinnen und Schüler, die sich im zeitlichen Verlauf (deutlich) verbessert haben, könnte die Note 2 (*gut*) oder sogar die Note 1 (*sehr gut*) erteilt werden, Schülerinnen und Schülern mit einem nach diesem Vergleichsmaßstab abnehmenden Leistungsniveau die Note 4 (*ausreichend*) oder gar schlechter; diejenigen Lernenden, deren Leistungen den bisherigen entsprachen, erhielten die Note 3 (*befriedigend*).⁴⁰

In der Schulpraxis spielt die ordinalskalierte Leistungsbewertung nach der individuellen Bezugsnorm zwar keine bedeutende Rolle, findet aber (zumindest begrenzt) gleichwohl Anwendung: Schülerinnen und Schüler, die beispielsweise bei der Ermittlung einer Zeugniszensur zwischen zwei Ziffernnoten stehen, erhalten bei einem über den Beurteilungszeitraum festzustellenden po-

36 Hierzu differenziert und ausführlich *Rheinberg* (Anm. 5), S. 67: „Diese Regelung drängt dem Lehrer im unteren Teil der Notenskala faktisch die soziale Bezugsnorm auf.“

37 Vgl. etwa *Paradies et al.* (Anm. 32), S. 32 oder *Wengert* (Anm. 3), S. 299. So auch *Rheinberg* (Anm. 5), S. 65: „Damit geht der individuelle Lernzuwachs direkt in die Leistungsbeurteilung ein und wird besonders deutlich gemacht.“ Bezogen auf die Funktionen von Noten sprechen vor allem motivationspsychologische Gründe dafür, sich dieser Fortschrittsnorm zu bedienen; vgl. etwa *Winter* (Anm. 2), S. 49 oder *Rheinberg* (Anm. 5), S. 69.

38 Vgl. etwa *Paradies et al.* (Anm. 32), S. 32 oder *Wengert* (Anm. 3), S. 299.

39 Zu möglichen Zensurierungsmodellen vgl. *Klauer* (Anm. 26), S. 50 ff.

40 Hierzu beispielhaft *Jürgens/Sacher* (Anm. 1), S. 65 (anknüpfend an Anm. 30): „Man kann ... 28 Punkte [von insgesamt 30 zu erreichenden; d.V.] aber auch mit dem Resultat vergleichen, das der Schüler bei der letzten Klausur erzielt hat (z. B. ebenfalls 28 Punkte). Übertrifft das jetzige Resultat das bisherige Abschneiden, so wäre das eine bessere, bleibt es dahinter zurück, eine schlechtere Leistung. ... Wendet man nun die Zensurengebung an, würde dies bedeuten, dass je nach Grad der Verbesserung bzw. Verschlechterung die Leistung benotet werden würde. Dies hätte zur Konsequenz, dass der Schüler in diesem Beispiel, bezogen auf das Lernziel, zwar eine ‚sehr gute‘ Klausur geschrieben hat, sich im Vergleich zum letzten Resultat aber nicht verbesserte und demnach nur eine Drei bekäme.“ Insofern ist auch die soeben dargestellte Motivationswirkung der Leistungsbewertung nach einer individuellen Bezugsnorm in Frage zu stellen, da Schülerinnen und Schüler auf einem gleichbleibend hohen Kompetenzniveau bestenfalls „befriedigende“ Noten erhalten würden. Aus diesem Grund ist auch *Rheinberg* (Anm. 5), S. 65 nur bedingt in der Auffassung zuzustimmen, „dass leistungsschwächere Schüler von der individuellen Bezugsnorm besonders profitieren, ohne dass leistungstärkere benachteiligt wären“.

sitiven Leistungsfortschritt häufig ohne weitere Maßnahmen der Leistungsdiagnostik die bessere Note (sog. positive Tendenz der erbrachte Leistungen).⁴¹

3.3 Kriteriumsorientierte Bezugsnorm

Eine Leistungsbewertung nach einer kriteriumsorientierten Bezugsnorm orientiert sich an der Erreichung eines *vor* der Leistungsmessung klar operationalisierten Lernziels bzw. dem Erwerb eines *zuvor* klar bestimmten angestrebten Kompetenzzuwachses;⁴² allein der Grad der Erfüllung dieser Sachnorm ist dann ausschlaggebend für die Vergabe der Leistungsnote, wobei diese Sachnorm für alle Schülerinnen und Schüler einer Schulform, einer Jahrgangsstufe bzw. eines Bildungsganges gelten muss und nicht individuell festgelegt werden darf. Es gilt mithin der Grundsatz: *Gut ist, wer allgemeine Erwartungen an den Kompetenzerwerb erfüllt, schlecht, wer sie nicht erfüllt.*

Dabei wird zum Teil angeführt, „dass die übliche Notenskala hier fehl am Platze“ sei, da sich die Beurteilung auf die Dichotomie *Lernziel erreicht* und *Lernziel (noch) nicht erreicht* beschränken müsse.⁴³ Dem ist entgegenzuhalten, dass das Erreichen eines Unterrichtsziels oder eines angestrebten Kompetenzerwerbs nicht ausschließlich schwarz und weiß gemalt werden kann.⁴⁴ Zum einen zeigt bereits die Schulpraxis, dass eine Schülerin oder ein Schüler bei einer Lernerfolgskontrolle den von der Lehrkraft skizzierten Erwartungshorizont in den seltensten Fällen entweder in Gänze oder gar nicht erreicht.⁴⁵ Zum anderen besteht das Ziel des Unterrichts nach neuerem, sich von einer klassischen und reinen Lernzielorientierung abwendendem Verständnis eher darin, unterschiedliche und zudem interdependent wirkende Kompetenzen zu fördern.⁴⁶ Auch bei kompetenzorientierten Prüfungen können daher graduelle Abstufungen gemacht werden, inwiefern das Ziel einer umfassenderen Handlungskompetenz (noch) nicht, teilweise oder vollständig erreicht ist, so dass auch hier hinsichtlich des Zielerreichungsgrades Unterschiede bei einer ordinalskalierten Leistungsbeurteilung gemacht werden können.⁴⁷

41 Interessanterweise wird in umgekehrt gelagerten Fällen selten entsprechend gehandelt; hier bewerten Lehrerinnen und Lehrer häufig nach dem Grundsatz *in dubio pro discipulo* oder verlangen zusätzliche Leistungsnachweise, wie beispielsweise das Halten eines Referats oder das Anfertigen eines Protokolls.

42 Vgl. exemplarisch Wengert (Anm. 3), S. 299. Zu betonen ist, dass diese Sachnorm tatsächlich *vor* der Leistungsmessung festzulegen ist.

43 So etwa Wengert (Anm. 3), S. 299.

44 Ähnlich Hofmeyer (Anm. 11), S. 223.

45 Darauf sind auch die meisten Lernerfolgskontrollen gar nicht angelegt – im Gegenteil: In den zahlreichen Ratschlägen und Handreichungen zur Erstellung von Lernerfolgskontrollen wird vielmehr angeregt, „Aufgaben auf unterschiedlichen Anforderungsniveaus zu stellen“ (Paradies et al. [Anm. 32], S. 85) und bei der Formulierung „auf die geeigneten Taxonomiestufen (Wissen, Verstehen, Anwenden, Analyse, Synthese, Evaluation) und Aufgabenarten (mit gebundenen Antworten, mit frei zu formulierenden Kurzantworten, mit längeren Einlassungen) zu achten“ (Wengert [Anm. 3], S. 307), um die Leistungsunterschiede der Schülerinnen und Schüler einer Klasse reliabel und valide messen zu können (hierzu Kostorz/Schlösser [Anm. 21], S. 129 f.).

