

Allgemeiner Teil

Britta Kohler/Thorsten Bohl/Nicolas Hübner

Sechzehn Länder, sechzehn Bestimmungen?

Aktuelle rechtliche Regelungen zur Benotung schriftlicher und mündlicher Leistungen an Gymnasien

Zusammenfassung: Vor dem Hintergrund der großen Relevanz schulischer Zensuren für Lebens- und Berufschancen einerseits und der vielfältigen wissenschaftlichen Kritik an ihrer Aussagekraft und Vergleichbarkeit andererseits arbeitet der vorliegende Beitrag aktuelle rechtliche Regelungen der Notengebung in Deutschland vergleichend auf. Hierbei erfolgt eine Fokussierung auf schriftliche und mündliche Leistungen an allgemeinbildenden Gymnasien. Mittels Dokumentenanalyse wird insbesondere eruiert und anschließend diskutiert, welche Leistungen die Grundlage der Fachnoten in den Bundesländern bilden, wie diese bestimmt und gewichtet werden und welche Anforderungen und Kriterien für die Bildung von Fachnoten gültig sind. Die Befunde verweisen bezüglich der rechtlichen Regelungen auf enorme Differenzen zwischen den Ländern, aber auch auf fehlende Spezifizierungen innerhalb von Ländern und variierende Begrifflichkeiten.

Schlagnvorte: Noten, Leistungen, Gymnasium, rechtliche Regelungen, Bundesländer

1. Einleitung

Noten bilden eine zentrale Informationsgrundlage bei Übergangsentscheidungen, beruflichen Auswahlprozessen oder der Zulassung zu Studiengängen (vgl. Hübner, Wagner, Hochweber, Neumann & Nagengast, 2020; Hübner, Jansen, Stanat, Bohl & Wagner, 2024). Diese hoch relevanten notenbasierten Auswahlentscheidungen im Bildungssystem fußen wesentlich auf der Idee, Noten seien über Schüler:innen aus unterschiedlichen Klassen und Schulen sowie über die Grenzen einzelner Bundesländer hinweg vergleichbar. Hierzu gibt es, spätestens seit der Publikation Karlheinz Ingenkamps aus dem Jahr 1971 zur „Fragwürdigkeit der Zensurengebung“, eine breite Debatte und zahlreiche wissenschaftliche Beiträge, die auf die unklare Aussagekraft und mangelnde Vergleichbarkeit von Noten hinweisen (z. B. Beutel & Pant, 2020; Bonefeld & Dickhäuser, 2018; Holder & Kessels, 2017; Hübner et al., 2022). Eine Zuspitzung erfährt die Thematik in der Sekundarstufe II und beim Abitur, welches im nationalen Bildungssystem nicht nur einen Abschluss, sondern zugleich eine Hochschulzugangsberechtigung im

Kontext vielfacher Zulassungsbeschränkungen darstellt. Statistiken der Kultusministerkonferenz (KMK, 2024) zeigen, dass sich die Quoten von Einserabschlüssen beim Abitur in den verschiedenen Bundesländern substantiell unterscheiden: So erreichte beispielsweise im Jahr 2023 in sechs Bundesländern mehr als ein Drittel der Abiturient:innen einen Durchschnitt von 1,0 bis 1,9, während dies in Schleswig-Holstein lediglich 22.8 % der Absolvent:innen gelang (KMK, 2024). Verschiedene Studien legen nahe, dass derartige Unterschiede am Ende der Sekundarstufe II, aber auch am Ende der Sekundarstufe I nicht mit objektiven Testleistungsunterschieden übereinstimmen, d. h. Schüler:innen mit identischen Noten in unterschiedlichen Bundesländern über substantiell unterschiedliche Kompetenzen verfügen (z. B. Baumert, Trautwein & Artelt, 2003; Neumann, Nagy, Trautwein & Köller, 2009; Klieme, 2003; Hoffmann, Schröter & Stanat, 2022). Damit berührt die skizzierte Problematik nicht nur erziehungswissenschaftliche Diskussionen, sondern ist mit Blick auf die Lebens- und Berufschancen der jungen Menschen auch von rechtlicher Relevanz. Nicht ohne Grund stellte das Bundesverfassungsgericht in einem diesbezüglichen Urteil vom 19. Dezember 2017 eine „eingeschränkte länderübergreifende Vergleichbarkeit“ (BVerfG, 2017) der Abiturnoten und damit eine Benachteiligung von Abiturient:innen aus bestimmten Bundesländern fest, worauf die Kultusministerkonferenz unter anderem mit einer stark veränderten Fassung der bereits seit 1972 bestehenden „Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung“ (KMK, 2023) im Jahr 2023 antwortete. Diese Vereinbarung zur „Sicherung der Vergleichbarkeit der Abiturergebnisse unter den Ländern“ (KMK, 2023, S. 5) beinhaltet beispielsweise eine Vereinheitlichung der Zahl der Klausuren sowie der Gewichtung der Kursleistungen. Aufgrund der Kulturhoheit der Länder sind die Vorgaben rechtlich nicht bindend und müssen daher von den Ländern sukzessive in entsprechende länderspezifische Regelungen überführt werden.

Festzuhalten ist, dass sich die Frage der Aussagekraft und Vergleichbarkeit von Noten am Schnittpunkt von Erziehungswissenschaft und (Schul-)Recht bewegt. Dabei liegen aus der Erziehungswissenschaft und affinen Disziplinen zahlreiche Forschungsergebnisse zur Thematik vor. So gibt es Studien, die Schulnoten als valide Prädiktoren für hochschulische Noten konturieren (z. B. Trapmann, Hell, Weigand & Schuler, 2007) bzw. sie allgemein als valide begreifen (z. B. Spinath & Brünken, 2016), und auch Beiträge, die große Zweifel an ihrer Aussagekraft nahelegen (vgl. Beutel & Pant, 2020). Publikationen hingegen, die aktuelle rechtliche Vorgaben zur Vergabe von Schulnoten systematisch aufarbeiten und vergleichend analysieren, sind kaum aufzufinden, obwohl diese Vorgaben einen bedeutsamen und bindenden Rahmen schulischer Notengebung bilden. Wie Beiträge und Modelle zur Schule als Mehrebenensystem (z. B. Ditton, 2017) und zum akteurzentrierten Institutionalismus (Fend, 2008, S. 157) zeigen, sind Bedingungen auf der Systemebene grundsätzlich relevant für institutionelles Handeln der Akteur:innen auf der Mikroebene, wird dieses also immer auch grundgelegt und limitiert von den bestehenden Voraussetzungen und Strukturen (vgl. auch Fend, 2008). Für diese theoretischen Überlegungen finden sich auch empirische Hinweise: So konnten beispielsweise in Studien zur Einführung des Zentralabiturs Veränderungen im Handeln von Lehrkräften als Folge der veränderten Regelungen gefunden werden. Mathe-

matiklehrkräfte zeigten nach der Reform eine verringerte Orientierung an der sozialen Bezugsnorm bei der Leistungsbeurteilung, was sich in einem reduzierten Referenzgruppeneffekt der mittleren Klassenleistung auf die Note nach der Reform ausdrückte. In Englisch fielen die Ergebnisse hingegen uneinheitlich aus, was nahelegt, dass Regelungen auf der Makroebene prinzipiell das Potenzial besitzen, die Handlungspraktiken von Lehrkräften zu verändern, allerdings fach- und leistungsgruppenspezifisch unterschiedlich ausfallen können (vgl. Neumann, Trautwein & Nagy, 2011).

Darüber hinaus zeigte sich in einer Studie von Hübner et al. (2020), dass eine strukturelle Veränderung des Oberstufensystems (von Leistungs- und Grundkursen zu Kernfächern) auch mit substanziellen Veränderungen der Benotung von Schüler:innenleistungen einherging. So wiesen beispielsweise Schülerinnen und Schüler in Leistungskursen vor der Strukturreform (mit identischen Noten) im Mittel substanziell höhere Leistungen auf als Schülerinnen und Schüler in Kernfächern nach der Reform. Ebenso zeigte sich, dass Lernende in Kernfächern nach der Reform substanziell höhere Testleistungen aufwiesen als Lernende in Grundkursen vor der Reform (bei identischen Noten). Zusammenfassend legen diese Befunde nahe, dass Veränderungen struktureller Regelungen auf der Makroebene durchaus mit Veränderungen auf der Mikroebene und der konkreten Benotungspraxis einhergehen können und für ein besseres Verständnis der mangelnden Vergleichbarkeit von Schulnoten stärker berücksichtigt werden sollten.

Der folgende Beitrag nimmt die beschriebene Forschungslücke bezüglich der schulischen Zensurengebung zum Anlass, um die länderspezifischen rechtlichen Regelungen für die Vergabe von Noten näher in den Blick zu nehmen. Da eine umfassende Darstellung und Analyse aller Regelungen in allen 16 Bundesländern den Rahmen eines Beitrags übersteigen würde, erfolgt hier eine bewusste Fokussierung auf allgemeinbildende Gymnasien. Diese stellen nicht nur die in der Sekundarstufe I am häufigsten gewählte Schulart dar (Autor:innengruppe Bildungsberichterstattung, 2022), sondern auch jene Schulart, bei der die Aussagekraft, die Vergleichbarkeit und die Konsequenzen von Zensuren – vor allem in den Abschlusszeugnissen mit Blick auf zulassungsbeschränkte Studiengänge – besonders kontrovers diskutiert werden. Außerdem soll hier ein Fokus auf die in den meisten Schulfächern vergebenen schriftlichen und mündlichen Noten gelegt werden, wobei Spezifika einzelner Fächer vernachlässigt werden. Zentral erscheinen bei der nachfolgenden Analyse mit Blick auf die Grundproblematik der fraglichen Aussagekraft und Vergleichbarkeit von Zensuren die Kernfragen, welche Leistungen in den einzelnen Bundesländern Grundlage der Fachnoten sind, wie diese in den Ländern gewichtet werden und welche Anforderungen und Kriterien bei der Notengebung jeweils gelten sollen.

2. Zur Benotung schriftlicher und mündlicher Leistungen

2.1 Arten und Vergabe von Noten

Die Güte fachlicher Leistungen, insbesondere bei summativen Leistungsstanderhebungen, wird an Schulen überwiegend in Noten ausgedrückt. In Deutschland bewegen diese sich zumeist zwischen den Ziffern Eins (sehr gut) und Sechs (ungenügend). Für die gymnasiale Oberstufe liegt zudem ein Punktesystem vor, welches von 0 bis 15 Punkten reicht (KMK, 2023). Im internationalen Raum finden sich noch weitere Skalierungen mit unterschiedlich vielen Abstufungen und verschieden benannten Polen. Im Allgemeinen steht im Zeugnis eine summarische Note pro Fach, Teilbereich oder Fächerverbund. Darüber hinaus gibt es beispielsweise in Alternativschulen und auch im internationalen Raum Zeugnisse ohne Noten (z. B. Falkenberg, Vogt & Waldow, 2017), die derzeit aber weniger an Gymnasien und mehr an anderen Schularten eingesetzt und weiterentwickelt werden (Beutel & Pant, 2020).

Die Vergabe von Noten ist ein verpflichtender Teil der Arbeit von Lehrkräften und ist eingebettet in den Aufgabenbereich des Beurteilens von Schüler:innenleistungen. Bundesweit besteht eine Übereinkunft darin, dass dieser Aufgabenbereich von hoher Relevanz ist: Das Beurteilen stellt einen der insgesamt vier Kompetenzbereiche – Unterrichten, Erziehen, Beurteilen und Innovieren – in den KMK-Standards für die Lehrerbildung dar, „die für die Ausbildung und den Berufsalltag von Lehrkräften von besonderer Bedeutung sind“ (KMK, 2022, S. 2) und schulische Bildung sichern und weiterentwickeln sollen. Die Kompetenz des Beurteilens besteht dabei aus zwei Teilkompetenzen, einer diagnostischen Kompetenz und einer Beurteilungskompetenz im engeren Sinne (KMK, 2022, S. 2). Die zugehörigen Standards bieten Anschlüsse an die Fachdiskussion, verbleiben aber auf einem für die praktische Umsetzung sehr allgemeinen Niveau: Hinweise zu konkreten Kriterien werden beispielsweise kaum expliziert; auch eine Unterscheidung von schriftlichen und mündlichen Leistungen findet nicht statt.