46 Hierzu etwa Matthes, E., Kompetenzorientierter Unterricht, in: Bildung und Erziehung 2013, S. 121–125; zur Kritik und der Frage, ob sich kompetenzorientierter Unterricht tatsächlich wesentlich vom Modell des tradierten lernzielorientierten Unterrichts unterscheidet Arnold, K.-H., Kompetenz- versus Lernzielorientierung von Unterricht?, in: Bildung und Erziehung 2013, S. 173–187 sowie Wiater, W., Kompetenzorientierung des Unterrichts – alter Wein in neuen Schläuchen? Anfragen Seitens der allgemeinen Didaktik, in: Bildung und Erziehung 2013, S. 145–161.

47 Hierzu liegen bereits eine Reihe fachdidaktischer Beiträge bzw. Konzepte vor; vgl. etwa Ralle, B. (Hrsg.), Lernaufgaben entwickeln, bearbeiten und überprüfen: Ergebnisse und Perspektiven der fachdidaktischen Forschung, Münster 2014; exemplarisch für die Pflegeausbildung Bonse-Rohmann, M. et al. (Hrsg.), Kompetenzorientiert prüfen. Lernen und Leistungsüberprüfungen in der Pflegeausbildung, München 2008 oder für den Deutschunterricht Kreutel, C.,

Der Abgleich der individuell erbrachten Schülerleistungen mit einem objektiv bestimmten Bewertungskriterium führt daher unter Umständen zu gänzlich unterschiedlichen Notenverteilungen innerhalb einer Lerngruppe als bei der Anwendung einer sozialen Bezugsnorm bei der Leistungsbewertung: Bei einer Orientierung an der Gauß'schen Normalverteilung, wie es bei der Berücksichtigung einer kollektiven Bezugsnorm üblich ist, erhalten die meisten Schülerinnen und Schüler die Note 3 (*befriedigend*) und entsprechend weniger die Noten 1 (*sehr gut*) oder 2 (*gut*) bzw. 4 (*ausreichend*) oder schlechter.⁴⁸ Eine Leistungsbewertung, die sich an einem sachbezogenen Beurteilungskriterium orientiert, kann im Extrem dazu führen, dass alle Schülerinnen und Schüler einer Klasse eine gleich „gute“ oder gleich „schlechte“ Zensur erhalten – jedenfalls dann, wenn alle Mitglieder der Lerngruppe die prospektiv festgelegte Sachnorm (trotz eventuell bestehender gradueller Unterschiede) mehr oder minder voll oder gar nicht erfüllen.⁴⁹

Die Leistungsbewertung an einer sachlichen Bezugsnorm zu orientieren ist immer dann sinnvoll, wenn „bestimmte Mindestkompetenzen, insbesondere wegen gravierender Folgen erreicht sein müssen und wo sich solche Mindestkompetenzen messen lassen“, was gerade bei Beurteilungen der Fall ist, „die auch Personen oder Instanzen außerhalb eines Ausbildungsabschnittes informieren“⁵⁰. Damit ist eine kriteriumsbezogene Zensurenvergabe immer dann wichtig, wenn Berechtigungszertifikate vergeben werden oder die erteilten Noten (auch) der Selektion und Allokation dienen sollen; in diesen Fällen müssen auch die „Abnehmer“ der Schülerinnen und Schüler (weiterführende Schulen, Hochschulen, Ausbildungsbetriebe etc.) möglichst objektiv über deren Kompetenzniveau in Kenntnis gesetzt werden.

4 Schulrechtliche Aspekte

In der pädagogischen Literatur zur Leistungsbewertung ist häufig der lapidare Hinweis zu finden, (allein) die Benotung nach der kriteriumsorientierten Bezugsnorm entspreche den geltenden schulrechtlichen Bestimmungen.⁵¹ Wenn dem (uneingeschränkt) so wäre, würden viele Noten, die sich (auch) an der kollektiven oder der individuellen Bezugsnorm orientieren, schlicht rechtswidrig erteilt werden – mit der Folge, dass die Schülerinnen und Schüler (oder ihre Eltern als gesetzliche Vertreter) juristisch gegen sie vorgehen könnten. Dann wäre zudem die gesamte Diskussion um die Funktion der Notengebung und die Bezugsnormen bei der Leistungsbewertung dem Grunde nach redundant und bestenfalls mit der eher rechtspolitischen Intention zu rechtfertigen, bestehende schulrechtliche Regelungen unter pädagogischen Gesichtspunkten neu zu ordnen.⁵²

Praxis der Bewertung im kompetenzorientierten Unterricht, in: Saxalber, A./Wintersteiner, W. (Hrsg.), Reifeprüfung Deutsch, Innsbruck et al. 2012, S. 140–148.

48 Vgl. hierzu oben Anm. 32.

49 So auch *Klawer* (Anm. 26), S. 43 sowie *Winter* (Anm. 2), S. 48; vgl. ebenfalls *Rheinberg* (Anm. 5), S. 67: „Dabei würde es sicher nicht nur theoretisch, sondern auch tatsächlich geschehen können, dass ganze Schulklassen ein ‚gut‘ oder ‚sehr gut‘, aber auch ganze Schulklassen (vielleicht sogar ganze Schulen) nur ‚mangelhaft‘ oder ‚ungenügend‘ erhielten.“

50 Beide Zitate *Rheinberg* (Anm. 5), S. 66.

51 Vgl. exemplarisch *Paradies* et al. (Anm. 32), S. 32 oder *Hofmeyer* (Anm. 11), S. 220 sowie S. 226 f.

52 So hält auch *Jürgens* (Anm. 8), S. 61 die zahlreichen Funktionszuschreibungen für erstaunlich, „weil Zensuren und Zeugnisse in ihrer Entstehung und in ihrer bisherigen Geschichte niemals in erster Linie pädagogische Instrumente gewesen sind, sondern überwiegend durch die gesellschaftliche Aufgabe der Selektion und die Aufgabe, Berechtigungen zu erteilen, geprägt waren“.

In der Bundesrepublik ist Schulrecht indes Landesrecht, so dass es nicht „die“ rechtliche Bestimmung zur Leistungsbewertung an Schulen gibt;⁵³ vielmehr existiert eine breite Fülle an entsprechenden juristischen Vorgaben, die im Folgenden dargestellt werden. Dabei zeigt sich, dass eine Leistungsbewertung nach der kriteriumsorientierten Bezugsnorm zwar den Grundsatz darstellt, Ausnahmen hiervon aber durchaus (auch rechtlich) möglich sind, wenn sie pädagogisch begründet werden können (vgl. zusammenfassend Tab. 1).

Tab. 1: Regelungen zu den Bezugsnormen in den Bundesländern

	Grundsatz: kriteriums- orientierte Bezugsnorm	Ausnahme: (gebundenes) Ermessen	Ausnahme: individuelle Bezugsnorm	Ausnahme: kollektive Bezugsnorm	Ausnahme: sog. Drittel- Erlaß*
Baden-Württemberg	X	X			
Bayern	X				
Berlin	X		X		X
Brandenburg	X	X	X	X	X
Bremen	X	X			X
Hamburg	X				X
Hessen	X				X
Mecklenburg-Vorpommern	X				X
Niedersachsen	X	X			X
Nordrhein-Westfalen	X	X			
Rheinland-Pfalz	X		X	X	X
Saarland	X				X
Sachsen	X				
Sachsen-Anhalt	X	X			X
Schleswig-Holstein	X	X			X
Thüringen	X	X			

(* Regelung zur Überprüfung/Wiederholung von Arbeiten bei einem bestimmten Notenspiegel)

4.1 Grundsatz: Leistungsbewertung nach der kriteriumsorientierten Bezugsnorm

Nach dem sogenannten Hamburger Abkommen von 1964 sind die Leistungen der Schülerinnen und Schüler nach Ziffernoten zwischen *sehr gut* und *ungenügend* zu bewerten,⁵⁴ wobei die Notestufen nach einem ergänzenden Beschluss der Kultusministerkonferenz von 1968 in allen Bun-

53 So auch *Bohl, T.*, Aktuelle Regelungen zur Leistungsbeurteilung und zu Zeugnissen an deutschen Sekundarschulen. Eine vergleichende Studie aller Bundesländer – Darstellung und Diskussion wesentlicher Ergebnisse, in: *Zeitschrift für Pädagogik* 2003, S. 550–556.

54 Vgl. § 19 Abs. 2 des Abkommens zwischen den Ländern der Bundesrepublik zur Vereinheitlichung auf dem Gebiete des Schulwesens vom 28. Oktober 1964 i. d. F. vom 14. Oktober 1971.

desländern bei nur graduellen Unterschieden in der Formulierung ähnlich aufgeschlüsselt werden (vgl. Tab. 2).⁵⁵

Tab. 2: Notenstufen und deren Bedeutung⁵⁶

#	Note	Bedeutung
1	sehr gut	Die Leistung entspricht den Anforderungen im besonderen Maße.
2	gut	Die Leistung entspricht den Anforderungen voll.
3	befriedigend	Die Leistung entspricht im Allgemeinen den Anforderungen.
4	ausreichend	Die Leistung weist zwar Mängel auf, entspricht aber im Ganzen noch den Anforderungen.
5	mangelhaft	Die Leistung entspricht nicht den Anforderungen, lässt jedoch erkennen, dass die notwendigen Grundkenntnisse vorhanden sind und die Mängel in absehbarer Zeit behoben werden können.
6	ungenügend	Die Leistung entspricht nicht den Anforderungen und selbst die Grundkenntnisse sind so lückenhaft, dass die Mängel in absehbarer Zeit nicht behoben werden können.

Gemeinsames Merkmal aller sechs Notenbeschreibungen ist die Anknüpfung der Leistungsbeurteilung an den Grad der Erfüllung von (zunächst nicht näher bestimmten⁵⁷) *Anforderungen*; sie muss daher grundsätzlich einer kriteriumsorientierten Bezugsnorm folgen. Diese jeweils schulform-, bildungsgang- und jahrgangsstufenspezifischen Anforderungen ergeben sich im Wesentlichen aus den entsprechend geltenden Richtlinien, Lehrplänen und Bildungsstandards,⁵⁸ in denen die Kompetenzen festgelegt sind, die die Schülerinnen und Schüler erwerben und in Lernerfolgskontrollen zeigen bzw. anwenden sollen.⁵⁹ In einigen Bundesländern wird auf diese Regelwerke sogar explizit Bezug genommen.⁶⁰

55 Vgl. *Baden-Württemberg*: § 5 Abs. 2 NotenbildungsVO; *Bayern*: Art. 52 Abs. 2 BAYEUG; *Berlin*: § 58 Abs. 3 SchulG; *Brandenburg*: § 57 Abs. 3 BbgSchulG; *Bremen*: § 5 Abs. 1 ZeugnisVO; *Hamburg* (exemplarisch): § 2 Abs. 2 APO-GrundstGY; *Hessen*: § 73 Abs. 4 HSchG; *Mecklenburg-Vorpommern*: § 62 Abs. 4 SchulG; *Niedersachsen*: Punkt 3.4.1 RdErl. Zeugnisse in den allgemein bildenden Schule; *Nordrhein-Westfalen*: § 48 Abs. 3 SchulG; *Rheinland-Pfalz*: § 53 Abs. 2 ÜSchO; *Saarland* (exemplarisch): § 5 Abs. 1 ZVO-Gym; *Sachsen* (exemplarisch): § 19 Abs. 1 SOGYA; *Schleswig-Holstein*: § 4 Abs. 1 ZVO; *Thüringen*: § 59 Abs. 2 ThürSchulO.

56 Vgl. Beschluss der Kultusministerkonferenz über die Erläuterung der Notenstufen bei Schulzeugnissen und Einzelergebnissen in staatlichen Prüfungszeugnissen vom 3. Oktober 1968 sowie *Avenarius/Füssel* (Anm. 23), S. 427 f.

57 *Rux, J./Niehues, N.*, Schulrecht, München 5. Aufl. 2013, Rdnr. 500 halten die Notenstufen daher zu Recht für unbestimmte Rechtsbegriffe.

58 Vgl. *Avenarius, H./Füssel, H.-P.*, Schulrecht im Überblick, Darmstadt 2008, S. 124 oder S. 446 f. und S. 428.

59 Hierzu allgemein *Wiater, W.*, Lehrplan, Curriculum, Bildungsstandards, in: Arnold, K.-H. et al. (Hrsg.), Handbuch Unterricht, Bad Heilbrunn 2006, S. 169: „Ein Lehrplan ist die staatlich verbindlich gemachte, geordnete Zusammenfassung von Lehrinhalten und Lehrzielen (Wissen, Können, Einstellungen, Verhalten), die während eines bestimmten Zeitraums an bestimmten Schulformen in bestimmten Fächern/Lernbereichen vermittelt werden sollen. Anstelle des Begriffs Lehrplan werden verschiedene andere Bezeichnungen verwendet: Bildungspläne, Richtlinien, Rahmenpläne, Curricula, curriculare Lehrpläne und (heute) Kerncurricula.“

60 Vgl. *Baden-Württemberg*: „Der Begriff ‚Anforderungen‘ ... bezieht sich auf die im Bildungsplan oder Lehrplan festgelegten Leitgedanken, Kompetenzen, Ziele und Inhalte, insbesondere auf den Umfang, auf die selbständige, richtige und prozessorientierte Anwendung der geforderten Kenntnisse, Fähigkeiten und Fertigkeiten sowie auf die Art der Darstellung“ (§ 5 Abs. 2 Satz 2 SchulG); *Berlin*: „Für die Leistungsbeurteilung maßgebend ist der nach Kriterien des Bil-

Diese immense grundsätzliche Bedeutung der kriteriumsorientierten Zensurenvergabe hängt mit dem hohen Gewicht der Berechtigungs- und Selektionsfunktion von Noten und „der zweifelsfrei belegten Tatsache zusammen, dass Noten zwar innerhalb einer Schulklasse einigermaßen aussagekräftig sind, was Leistungsniveauunterschiede angeht, nicht aber, wenn man die Noten mit denen aus anderen Klassen, Schulen oder Bundesländern vergleicht“⁶¹: Hochschulen bzw. Arbeitgeber etc., die bei ihrer Entscheidung zur Zulassung bzw. Einstellung auf die Gültigkeit und Aussagekraft von Zertifikaten und Zensuren vertrauen müssen, müssen sich ein möglichst objektives Bild von den Kompetenzen der Bewerberinnen und Bewerber machen können,⁶² weshalb Lehrerinnen und Lehrer „von dem Bemühen geleitet sein [müssen], einen ‚absoluten‘ Maßstab (‚Standard‘) zugrunde zu legen, also zu ermitteln, ob und bis zu welchem Grade die Schüler die Lernziele erreicht haben“⁶³.