Lehrkräfte fühlen sich Studien zufolge häufig nicht angemessen auf das Beurteilen von Schüler:innenleistungen vorbereitet (vgl. Terhart, 2011). In qualitativen Studien wird deutlich, wie Lehrpersonen angesichts der Anforderungen ihnen praktikabel erscheinende Lösungen zu entwickeln versuchen. So konnte bereits mehrfach herausgearbeitet werden, dass Lehrkräfte „Objektivierungspraktiken“ (Meister & Hollstein, 2018, S. 124) anwenden, indem sie beispielsweise möglichst viele Einzelnoten erzeugen, hinter denen die Subjektivität des Beurteilungsprozesses (vermeintlich) verschwindet (vgl. Bräu & Fuhrmann, 2015; Zabarovski, Meier & Breidenstein, 2011).

2.2 Noten für schriftliche und mündliche Leistungen

In alltäglichen Diskussionen, aber auch in Lehr- und Handbüchern sowie empirischen Studien zur Notengebung in der Schule sind mit Noten überwiegend schriftliche Noten gemeint, d. h. Ziffernzensuren, die die Güte schriftlicher Arbeiten im Sinne von Klau-

suren in numerischer Form umschreiben, oder Zeugnisnoten, die Konglomerate von Bewertungen unterschiedlicher Leistungen darstellen (z.B. Harks, Rakoczy, Klieme, Hattie & Besser, 2014; Lintorf & Buch, 2021; Sacher, 2014). Zur Aussagekraft und Vergleichbarkeit dieser Noten liegen zahlreiche empirische Studien vor, die beispielsweise erzielte Fachnoten im Sinne von Konglomeraten mit Leistungen in standardisierten Tests in Beziehung setzen und überwiegend nur mittlere Zusammenhänge finden (Bowers, 2011; Hübner et al. 2022; Hübner et al., 2024). Zu präzisieren ist an dieser Stelle, dass Zusammenhänge zwischen $r = .40\text{--}.50$ (vgl. Hübner et al. 2024) zwischen Noten und Ergebnissen aus Leistungstests in Sozial- und Bildungswissenschaftlichen Studien typischerweise als bedeutsam eingeschätzt werden und nach Cohen (1988) im mittleren Bereich liegen. Gleichzeitig drücken derartige Zusammenhänge aus, dass lediglich 16 % – 25 % der Unterschiede in den Noten durch Testleistungen erklärt werden können, was in Anbetracht der großen Bedeutung von Schulnoten für individuelle Bildungswege eher als gering beurteilt werden kann.

Mit Blick auf die zuvor skizzierte Unterscheidung in den KMK-Standards (KMK, 2022) kann zudem festgestellt werden, dass sich hier zum einen Untersuchungen finden lassen, die die diagnostische Kompetenz von Lehrkräften im Sinne einer möglichst akkuraten Einschätzung lernrelevanter Merkmale von Schüler:innen in den Blick nehmen (vgl. z.B. Südkamp, Kaiser & Möller, 2012) – bzw. das Konstrukt der diagnostischen Kompetenz grundsätzlich in Frage stellen (Spinath, 2005) –, und zum anderen Studien, die auf die Notenfindung fokussieren und hier enorme Differenzen zwischen Lehrpersonen finden (z.B. Birkel, 2005; Schröter et al., 2022).

Mündliche Leistungen bzw. mündliche Noten geraten deutlich seltener oder nur als kurz abgehandelter „Spezialfall“ (Jürgens & Sacher, 2008, S. 116) in den Fokus von Forschungsvorhaben oder von Kompendien für (angehende) Lehrpersonen (z.B. Hosenfeld & Schrader, 2006; Maier, 2015; Sacher, 2014). So ist auch über ihr Zustandekommen und ihre Relevanz für die Gesamtnote wenig bekannt. In einer qualitativen Untersuchung mit 21 Lehrkräften aus der Schweiz zeigte sich, dass die Zeugnisnote in Mathematik überwiegend „aus der Durchschnittsnote der schriftlichen Prüfungen unter geringfügiger Berücksichtigung der mündlichen Beteiligung im Unterricht“ (Weich & Stalder, 2019, S. 4) berechnet wird. Vielfach wird auch berichtet, mündliche Noten würden negativ bestimmt und würden alles umfassen, was nicht den Klassenarbeiten zugerechnet werden kann (z.B. Krieger, 2003). Sofern die Benotung mündlicher Leistungen außerhalb des deutschsprachigen Raums diskutiert wird, geht es vornehmlich um punktuell stattfindende mündliche Prüfungen und weniger um die Beurteilung kontinuierlich erbrachter mündlicher Beiträge im alltäglichen Unterricht (z.B. Herzmann & Liegmann, 2020; Sparfeldt, Rost, Baumeister & Christ, 2013). Internationale Studien beziehen sich außerdem häufig auf Studierende an Hochschulen und nicht auf Schüler:innen (z.B. Goodman, 2020; Schürmann, Kärner & Ringeisen, 2022). Insgesamt kann festgestellt werden, dass die Vergabe und Rezeption mündlicher Schulnoten theoretisch und empirisch kaum untersucht sind. Darüber hinaus ist unklar, wo genau die Grenze zwischen Mündlichem und Schriftlichem verläuft. Im praktischen Vollzug besteht für die Beurteilung mündlicher Beiträge die Herausforderung der Flüchtigkeit des

gesprochenen Wortes und auch das Phänomen, dass die Beiträge einzelner Lernender im Unterricht oftmals nur schwer voneinander abgegrenzt werden können. Überdies erzeugen Lehrkräfte über die Zuweisung unterschiedlich schwerer Fragen an verschiedene Lernende und abgestufte Hinweise jene Leistung mit, die sie anschließend beurteilen (vgl. z.B. Bräu & Fuhrmann, 2015; Kalthoff, 2000; Krelle, 2014; Zabarowski et al., 2011). Qualitative und quantitative Studien zeigen weiterhin, dass ein regelkonformes Verhalten und damit ein nicht leistungsbezogenes Kriterium für die Benotung fachlicher Leistungen relevant ist (Krämer & Zimmermann, 2020; Meister & Hollstein, 2018).