Dies ist auch und gerade vor dem Hintergrund des Grundsatzes der Gleichbehandlung nach Art. 3 GG wichtig.⁶⁴ Sofern Schülerinnen und Schüler die im Verhältnis zu einer kriteriumsorien-

dungsgangs festgestellte Entwicklungsstand der Kenntnisse, Kompetenzen, Fähigkeiten und Fertigkeiten der Schülerin oder des Schülers“ (§ 58 Abs. 5 Satz 2 SchulG); *Hamburg*: „Die Anforderungen und die Anforderungsebenen, auf die sich die Leistungsbewertung bezieht, ergeben sich aus den Bildungsplänen“ (§ 2 Abs. 1 Satz 2 APO-GrundGy); *Sachsen*: „Die Leistung des Schülers ist bezogen auf die Anforderungen der im Lehrplan festgelegten Ziele und Inhalte zu beurteilen“ (§ 19 Abs. 1 1. Halbsatz SchulG); *Sachsen-Anhalt*: „Der Begriff ‚Anforderungen‘ bezieht sich auf das jeweilige abschlussbezogene Niveau, die im Lehrplan oder in den Rahmenrichtlinien festgelegten Ziele, Kompetenzen und Wissensbestände“ (Punkt 6.2 des RdErl. Leistungsbewertung und Beurteilung an allgemeinbildenden Schulen und Schulen des Zweiten Bildungsweges der Sekundarstufen I und II). In *Rheinland-Pfalz* wird demgegenüber recht unbestimmt ausgeführt, die Leistungen würden „nach dem Grad des Erreichens von Lernanforderungen beurteilt“ (§ 53 Abs. 1 Satz 1 SchulG); ebenso unbestimmt die entsprechende Vorschrift in *Thüringen*, nach der sich der Begriff der Anforderungen „auf den Umfang sowie auf die selbständige und richtige Anwendung der Kenntnisse, Fähigkeiten und Fertigkeiten sowie auf die Art der Darstellung“ bezieht (§ 59 Abs. 2 Satz 2 ThürSchulO).

61 *Winter* (Anm. 2), S. 42.

62 Hierzu indes kritisch *Breidenstein et al.* (Anm. 6), S. 22 ff., die die Selektionsentscheide der Schule keinesfalls allein als gesellschaftlich determiniert ansehen, sondern insofern als dem Erziehungssystem immanent, als Berechtigungen in erster Linie für den Verbleib bzw. Optionen innerhalb des Bildungswesens ausgestellt würden: „Die Schule sortiert die Schüler nach Leistung für die Schule. Selektionsentscheidungen betreffen – empirisch – fast immer Optionen, Schwellen und ‚Karrieren‘ innerhalb des Erziehungssystems: Die Versetzung in die nächste Jahrgangsstufe, die Schullaufbahnentscheidung nach der vierten Klasse, die Einteilung in verschiedene Kurse und Niveaustufen und schließlich die Berechtigung für das Studium an einer Hochschule bzw. der Zugang zu bestimmten Studiengängen. Das Abitur berechtigt zu nichts anderem als zum Verbleib im Bildungssystem und die konkrete Abiturnote entscheidet über nichts anderes als über die Studienwahloptionen. – Alle anderen ‚Abnehmer‘ des Schulsystems habe [sic!] eigene und zusätzliche Kriterien und zunehmend auch eigene Prüfungen und Verfahren der Auswahl“ (a. a. O., S. 23).

63 *Avenarius/Füssel* (Anm. 23), S. 447. Aus diesem Grund muss die Kriteriumsorientierung bei der Notenerteilung auch vor allem bei Abschlussprüfungen eingehalten werden, „die für den weiteren Bildungs- und Berufsweg von ausschlaggebender Bedeutung sind“ (ebd.).

64 Art. 3 Abs. 1 GG gebietet es vor allem, wesentlich Gleiches gleich zu behandeln; vgl. exemplarisch *Detterbeck*, S., Öffentliches Recht. Ein Basislehrbuch zum Staatsrecht, Verwaltungsrecht und Europarecht, München 9. Aufl. 2013, Rdnr. 342 ff. Dabei müssen Personen(gruppen) miteinander verglichen werden, „die solche gemeinsamen Eigenschaften, Kriterien oder Oberbegriffe aufweisen, die die Personengruppe (trotz sonstiger Unterschiede) in bezug [sic!] auf das in Rede stehende Gesetz als im wesentlichen [sic!] gleich erscheinen lassen“ (a. a. O., Rdnr. 345). Eine solche taugliche Vergleichsgruppe ist vorliegend gegeben, wenn es sich um Schülerinnen und Schüler der gleichen Schulform und der gleichen Jahrgangsstufe bzw. des gleichen Bildungsgangs handelt, für die die gleichen Richtlinien und Lehrpläne etc. gelten. Eine Ungleichbehandlung von Personen dieser Vergleichsgruppe – etwa bei der Zensurenfindung bzw. -erteilung – wäre nur dann verfassungsrechtlich gerechtfertigt, „wenn zwischen ihnen Unterschiede von solcher Art und solchem Gewicht bestehen, daß sie die ungleiche Behandlung rechtfertigen können“ (BVerfG vom 30. Mai 1990 [Az. 1BvL2/83]). Dies ist nach allgemeiner Auffassung dann nicht der Fall, wenn die Ungleichbehandlung zugleich in andere Grundrechte eingreift (a. a. O., Rdnr. 347). Da jede Notengebung auch den Schutzbereich des Art. 12 GG tangiert (Ausbildungs- und Berufsfreiheit; vgl. oben Anm. 17), kann die Zugehörigkeit zu einer bestimmten Klasse und damit zu einer relativ kleinen Vergleichsgruppe oder der individuelle, von der Vergleichsgruppe gänz-

tierten Bezugsnorm gleichen Leistung erbracht haben, müssen sie auch gleich bewertet werden, um bei späteren Bewerbungen (etwa um einen Hochschul- oder Ausbildungsplatz) die insoweit gleichen Chancen zu haben. Das schließt eine Orientierung an einem sozialen bzw. individuellen Bewertungsmaßstab dem Grund nach aus – sie wäre insofern verfassungsrechtlich zu beanstanden, als objektiv gleichwertig erbrachte Leistungen unterschiedliche Teilhabemöglichkeiten im Erwerbsleben nach sich zögen.⁶⁵

4.2 Ausnahme: Leistungsbewertung nach Ermessen

Sogar in der juristischen Literatur wird angeführt, dass es im Schulalltag mehr als problematisch sein könne, bei der Notenvergabe allein auf Richtlinien und Lehrpläne als Bezugspunkt für eine kriteriumsorientierte Leistungsbewertung zu setzen, da nicht nur die Notendefinitionen selbst unbestimmte Rechtsbegriffe seien, sondern häufig auch der Grad des Erreichens bestimmter Anforderungen nur schwer zu operationalisieren sei.⁶⁶ Umso stärker wird das in der Schulpädagogik diskutiert: „Liest man amtliche Zensurendefinitionen, so erkennt man häufig den halbherzigen Versuch, Zensuren über solche sachliche Bezugsnormen zu bestimmen.“⁶⁷

Viele Bundesländer setzen daher eine Leistungsbewertung nach der kriteriumsorientierten Bezugsnorm zwar als Grundsatz, räumen den Lehrkräften aber zugleich die Möglichkeit ein, in begründeten Ausnahmefällen nach Ermessen von dieser Richtschnur abzuweichen. Gesetzessystematisch geschieht dies mit dem Wörtchen *soll*, das sich in den Bestimmungen zur Zensurenvergabe einiger Bundesländer findet, wenn es etwa heißt, dass eine bestimmte Ziffernnote bei einem gewissen Grad der Erfüllung der Leistungsanforderungen erteilt werden *soll* und eben nicht erteilt werden *muss* bzw. zu erteilen *ist*. Im Verwaltungsrecht, zu dem auch das Schulrecht gehört, wird im Zusammenhang mit sog. Soll-Bestimmungen von einem gebundenen bzw. eingeschränkten Ermessen ausgegangen: Die Behörde ist angehalten, den grundsätzlich beschriebenen Rechtsbefehl zu befolgen und nur in atypischen Ausnahmefällen von ihm abzuweichen.⁶⁸

lich entkoppelte Lernfortschritt einer einzelnen Schülerin oder eines einzelnen Schülers dem Grunde nach kein Kriterium für eine ungleiche Bewertung gleicher sachorientiert gemessener Leistungen sein; so auch *Hofmeyer* (Anm. 11), S. 36 f., nach dem unter dem Gesichtspunkt des Allgemeinen Gleichheitssatzes sicherzustellen ist, dass „objektiv gleiche bzw. gleichwertige Leistungen auch gleich beurteilt worden sind, mit der Konsequenz, daß das Gebot der Chancengleichheit eine pädagogische Bewertung nach dem subjektiven Leistungspotential unterbindet“.

65 Vgl. etwa BVerwG vom 9. August 1996 (Az. 6C3/95): „Prüfungsaufgaben müssen geeignet sein, Kandidaten, die das Ausbildungsziel erreicht haben, von denen zu unterscheiden, die es nicht erreicht haben. ... Der Grundsatz der Chancengleichheit im Prüfungsverfahren gebietet es, daß bei allen Prüflingen die gleichen Maßstäbe für die Bewertung einer Prüfungsleistung eingehalten werden“; so auch *Rux/Niehues* (Anm. 57), Rdnr. 501.

66 Vgl. *Rux/Niehues* (Anm. 57), Rdnr. 500, nach denen bei der „Antwort auf die Frage [sic!] ob eine Leistung ‚den Anforderungen in besonderem Maße entspricht‘ – und daher mit der Note ‚sehr gut‘ zu bewerten ist – oder ob sie den Anforderungen nur ‚voll‘ entspricht – und daher die Note ‚gut‘ verdient, ... den persönlichen Einstellungen und Erfahrungen der einzelnen Fachkraft entscheidende Bedeutung“ zukommt. Zur fehlenden begrifflichen Schärfe der amtlichen Notendefinitionen auch *Hofmeyer* (Anm. 11), S. 221 f.

67 *Rheinberg* (Anm. 5), S. 66, der fortführt: „Wollte man die Zensurengebung tatsächlich an sachliche Bezugsnormen knüpfen, so müsste man den beurteilenden Lehrern pro Fach, Jahrgangsstufe und Schulform sehr genau sagen, was jemand können muss, um ein ‚ausreichend‘ oder ein ‚gut‘ zu bekommen.“

68 Vgl. *Detterbeck* (Anm. 64), Rdnr. 686 m. w. N. der höchstrichterlichen Rechtsprechung des Bundesverwaltungsgerichts. *Soll*-Vorschriften, die sich hinsichtlich ihrer rechtlichen Qualität zwischen *Kann*- (sog. Ermessensvorschriften) und *Muss*-Vorschriften (sog. Vorschriften der gebundenen Verwaltungstätigkeit) bewegen, beinhalten umgangssprachlich mithin einen Normbefehl, der in dem Sinne umgesetzt werden *soll*, als er umgesetzt werden *muss*, wenn er umgesetzt werden *kann*.

Konkret bedeutet das, dass sich Lehrkräfte immer dann, wenn das landeseigene Schulrecht ein gewisses Ermessen bei der Anwendung von Bezugsmaßstäben bei der Leistungsbewertung einräumt, neben dem kriterienbezogenen Maßstab der Notenfindung auch an einer sozialen oder individuellen Bezugsnorm bei der Zensurenvergabe orientieren können, wenn dies aus sachlichen Gründen im Sinne eines atypischen Sonderfalls geboten ist, wobei insbesondere die eng verbundenen Aspekte der Teleologie der Vergabe von Leistungsnoten und die verfassungsmäßig gebotene Verpflichtung zur Gleichbehandlung die Anwendung anderer als der grundsätzlich gebotenen kriteriumsorientierten Bezugsnorm restringieren: So muss zum einen jede Abweichung der Zensurenvergabe vom grundsätzlich vorgegebenen Vergleichsmaßstab der in den Richtlinien, Lehrplänen etc. beschriebenen anzustrebenden Kompetenzerweiterungen der Schülerinnen und Schüler dem vordringlichen Sinn und Zweck der Notenerteilung zur Selektion und zur Erteilung von Berechtigungen Genüge leisten. Die pädagogische Funktion von Noten tritt insofern hinter der mit ihr konkurrierenden gesellschaftlichen zurück.⁶⁹ Zum anderen muss sich auch eine Leistungsbewertung, die sich neben der kriteriumsorientierten Bezugsnorm bewegt, am Maßstab des verfassungsrechtlich gebotenen Gleichheitssatzes messen lassen,⁷⁰ was – wie sogleich noch zu zeigen sein wird – nur schwer zu realisieren ist.

Was bleibt also für den Fall, dass das Landesschulrecht den Maßstab für die Leistungsbewertung zumindest bedingt in das Ermessen der Lehrkräfte stellt? Zunächst muss sie sich der Soll-Bestimmung zum Trotz stark an der Sachnorm der Lehrpläne und Richtlinien etc. orientieren und darf hiervon nur in eng umrissenen Ausnahmefällen abweichen – bezogen auf den kollektiven Maßstab etwa (nur) dann, wenn die curricularen Regularien (noch) keine Kompetenz- oder Lernzielbeschreibungen enthalten sollten; in Bezug auf den individuellen Maßstab unter anderem dann, wenn eine Schülerin bzw. ein Schüler bei der rechnerischen Ermittlung einer Gesamtnote zwischen zwei Notenstufen steht und zur Findung der Endnote ihr bzw. sein Leistungsfort- bzw. -rückschritt berücksichtigt werden soll. In jedem Fall muss sich die Leistungsbewertung aber an einem operationalisierbaren Bewertungsmaßstab orientieren und darf keinen sachfremden Erwägungen folgen, die im Zweifel nicht oder nur mittelbar mit der gezeigten Leistung der Schülerin bzw. des Schülers in Verbindung stehen (sog. Willkürverbot).⁷¹ Insgesamt rücken entsprechende, ermessenseinräumende Soll-Vorschriften damit deutlich in die Nähe von verpflichtenden Muss-Vorschriften.⁷²

69 Aus pädagogischer Sicht so auch *Jürgens/Lissmann* (Anm. 4), S. 165: „Weil Schulleistungskontrollen und -beurteilungen nicht ausschließlich der optimalen individuellen Förderung und der chancengleichen Gewährung des ‚Rechts auf Bildung‘ dienen, sondern ebenso der ‚Auslese befähigter Anwärter auf höhere Bildungslaufbahnen, begehrte Abschlüsse und angesehene berufliche und gesellschaftliche Positionen‘ (*Sacher* [W., Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe, Bad Heilbrunn 5. Aufl.] 2009, S. 22), wird Schule nolens volens zu einer Verteilungsagentur für Bildungslaufbahnen und Lebenschancen. Ihre genuin pädagogische Aufgabe, allen Schülerinnen und Schülern so viel Allgemein- und Persönlichkeitsbildung wie irgend möglich angedeihen zu lassen, gerät in Konkurrenz zu der parallel verlangten Funktion, nahezu durchgängig im gesamten Schulsystem Auswahlen, Zuteilungen und Klassifizierungen vorzunehmen oder Zugangsberechtigungen vergeben zu sollen. Das führt dazu, dass mit der Beurteilungsaufgabe in der Schule stets eine Doppelrolle zu erfüllen ist, indem sowohl der pädagogischen als auch gesellschaftlichen Funktion der Leistungsbeurteilung entsprochen werden soll.“

70 Vgl. oben Punkt 4.1.

71 Generell unzulässig wäre danach beispielsweise die Erteilung einer vergleichsweise privilegierenden Note wegen einer längeren Erkrankung oder des Todes eines Elternteils der Schülerin bzw. den Schülers (so aber *Hoegg, G.*, Schul-Recht! Von A bis Z, Weinheim und Basel 2012, S. 47) oder die Vergabe einer schlechteren Note zur Bestrafung bzw. Sanktionierung einer Schülerin bzw. eines Schülers (so auch *Jürgens/Sacher* [Anm. 1], S. 57), wobei in diesem Zusammenhang klar zwischen der erlaubten Motivierung/Disziplinierung durch Schulnoten und der unerlaubten Bestrafung/Sanktionierung durch Leistungsbewertung zu unterscheiden ist.