2.3 Arten von und Umgang mit rechtlichen Vorgaben zur Beurteilung schriftlicher und mündlicher Leistungen

Aufgrund der Kulturhoheit der Länder liegen in Deutschland 16 verschiedene Schulgesetze und zahlreiche weitere rechtliche Vorgaben für den schulischen Kontext vor. Sie bilden einen länderspezifischen verbindlichen Ordnungsrahmen für die schulische Arbeit und damit auch die Vergabe von Noten und sollen Rechtsstaatlichkeit im Bildungswesen gewährleisten (Avenarius, 2019). Gleichzeitig stellen sie Instrumente der Qualitätssicherung dar (Fend, 2008). Der geltenden Normenhierarchie folgend darf keine Regelung einer ranghöheren Norm widersprechen. An oberster Stelle stehen Gesetze, die von Parlamenten im Rahmen demokratischer Entscheidungsprozesse verabschiedet werden. Ihnen folgen Rechtsverordnungen, die von den Kultusministerien erlassen werden und von gleicher Verbindlichkeit wie gesetzliche Regelungen sind, und schließlich behördliche Erlasse. Grundsätzlich werden auf höherer Ebene vorwiegend allgemeine und grundlegende Fragen entschieden, während (zusätzliche) Konkretisierungen, insbesondere für einzelne Schularten, überwiegend von den zuständigen Ministerien und Behörden geleistet werden.

In der Literatur zur Frage der Leistungsbeurteilung wird in der Regel nicht auf rechtliche Vorgaben Bezug genommen, was kein Spezifikum dieses Themenbereichs zu sein scheint: Für Füßel (2020) sind Fragen des Rechts generell ein „blinder Fleck in der Lehrerinnen- und Lehrerbildung“ (S. 114) und vielfach wird sowohl mit Blick auf die Erziehungswissenschaft als auch die Schulpraxis von einer distanten Haltung rechtlichen Regelungen gegenüber ausgegangen (vgl. z.B. Fend, 2008; Helbig & Nikolai, 2015; Salgo, 2021). Prominente und breit angelegte Bestrebungen, „den weitgehend brachliegenden Diskurs zwischen Erziehungswissenschaft und Bildungsrecht zu revitalisieren“ (Wiezorek & Reimer, 2021, S. 32), sind selten. Benannt werden kann hier aus neuerer Zeit insbesondere ein Band der Deutschen Gesellschaft für Erziehungswissenschaft mit dem Titel „Erziehungswissenschaft und Bildungsrecht – Anknüpfungspunkte zur Belebung eines brachliegenden Diskurses“. Hugo und Heinrich (2021) sprechen darin von einer „seltsam anmutende[n] Abstinenz im Umgang mit Bildungsrecht, die der faktischen Relevanz von rechtlichen Rahmenbedingungen für die ‚Entwicklung des Bildungswesens‘ zuwider läuft“ (S. 11). Auch wird in diesem Band darauf hingewiesen, Einzelschulen und Lehrkräfte könnten davon profitieren, wenn sie über Kennt-

nisse schulrechtlicher Bestimmungen verfügten, die ihr Handeln legitimieren und auch limitieren und dabei Sicherheit und Verlässlichkeit schaffen (Salgo, 2021). Gerade bei der Zensurengebung und damit möglicherweise der Zuteilung von Lebens- und Berufschancen erscheint dies von hoher Relevanz. Unklar ist hierbei, ob fehlende Bezugnahmen auch mit unterschiedlichen Zugängen und einem unterschiedlichen Vokabular von Erziehungs- und Rechtswissenschaft zu tun haben.

3. Darstellung und Analyse rechtlicher Regelungen bezüglich der Benotung schriftlicher und mündlicher Leistungen an Gymnasien in den verschiedenen Bundesländern

3.1 Fragestellungen

Eine Darstellung und Analyse rechtlicher Regelungen bezüglich der Benotung schriftlicher und mündlicher Leistungen erscheint mit Blick auf deren Relevanz einerseits und die bislang fehlenden Auseinandersetzungen andererseits überaus bedeutsam. In Anbetracht der Kulturhoheit der Länder ist dieses Vorhaben auch bei einer Begrenzung auf die Notengebung an allgemeinbildenden Gymnasien ein aufwendiges Unterfangen, das im Rahmen eines Einzelbeitrags nicht vollständig abbildbar ist. Im Folgenden werden die Vorschriften zur Benotung schriftlicher und mündlicher Leistungen an Gymnasien aus den 16 Bundesländern entsprechend auszugsweise dargestellt und vergleichend analysiert. Mit Blick auf die weiter oben skizzierte Problematik der hohen Relevanz sowie mangelnden Aussagekraft und Vergleichbarkeit von Zensuren einerseits, die hohen Anforderungen an Lehrpersonen beim Beurteilen andererseits und auch die fehlenden gegenseitigen Bezugnahmen von Wissenschaft und Recht fokussiert der vorliegende Beitrag insbesondere die Fragen, welcher rechtliche Rahmen durch die Vorgaben in den verschiedenen Bundesländern abgesteckt wird und ob dieser anschlussfähig an fachliche Diskussionen erscheint. Im Einzelnen soll folgenden Fragestellungen nachgegangen werden:

- 1) Welche Leistungen bilden die Grundlage der Fachnoten an Gymnasien in den 16 Bundesländern?
- 2) Wie werden die verschiedenen Leistungen bei der Bildung der Fachnoten gewichtet?
- 3) Welche Anforderungen und Kriterien sollen bei der Bildung der Noten gelten?