72 Dies gilt vor allem vor dem Hintergrund des allgemeinen Gleichbehandlungsgrundsatzes nach Art. 3 GG (vgl. hierzu oben Anm. 64), der das Ermessen behördlichen Handelns immens einschränkt und im Einzelfall sogar auf Null reduzieren kann; vgl. etwa *Detterbeck* (Anm. 64), Rdnr. 692 f.

4.3 Ausnahme: Leistungsbewertung nach der kollektiven Bezugsnorm

Bis zum Inkrafttreten des Hamburger Abkommens und dem daraufhin erlassenen Beschluss der Kultusministerkonferenz zur Erläuterung der Notenstufen⁷³ wurde die soziale Bezugsnorm bei der Zensurenvergabe vielfach als verbindlicher Bewertungsmaßstab vorgegeben.⁷⁴ Noch heute wird in einigen Bundesländern neben der kriteriumsorientierten Bezugsnorm bei der Leistungsbewertung auch eine Berücksichtigung der kollektiven Bezugsnorm gefordert bzw. in Form einer bindenden Verwaltungsvorschrift für die Notenfindung zwangsweise vorgegeben.⁷⁵ Dies entspricht zwar grundsätzlich der Sozialisations- bzw. Rückmeldungs- und Orientierungsfunktion von Noten, stößt aber dennoch auf sowohl pädagogische als auch juristische Kritik.

So wird in aller Regel insbesondere darauf hingewiesen, dass bei einer Anwendung der kollektiven Bezugsnorm die Zensurenvergabe zu sehr von der rein zufälligen Gruppen- bzw. Klassenzugehörigkeit abhängt, „denn während ein mittelmäßiger Schüler in einer schwachen Klasse gute Noten bekommt, muss er bei gleichen Leistungen in einer sehr guten Klasse vielleicht um die Versetzung bangen“⁷⁶. Bezogen auf eine größere Vergleichsgruppe (wie etwa alle Schülerinnen und Schüler eines Jahrgangs oder eines Bildungsgangs) verlieren die so ermittelten Noten stark an Aussagekraft und sind für einen klassenübergreifenden Leistungsvergleich gänzlich unbrauchbar.

Eine solche Leistungsbewertung steht indes auch der Evaluations- und Kontrollfunktion von Noten entgegen: Wird das Spektrum der erteilten Noten am Klassendurchschnitt ausgerichtet, geht jedwedes Gespür dafür verloren, ob der vor der Lernerfolgskontrolle erteilte Unterricht „gut“ oder „schlecht“ war; die Klasse erhält immer eine (annähernd) gleiche Durchschnittsnote – unabhängig davon, ob es die Lehrkraft verstand, die erwünschte Kompetenzerweiterung bei der Mehrzahl der Schülerinnen und Schüler zu erzielen, oder die Mehrheit der Klasse die standardisierten Anforderungen nicht erfüllt.⁷⁷ Besonders prekär ist dies beim Bestehen einer Quote, bei der eine

73 Vgl. Anm. 54 und 56.

74 Vgl. *Klauer* (Anm. 26), S. 44. Zwar wurden mit dem Düsseldorfer Abkommen bundesweit verbindliche Notenstufen festgelegt (vgl. Anm. 1), die Kultusministerkonferenz verzichtete aber zunächst darauf, sie einheitlich zu operationalisieren. In den landesspezifischen Schulgesetzen fanden sich entsprechend unterschiedliche Definitionen, die zum Teil an den Klassendurchschnitt anknüpfen. So galt in Baden-Württemberg eine Leistung als *befriedigend*, wenn sie einer *tüchtigen Leistung des guten Durchschnitts* entsprach, in Hamburg wurde eine Leistung mit *gut* bewertet, wenn sie *deutlich über dem Leistungsdurchschnitt der Klassenstufe lag*; vgl. *Lang, N.*, Sind Schwaben dümmter?, in: Die Zeit, Ausgabe 9/1968.

75 So in *Brandenburg*: „Die Leistungsbewertung bezieht sich auf die im Unterricht vermittelten Kenntnisse, Fähigkeiten und Fertigkeiten. Dabei [wird] ... der Leistungsstand der Lerngruppe ... berücksichtigt“ (§ 57 Abs. 2 Satz 1 und 2 BbgSchulG) sowie in *Rheinland-Pfalz*: „Leistungen werden nach dem Grad des Erreichens von Lernanforderungen beurteilt. Die Beurteilung berücksichtigt ... auch die Lerngruppe, in der die Leistung erbracht wird“ (§ 53 Abs. 1 ÜSchO).

76 *Wengert* (Anm. 3), S. 303; ähnlich *Rheinberg* (Anm. 5), S. 64, nach dem *Ingenkamp* (Anm. 2) bereits in den 1970er Jahren durch empirische Untersuchungen belegen konnte, dass die gleiche Leistung „mit ‚gut‘ oder ‚mangelhaft‘ beurteilt [wird], je nachdem, ob der Schüler in einer leistungsstarken oder leistungsschwachen Klasse ist“. Probleme ergeben sich nach *Breidenstein et al.* (Anm. 6), S. 18 f. sowie *Klauer* (Anm. 26), S. 42 dabei vor allem bei einem Klassenwechsel, wenn sich die Bezugsgruppe ändert und eine Schülerin bzw. ein Schüler dadurch (bei gleicher Leistung) Veränderungen in der Leistungsbewertung erfährt; hier kann auch nicht mehr davon ausgegangen werden, dass die so erteilten Noten eine Rückmeldungs- oder Orientierungsfunktion haben.

77 Hierzu bereits oben Anm. 33. Zu diesem Grunde nach pädagogischen Offenbarungseid auch *Jürgens/Sacher* (Anm. 1), S. 66, die die Gefahr darin sehen, dass Lehrende und Lehrer das Bewertungsschema auch zu ihren Gunsten nach unten verschieben könnten. „Durch eine derartige, gleichermaßen unpädagogisch [sic!] wie unprofessionelle Reaktion würde die konsequente Anwendung der sachlichen Bezugsnorm in Verbindung mit der Zensurierung unterlaufen: Tatsächlicher Misserfolg bezogen auf die Lehrplananforderungen würde durch diesen ‚Trick‘ auf dem Papier in einen ‚Erfolg‘ umgewertet.“

schriftliche Arbeit grundsätzlich wiederholt werden muss, wenn ein bestimmter Prozentsatz der Schülerinnen und Schüler nicht mindestens die Note 4 (*ausreichend*) erhalten hat.⁷⁸ „Diese Regelung drängt dem Lehrer im unteren Teil der Notenskala faktisch die soziale Bezugsnorm auf.“⁷⁹

Juristisch betrachtet dürfte es mit dem allgemeinen Gleichheitssatz des Art. 3 Abs. 1 GG nur schwer vereinbar sein, „dass die Zugehörigkeit zu dieser oder jener ... Klasse sich positiv oder negativ auf den Schulerfolg auswirkt“⁸⁰ – gleichwohl schreiben einige Bundesländer die Berücksichtigung der sozialen Bezugsnorm bei der Leistungsbewertung vor.⁸¹ Dieser Widerspruch kann nur dadurch aufgelöst werden, dass bei der Zensurenbildung von der sozialen Bezugsnorm nur äußerst restriktiv Gebrauch gemacht wird; so kann die soziale Bezugsnorm bestenfalls als Ergänzung zur kriteriumsorientierten herangezogen werden. Denkbar ist dies zum Beispiel bei Zensuren in niedrigen Jahrgangsstufen oder bei Einzelnoten, die keinen versetzerheblichen Charakter bzw. keine nennenswerten Auswirkungen auf die weitere Schullaufbahn haben. Mit zunehmender Bedeutung der Zensurenvergabe für die Schul- bzw. Berufskarriere muss bei der Leistungsbewertung der kollektive Maßstab zugunsten des kriteriumsorientierten immer mehr an Bedeutung verlieren; bei Abschlussprüfungen oder Noten, mit denen eine bestimmte Berechtigung verliehen wird, muss sich dessen Berücksichtigung sogar auf Null reduzieren.⁸²

4.4 Ausnahme: Leistungsbewertung nach der individuellen Bezugsnorm

In einigen Bundesländern ist bei der Notenfindung zusätzlich zur kriteriumsorientierten Bezugsnorm auch der individuelle Lernfortschritt der Schülerinnen und Schüler zu berücksichtigen (individuelle Bezugsnorm).⁸³ Damit wird der Erkenntnis Rechnung getragen, dass Leistung wesentlich prozessorientiert ist und demzufolge auch die Leistungsbewertung prozess- und nicht nur ergebnisorientiert erfolgen soll.⁸⁴

Als alleinige Beurteilungsperspektive würde die individuelle Bezugsnorm zu „irrationalen Konsequenzen“⁸⁵ führen – insbesondere dann, wenn an eine Leistungsbeurteilung Berechtigungen

78 Vgl. hierzu bereits oben Anm. 35.

79 *Rheinberg* (Anm. 5), S. 67, der weiter ausführt: „In leistungsschwachen Schulklassen bringt ihn [den Lehrer; d.V.] das zwangsläufig in Widerspruch zur Zensurendefinition, die den Anker ja bei den Anforderungen (sachliche Bezugsnorm) und nicht beim Klassendurchschnitt (soziale Bezugsnorm) vorsieht.“ Hierzu kritisch auch *Hofmeyer* (Anm. 11), S. 215: „Nicht zuletzt hat auch die Schulverwaltung mit fragwürdigen Verwaltungsvorschriften dazu beigetragen, wonach bspw. Klassenarbeiten bei Überschreiten gewisser Grenzen (z. B. 30 % ‚mangelhaft‘ oder ‚schlechter‘) nicht ohne weiteres gewertet werden dürfen“.

80 *Avenarius/Füssel* (Anm. 58), S. 124; weniger absolut *Avenarius/Füssel* (Anm. 23), S. 428: „Eine völlige Relativierung der Leistungsbewertung ist indes mit dem Gebot der Chancengleichheit (Art. 3 Abs. 1 GG) nicht vereinbar.“

81 Vgl. oben Anm. 75.

82 Vgl. hierzu bereits oben Anm. 63.

83 So etwa in *Berlin*: „Für die Leistungsbeurteilung maßgebend ist der nach Kriterien des Bildungsgangs festgestellte Entwicklungsstand der Kenntnisse, Kompetenzen, Fähigkeiten und Fertigkeiten der Schülerin oder des Schülers. Die individuelle Lernentwicklung ist zu berücksichtigen“ (§ 58 Abs. 5 Satz 2 und 3 SchulG), *Brandenburg*: „Die Leistungsbewertung bezieht sich auf die im Unterricht vermittelten Kenntnisse, Fähigkeiten und Fertigkeiten. Dabei [wird] ... die Lernentwicklung der Schülerin oder des Schülers berücksichtigt“ (§ 57 Abs. 2 Satz 1 und 2 BbgSchulG) sowie *Rheinland-Pfalz*: „Leistungen werden nach dem Grad des Erreichens von Lernanforderungen beurteilt. Die Beurteilung berücksichtigt den individuellen Lernfortschritt der Schülerinnen und Schüler“ (§ 53 Abs. 1 ÜSchO).

84 Vgl. *Jürgens* (Anm. 8), S. 31 f. sowie *Paradies et al.* (Anm. 32), S. 29, nach denen die traditionellen Formen der Leistungsbewertung häufig ausschließlich das Produkt im Auge haben; „die Beurteilung des (Lern-)Prozesses, der zu diesem Produkt geführt hat, bleibt außerhalb des Wahrnehmungshorizontes“.

85 *Rheinberg* (Anm. 5), S. 65.

geknüpft sind oder sie Auswirkungen auf den weiteren Bildungs- und Berufsweg einer Schülerin bzw. eines Schülers hat: „So würde der Schüler, der sich im letzten Jahr vom ‚mangelhaft‘ auf ‚ausreichend‘ hochgearbeitet hat, einen Studienplatz in einem anspruchsvollen Numerus Clausus Fach erhalten, der Schüler, der konstant bei ‚sehr gut‘ läge jedoch nicht.“⁸⁶ Von daher sind Beurteilungen unter individueller Bezugsnorm bestens geeignet, wenn es darum geht, möglichst veränderungssensible, detaillierte Rückmeldungen *innerhalb* eines Ausbildungsabschnittes zu geben, die über günstige Motivationsauswirkungen den Lernerfolg fördern. Bei Leistungsbeurteilungen dagegen, die dauerhafte Berechtigungen *außerhalb* dieses Ausbildungsabschnittes vermitteln (z. B. Studienplätze oder andere Zugangsberechtigungen), können individuelle Bezugsnormen keine nennenswerte Rolle spielen.⁸⁷

Diese „irrationalen Konsequenzen“ bedingen auch die verfassungsrechtlichen Bedenken gegen „die individuelle Bezugsnorm als alleinige Beurteilungsperspektive“⁸⁸; sie wäre vor dem Hintergrund des allgemeinen Gleichheitssatzes nach Art. 3 Abs. 1 GG ebenso unzulässig wie eine alleinige Bewertung nach der sozialen Bezugsnorm – mehr noch: Während die Anwendung des kollektiven Bewertungsmaßstabes einen (wenn auch immens eingeschränkten) interpersonellen Vergleich ermöglicht, ist dieser bei einer Zugrundelegung der individuellen Bezugsnorm bei der Notenvergabe gänzlich ausgeschlossen; insofern scheidet sie als Maßstab bei der Zensurenvergabe vor dem Hintergrund der Chancengleichheit als juristischem Moment der Berechtigungs- und Selektionsfunktion von Noten dem Grunde nach aus.

Sofern schulrechtliche Bestimmungen dennoch auf sie rekurren, gelten die gleichen Grundsätze wie bei der Anwendung der kollektiven Bezugsnorm.⁸⁹ Der kriteriumsorientierte Maßstab muss das ausschlaggebende Moment bei der Notenfindung sein, eine pädagogische Berücksichtigung des individuellen Lernfortschritts muss deutlich dahinter zurückstehen und kann bestenfalls dann erfolgen, wenn die Note allein gemessen am Grad der in Bezug auf die curricularen Anforderungen erworbenen Kompetenzen nicht sicher ermittelt werden kann.⁹⁰

5 Resümee

Unter pädagogischen Gesichtspunkten können bei der Zensurenvergabe alle drei möglichen Bezugsnormen ihre Berechtigung haben – je nachdem, welcher Zweck mit der Notengebung verfolgt wird: Geht es in erster Linie um die Sozialisation der Schülerinnen und Schüler für die bestehende Leistungsgesellschaft oder darum, ihnen Auskunft über ihren Leistungsstand innerhalb einer Referenzgruppe zu geben, macht eine Orientierung an der kollektiven Bezugsnorm Sinn. Werden beispielsweise bloße Übungsaufgaben bewertet oder dient die Benotung der Dokumentation des individuellen Lernfortschritts, kann die Zensurenvergabe nach der individuellen Be-

86 Nach Hofmeyer (Anm. 11), S. 219 kann durch eine solche Verfahrensweise bei der Leistungsbewertung „in letzter Konsequenz die Funktionsfähigkeit des schulischen Selektions- und Allokationsprozesses außer Kraft“ gesetzt werden.