3.2 Methode

Zunächst erfolgte eine Recherche zu zentralen rechtlichen Vorgaben bezüglich der Notenbildung an Gymnasien in allen 16 Bundesländern. Es zeigte sich, dass nicht alle Schulgesetze entsprechende Regelungen beinhalten und die vorhandenen Regelungen unterschiedlichen Umfangs und unterschiedlicher Reichweite sind. Die Schulgesetze

werden zur Frage der Notenbildung in allen Ländern durch Verordnungen und Erlasse ergänzt oder ersetzt, die in den Ländern unterschiedlich aufgebaut sind. Um die Vollständigkeit der identifizierten Dokumente so weit wie möglich sicherzustellen, wurden zusätzlich alle Kultusministerien der 16 Bundesländer angeschrieben und um die Einsendung von Hinweisen und Dokumenten zu diesbezüglichen zentralen rechtlichen Vorgaben gebeten. Dieser Anfrage kamen 12 von 16 Bundesländern nach. Insgesamt zeigte sich eine große Übereinstimmung zwischen den vorab identifizierten Regelungen und den durch die Ministerien bereitgestellten Hinweisen und Dokumenten. Bei der Dokumentenanalyse wurden schließlich 54 Textgrundlagen berücksichtigt. Hierbei handelte es sich neben den 16 Schulgesetzen um 26 Verordnungen und 12 Erlasse (vgl. Online-Supplement 1). Verordnungen und Erlasse zur Leistungsbeurteilung, die keine Antworten auf die in dieser Studie aufgeworfenen Fragen beinhalten, wurden aus der Studie ausgeschlossen. Zusammengefasst ergab sich bei der Dokumentenanalyse das in Abbildung 1 zusammengefasste Vorgehen, wobei mit Ausnahme von Schritt 2 alle Schritte von jeweils zwei Personen unternommen wurden.

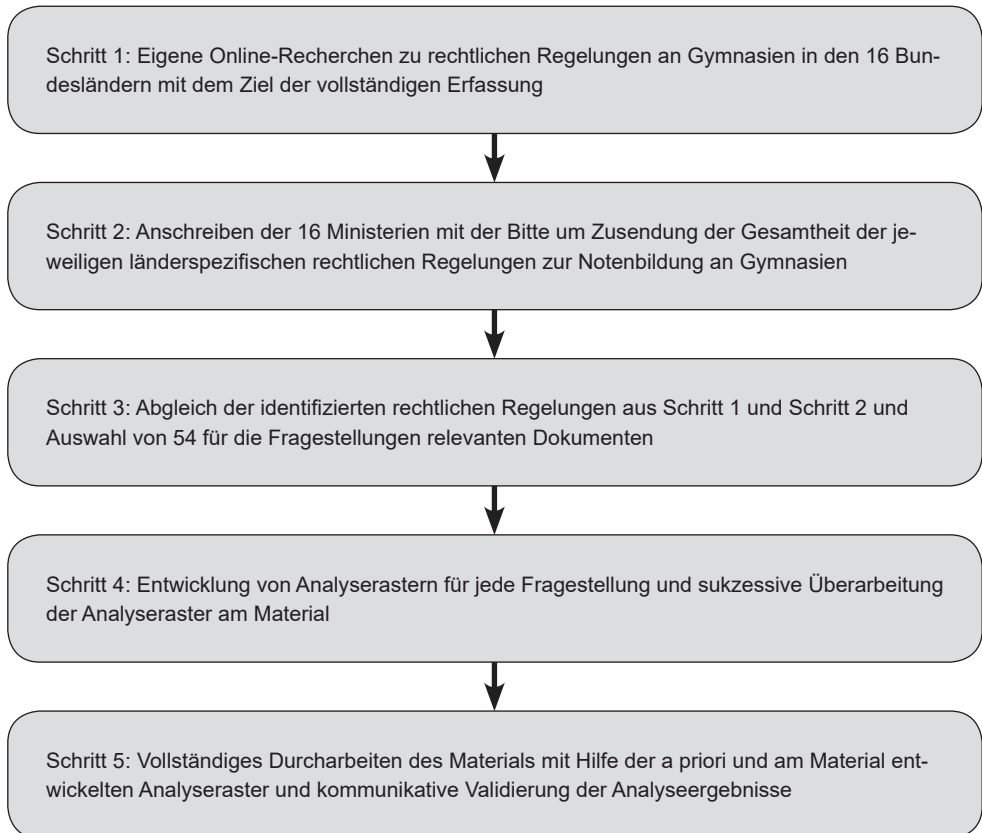


Abb. 1: Ablauf der Erfassung und Analyse von Dokumenten zu rechtlichen Regelungen bezüglich der Notenbildung an Gymnasien.

Die Analyse der Dokumente erfolgte schließlich entlang der Fragestellungen in niedrig-inferenter Weise. Nicht nur aufgrund der Art der Fragestellungen, sondern insbesondere wegen der Struktur und Sprache der Dokumente waren kaum interpretative Leistungen zu erbringen und konnten vielfach wörtliche Formulierungen in die länderspezifischen Tabellen übernommen werden (vgl. Online-Supplement 2). Es galt nicht – wie in anderen Dokumentenanalysen –, latente Sinnebenen aufzudecken und situative Kontexte der Entstehung zu berücksichtigen (z.B. Salheiser, 2014). Dennoch wurden die Analyseraster zunächst a priori (z. B. „Werden schriftliche Leistungen erwähnt?“) und dann am Material (z. B. „Gibt es Hinweise auf schulinterne Vergleichsarbeiten?“) von zwei Personen entwickelt und wurden die Dokumente auch parallel von zwei Personen analysiert, um relevante Aussagen mit Blick auf die Fragestellungen nicht zu übersehen und Fehler zu minimieren (vgl. Kuckartz & Rädiker, 2022). Die wenigen Unterschiede in den Ergebnissen wurden kommunikativ aufgelöst.

3.3 Ergebnisse

Fragestellung 1: Welche Leistungen bilden die Grundlage der Fachnoten an Gymnasien in den 16 Bundesländern?

Trotz der Vielfaltigkeit und Uneinheitlichkeit der länderspezifischen Regelungen findet sich in allen Bundesländern ein gemeinsamer Kern. Es besteht Einigkeit darin, dass verschiedene Leistungen zu bewerten sind und in die Gesamtnote eingehen. In allen Bundesländern werden hierbei schriftliche und mündliche Leistungen explizit benannt. Differenzen werden jedoch bei der Frage sichtbar, ob bzw. in welcher Weise und auf welcher Ebene der Normenhierarchie diese Leistungen lediglich benannt oder weiter konkretisiert werden: In den verschiedenen Bundesländern wird in Gesetzen, Verordnungen und Erlassen für Gymnasien oder alle allgemeinbildende Schularten insbesondere von ‚schriftlichen Leistungen‘ (z. B. BW, NI), ‚schriftlichen Arbeiten‘ (z. B. BB, HB), ‚schriftlichen Leistungsnachweisen‘ (HE) oder ‚bewerteten schriftlichen Arbeiten‘ (NI) gesprochen, die dann weiter ausgeführt oder mit Beispielen unterlegt werden. Besonders häufig finden sich dabei die Begriffe ‚Klassenarbeiten‘ und ‚Klausuren‘. Bayern und das Saarland gliedern zunächst nicht in schriftliche, mündliche und weitere Leistungen, sondern in ‚große und kleine Leistungsnachweise‘, die sie dann in länderspezifischer Weise ausführen.

Einige Bundesländer präzisieren die zu erbringenden schriftlichen Leistungen oder schriftlichen Arbeiten kaum (z. B. HH, TH), andere führen zusätzlich zu oder anstelle von Klassenarbeiten und Klausuren weitere Beispiele an, so z. B. ‚Kurzkontrollen‘, ‚Projektarbeiten‘ und ‚Schulleistungstests‘ (BE), inkludieren explizit Hausaufgaben (z. B. MV, BB) oder verweisen auf ‚Diktate, Vokabelkontrollen und Tests sowie die Aufbereitung von Materialien, Protokolle, Dokumentationen, Prozessberichte, Projektskizzen, Portfolios, Belegarbeiten, Belegmappen oder Exposés‘ (SN). Teilweise werden auch diese Konkretisierungen noch einmal präzisiert, so z. B. in der bayerischen Schulordnung für Gymnasien, in der ‚Kurzarbeiten‘ und ‚Stegreifaufgaben‘ als Beispiele für

„schriftliche Leistungsnachweise“ im Rahmen „kleiner Leistungsnachweise“ voneinander abgegrenzt werden. Schließlich unterscheiden sich die bundeslandspezifischen Regelungen noch bezüglich der Frage, ob alle erbrachten Leistungen Grundlage der Bewertung sind (z. B. in BW) oder spezifisch zum Nachweis des Leistungsstandes erbracht werden sollen (z. B. in BY): „Grundlage der Leistungsbewertung in einem Unterrichtsfach sind alle vom Schüler im Zusammenhang mit dem Unterricht erbrachten Leistungen“ (§ 7 Abs. 1 NVO). „Zum Nachweis des Leistungsstands erbringen die Schülerinnen und Schüler in angemessenen Zeitabständen entsprechend der Art des Fachs schriftliche, mündliche und praktische Leistungen“ (Art. 52 Abs. 1 BayEUG).

Auch mündliche Leistungen werden in einigen Bundesländern noch einmal genauer aufgeschlüsselt, so zum Beispiel in der bayerischen Regelung, in der „Rechenschaftsablagen, Unterrichtsbeiträge und Referate“ genannt werden, oder in Berlin, wo von „Beiträgen zum Unterrichtsgeschehen, mündlichen Kurzkontrollen, mündlichen Teilen von Projektarbeiten oder mündlichen Leistungsfeststellungen im Zusammenhang mit Vergleichsarbeiten“ die Rede ist. Differenzen zwischen den Bundesländern und unscharfe Bestimmungen zeigen sich bei der Abgrenzung von Mitarbeit und mündlicher Fachleistung: Neben Bundesländern, die hier eine klare Trennung vollziehen (z. B. BW, BY), gibt es Bundesländer, die mündliche Leistungen, mündliche Beiträge und Mitarbeit in spezifischer Weise aufeinander beziehen. Während beispielsweise in Brandenburg „mündliche Beiträge im Unterricht“ einen Teil der Leistungen bei der Mitarbeit im Fach darstellen (ebenso NI), gilt in Hessen die Mitarbeit im Unterricht der Oberstufe gleichrangig mit Protokollen oder Hausaufgaben als Teil der „kontinuierlich erbrachten Leistungen“. In Mecklenburg-Vorpommern hingegen werden kontinuierlich erbrachte mündliche Leistungen den sonstigen Leistungen zugerechnet (ebenso NW, SN).

Diese bundeslandspezifischen Besonderheiten korrespondieren mit Differenzen bei der Frage, ob die Mitarbeit als Teil der Fachleistung gilt oder wie das Verhalten als eigener Bereich in (Abschluss-)Zeugnissen ausgewiesen werden kann, muss oder nicht darf. Tatsächlich gibt es Bundesländer, die Beurteilungen von „Mitarbeit“ und „Verhalten“ (BW, BY, RP, SL, TH) oder von „Betragen, Fleiß, Mitarbeit und Ordnung“ (SN) überwiegend als eigenständige Bereiche in den Zeugnissen vorsehen, während andere Länder beim überfachlichen Teil der Beurteilungen das „Arbeits- und Sozialverhalten“ (BE, BB, HB, HE, MV, NI, NW) oder das „Lern- und Sozialverhalten“ (SH, ST) fokussieren. Hier gilt wiederum, dass sich unter gleichen Oberbegriffen in verschiedenen Bundesländern verschiedene Konkretisierungen finden lassen, die einen Vergleich erschweren. Bei der konkreten Beurteilung schließlich gibt es je nach Bundesland mehr oder weniger große Spielräume für die Einzelschule, die gefüllt werden können oder müssen (z. B. in BE, TH).