87 Rheinberg (Anm. 5), S. 65 f.

88 Rheinberg (Anm. 5), S. 65.

89 Vgl. soeben unter Punkt 4.3.

90 Ähnlich Rheinberg (Anm. 5), S. 66: „Man könnte sie allenfalls als ‚letzten Trend‘ mit berücksichtigen, der zu einer leichten Auf- oder Abwertung einer anders vorgenommenen Beurteilung führt.“ Ähnliches kann im Einzelfall bei einer „Ausreißernote“ gelten, wenn die Lehrkraft im Rahmen der Gesamtbewertung einer Schülerleistung pädagogisch begründet zu der Auffassung kommt, dass die errechnete Note den Leistungsstand der Schülerin bzw. des Schülers gemessen an den Anforderungen nicht angemessen wiedergibt.

zugsnorm das Mittel der Wahl sein, da sie am stärksten motivationsfördernd wirkt. Immer dann, wenn es (auch) um die Berechtigungsfunktion von Noten geht, scheint die Heranziehung der kriteriumsorientierten Bezugsnorm unverzichtbar. Gute Lehrkräfte wissen daher um die Funktionen von Noten und um die möglichen Bezugsnormen bei der Notenvergabe, berücksichtigen sie entsprechend und wenden sie situationsadäquat an.⁹¹ Der generellen Informationsfunktion von Noten wäre indes dann am besten Rechnung getragen, wenn alle drei Bezugsnormen berücksichtigt würden und Lehrerinnen und Lehrer drei Noten gemäß den drei Bezugsnormen erteilen dürften.⁹²

Aus schulrechtlicher Perspektive ergibt sich indes eine andere Einschätzung: Die jeweils einschlägigen Vorschriften aller Bundesländer statuieren ein Primat der kriteriumsorientierten Bezugsnorm bei der Zensurenvergabe.⁹³ Dies geschieht zum Teil – wie etwa in Bayern oder Sachsen – in Form eines generellen Normbefehls, der keine Ausnahmen zulässt; in diesem Fall *muss* die Notenfindung kriteriumsorientiert erfolgen. In anderen Bundesländern werden hiervon mehr oder weniger weitreichende Ausnahmen gemacht, auch wenn die Berücksichtigung der kriteriumsorientierten Bezugsnorm der Grundsatz bleibt. So gibt es die Möglichkeit, in begründeten Ausnahmefällen nach (gebundenem) Ermessen von dieser Richtschnur abzuweichen und/oder neben der kriteriumsorientierten Bezugsnorm auch die individuelle und/oder die soziale Bezugsnorm bei der Zensurenvergabe zu berücksichtigen (vgl. Tab. 1).

Dabei ist Folgendes zu beachten: Sofern Lehrkräften bei der Zensurenerteilung vom geltenden Schulrecht ein gewisser Ermessens- bzw. Handlungsspielraum eingeräumt wird, ist eine Abweichung von der kriteriumsorientierten Bezugsnorm zwar grundsätzlich möglich, sie darf aber nicht willkürlich sein, sondern muss sich an vernünftigen und sachgerechten Erwägungen orientieren.⁹⁴ Zu berücksichtigen sind ferner die grundrechtlichen Positionen der zu bewertenden Schülerinnen und Schüler sowie ihrer Mitlernenden, die sich aus Art. 3 und 12 GG ergeben. Vor diesem Hintergrund ist von anderen als der kriteriumsorientierten Bezugsnorm umso restriktiver Gebrauch zu machen, je mehr die zu vergebende Zensur den Rahmen einer rein pädagogisch motivierten Leistungsbeurteilung verlässt und sie bei der Erteilung von Berechtigungen oder Zertifikaten zu berücksichtigen ist. Bei versetzungserheblichen Noten, spätestens aber bei der Vergabe von Noten im Rahmen einer (zentralen) Abschlussprüfung reduziert sich das eventuell eingeräumte Ermessen zur Berücksichtigung einer nicht-kriteriumsorientierten Bezugsnorm auf Null.⁹⁵ Die in diesem Sinne vermeintlich eingeräumte pädagogische Freiheit ist dann bestenfalls noch pädagogische Poesie, die keinen Niederschlag in der schulverwaltungsrechtlichen Praxis finden darf – und das vor dem Hintergrund der immensen Bedeutung der Berechtigungs- und Selektionsfunktion der Zensurenvergabe und des Bildungswesens völlig zu Recht.

91 So auch *Klawer* (Anm. 26), S. 44: „Man muss also davon ausgehen, dass heute alle drei Bezugsnormen faktisch in Gebrauch sind, wobei der einzelne Lehrer im allgemeinen vielleicht eine bestimmte bevorzugen wird. Es ist nicht einmal auszuschließen, dass manche Lehrer die Bezugsnorm in Abhängigkeit von den Bedingungen des Einzelfalls wechseln, um jedem Schüler gerecht werden zu können.“

92 Hierzu bereits oben Anm. 27. *Richter*; I., Die sieben Todsünden der Bildungspolitik, München 1999, 121 und 129 ff. schlägt gar die Erteilung von vier Zensuren vor, die an einem absoluten („Was kann ich?“), einem relativen („Wo stehe ich?“), einem retrospektiven („was habe ich geschafft?“) und einem prognostischen Leistungsmaßstab („Was schaffe ich in Zukunft?“) gemessen werden sollen (a. a. O., S. 139).

93 Hierzu zusammenfassend *Hofmeyer* (Anm. 11), S. 227: „Die rechtlich zwingende Notwendigkeit kriteriumsorientierter Bewertung ergibt sich aus der Tatsache, dass die Gesellschaft der Schule Allokations- und Selektionsentscheidungen abverlangt.“

94 Vgl. *Detterbeck* (Anm. 64), Rdnr. 346 und 358.

95 Hierzu *Detterbeck* (Anm. 64), Rdnr. 693.

Kritisch zu betrachten sind daher auch die in vielen Bundesländern (noch) bestehenden sog. Drittel-Erlasse, deren Geltung Lehrkräfte dazu verleiten könnte, sich von der Anwendung der kriteriumsorientierten Bezugsnorm abzuwenden und – dem Grunde nach contra legem – zu stark die kollektive Bezugsnorm zu berücksichtigen. An dieser Stelle kann an Schulleiterinnen und Schulleiter nur appelliert werden, sich ihrer Verantwortung für die gesellschaftliche Funktion von Schule und der gegebenen schul- und verfassungsrechtlichen Vorgaben bewusst zu werden und auch Notenspiegel zu akzeptieren, die zwar nicht der Gauß'schen Normalverteilung entsprechen, dafür aber ermessens- und beurteilungsfehlerfrei die kriteriumsorientierte Bezugsnorm bei der Notenvergabe berücksichtigen – es gibt nun einmal „gute“ und „schlechte“ Klassen!

Insgesamt kann also die Frage, ob bei der Leistungsbewertung eine Diskrepanz zwischen pädagogischen Wünschen und schulrechtlichen Möglichkeiten besteht, nur in Abhängigkeit davon beantwortet werden, in welchem Bundesland die Note erteilt werden soll, welchem Zweck die Zensurenvergabe dient und vor allem welche Bezugsnorm angewendet wird. Je stärker sich Lehrkräfte bei der Zensurenvergabe an der kriteriumsorientierten Bezugsnorm ausrichten, umso sicherer stehen sie auf dem Boden des geltenden Schulrechts.

Verf.: Prof. Dr. Peter Kostorz, Fachhochschule Münster, Fachbereich Gesundheit, Leonardo-Campus 8, 48149 Münster, E-Mail: kostorz@fh-muenster.de