Fragestellung 2: Wie werden die verschiedenen Leistungen bei der Bildung der Fachnoten gewichtet?

Auch die Gewichtung der verschiedenen Leistungen bei der Bildung der Fachnoten wird, wie die Analyse zeigt, in den verschiedenen Bundesländern auf unterschiedlichen Ebenen und unterschiedlich präzise geregelt. Teilweise wird hierbei zwischen Sekun-

darstufe I und II (z. B. BE, ST) sowie zwischen verschiedenen Fächern oder Gruppen von Fächern (z. B. BY, HE) unterschieden; teilweise finden sich in mehreren Bundesländern innerhalb der auf der Systemebene beschlossenen Regelungen explizite Verweise auf die Einzelschule, die konkrete Entscheidungen zur Gewichtung verschiedener Leistungen treffen kann oder muss (z. B. BE, NW). Um in Anbetracht dieser Vielfalt einen übersichtlichen Vergleich zwischen den Bundesländern zu ermöglichen, erfolgte eine Begrenzung auf die Sekundarstufe I und hier auf schriftliche Leistungen im Sinne von Klassenarbeiten und Klausuren in den Fächern Deutsch und Mathematik, in denen schriftlichen Leistungen häufig ein größeres Gewicht als in anderen Fächern zukommt, und es wurden Zusätze wie „in der Regel“ oder „etwa“ vernachlässigt. Wie Abbildung 2 entnommen werden kann, zeigt sich hier eine große Uneinheitlichkeit zwischen den Bundesländern, in denen die Gewichtung unterschiedlich oder auch gar nicht geregelt ist. Hinzu kommt, dass in einigen Bundesländern die prozentualen Anteile bestimmter Leistungen für die Gesamtnote exakt benannt werden (z. B. BB, MV), während in Nordrhein-Westfalen eine „rein rechnerische Bildung der Kursabschlussnote“ (§ 13 APO-GOST) als unzulässig erklärt wird.

Nicht leistbar war an dieser Stelle eine vergleichende Darstellung der Gewichtung mündlicher Leistungen, weil – wie weiter oben ausgeführt wurde – in den verschiedenen Ländern kein einheitliches Verständnis von diesen Leistungen besteht und sie, wie auch die insgesamt zu erbringenden schriftlichen Leistungen, nicht überall einen abgrenzbaren Bereich bilden.

Fragestellung 3: Welche Anforderungen und Kriterien sollen bei der Bildung der Noten gelten?

In den verschiedenen rechtlichen Regelungen finden sich überwiegend sehr allgemeine Hinweise auf zugrundeliegende Anforderungen und relevante Kriterien für die Notenbildung. Explizit benannt werden in den meisten Bundesländern die Anzahlen der jährlich oder halbjährlich und minimal oder maximal zu schreibenden Klassenarbeiten bzw. Klausuren pro Fach oder Kurs. Beispielhaft finden sich in Tabelle 1 die minimal und/oder maximal vorgeschriebenen Klassenarbeiten für das fünfte Schuljahr.

Hierbei zeigt sich eine große Spannweite möglicher Minimalanforderungen, die von 9 bis 22 Klassenarbeiten pro Schuljahr reicht. Außerdem wird deutlich, dass sich die Länder darin unterscheiden, ob sie minimale, maximale oder exakte Anzahlen vorschreiben. Enorme Differenzen werden insbesondere bei einer vergleichenden Erfassung der vorgeschriebenen Klausuren in der Oberstufe sichtbar (vgl. Tabelle 2): Während beispielsweise in Thüringen eine Minimalzahl von 20 Klausuren gilt und in mehreren Ländern in etwa 40 Klausuren zu schreiben sind, werden im Saarland mindestens 72 Klausuren verlangt.

Über diese Quantifizierung hinaus bleiben die Angaben auf einem zumeist wenig konkreten Niveau, wobei überwiegend auf die in den Bildungsplänen formulierten Anforderungen verwiesen wird (z. B. SN, ST). Teilweise wird hier von ‚Kenntnissen, Fähigkeiten und Fertigkeiten‘ gesprochen (z. B. HB, TH), teilweise wird der Kompetenzbegriff anstelle dieser Begriffe (z. B. SL, SH) oder zusätzlich zu diesen (z. B. BE, BB)

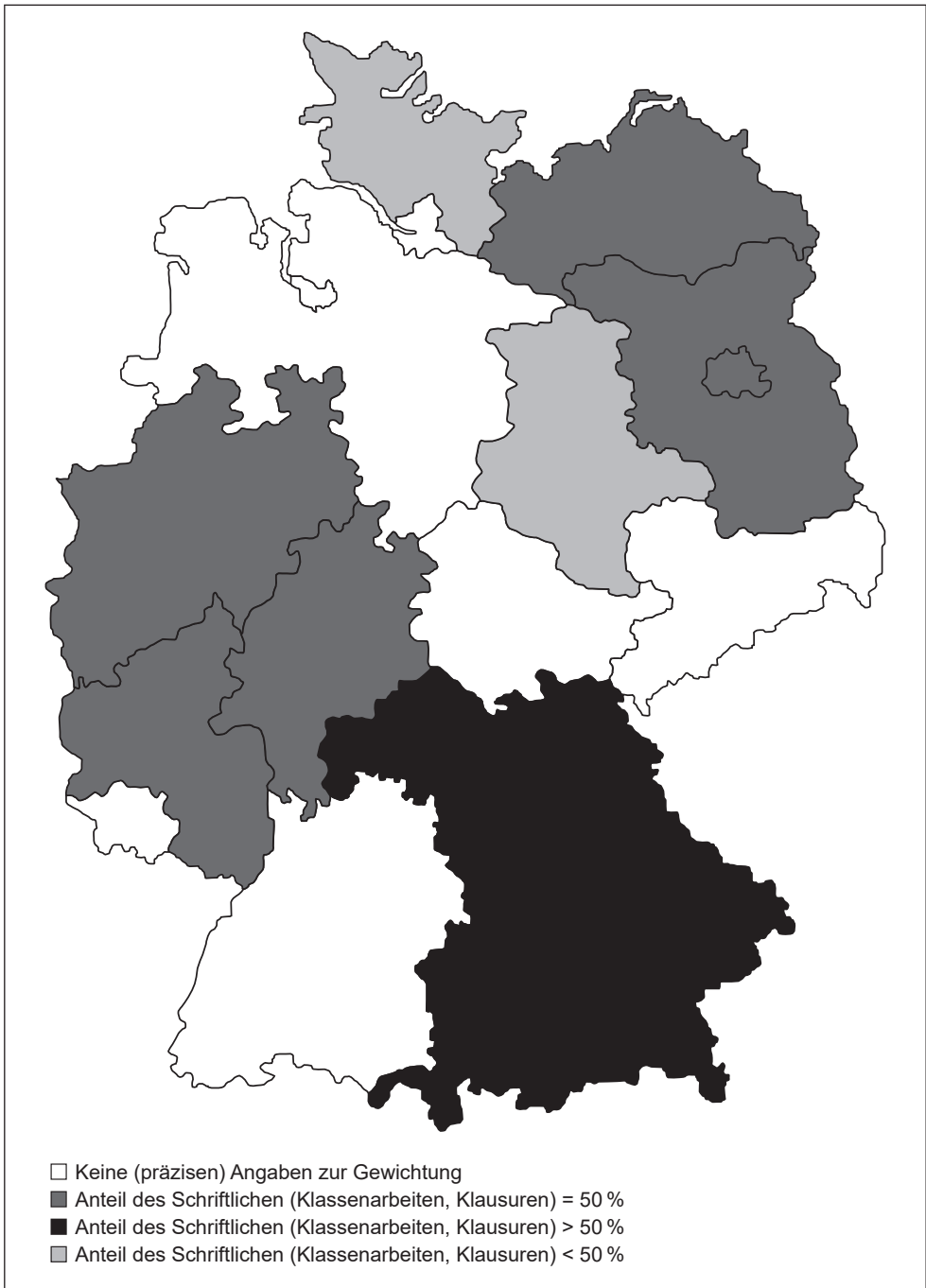


Abb. 2: Gewichtung schriftlicher Leistungen (Klassenarbeiten und Klausuren) bei der Bildung der Fachnoten in Deutsch und Mathematik in der Sekundarstufe I in den verschiedenen Bundesländern

	BW	BY	HB	HH	HE	MV	NI	NW	RP	SL	SN	ST	SH	TH
Zahl der Klassenarbeiten pro Kernfach	≥ 4	≥ 4	3–6	–	≥ 5	≥ 3	4–6	6	3–4	4	–	≥ 2	4–6	–
Gesamtzahl der Klassenarbeiten für alle Kernfächer	≥ 12	≥ 12	11–12	–	≥ 15	≥ 9	12–18	18	11–12	12	–	≥ 6	12–16	–
Zahl der Klassenarbeiten pro weiteres Fach	≤ 4	–	–	–	–	0–2	2	–	–	0	–	≥ 1	–	–
Gesamtzahl der Klassenarbeiten für alle weiteren Fächer	≤ 16	–	–	–	–	0–8	8	–	–	0	–	≥ 4	–	–
Gesamtzahl der Klassenarbeiten pro Schuljahr	≥ 12	≥ 12	11–12	–	≥ 15	≥ 9	22–30	18	11–12	12	≤ 25	≥ 10	12–16	–

Anmerkungen. Angaben basieren auf den rechtlichen Vorgaben zur Leistungsbewertung und den Stundentafeln der Länder. Fehlende Angaben in der Tabelle beruhen auf fehlenden Angaben in den rechtlichen Regelungen. Als weitere Fächer wurden mindestens zweistündig erteilte Fächer ohne Sport, Musik und Kunst erfasst. Zumeist können einzelne Klassenarbeiten durch andere Leistungen ersetzt werden; teilweise gelten für spezielle Profile spezifische Vorgaben, die hier vernachlässigt wurden. Berlin und Brandenburg wurden aufgrund ihrer sechsjährigen Grundschule nicht in die Darstellung einbezogen.

Tab. 1: Zahl der Klassenarbeiten pro Schuljahr (Klassenstufe 5 am Gymnasium).

	BW	BY	BE	BB	HB	HE	MV	NI	NW	RP	SL	SN	ST	TH
Zahl der Klausuren pro Leistungskurs (1. bis 3. Halbjahr)	≥ 2	1	2	1	≥ 2	2	1–2	1,3	2	1,75	2	≥ 2	1	1
Zahl der Klausuren pro Leistungskurs (4. Halbjahr)	≥ 1	1	1	1	≥ 1	1	1	1	1	1	2	≥ 1	1	1
Gesamtzahl der Leistungskurse	3	3	2	2	2–3	2	2	3	2	3	2	2	3	4
Gesamtzahl der Klausuren für alle Leistungskurse	≥ 21	12	14	8	≥ 14–21	14	8–14	25	14	24	16	≥ 14	12	16
Zahl der Klausuren pro Basis-kurs (1. bis 3. Halbjahr)	≥ 1	1	1	0–1	≥ 1	1	1–2	1	0–2	1	2	≥ 1	1	≥ 0
Zahl der Klausuren pro Basis-kurs (4. Halbjahr)	0–1	1	1	1	≥ 1	1	1	1	1	1	1	1	1	≥ 0
Gesamtzahl der Basiskurse	≥ 7	≥ 8	≥ 8	≥ 8	≥ 5	≥ 7	≥ 10	≥ 9	7	≥ 7	≥ 9	≥ 10	≥ 8	8
Gesamtzahl der Klausuren für alle Basiskurse	≥ 28	≥ 24	≥ 32	14	≥ 20–24	≥ 28	≥ 38	≥ 20	≥ 25	≥ 35	≥ 63	40–58	≥ 32	≥ 4
Gesamtzahlen der Klausuren für alle Leistungs- und Basis-kurse	≥ 49	≥ 36	≥ 46	22	≥ 38	≥ 42	≥ 46	≥ 45	≥ 39	≥ 59	≥ 79	54–72	≥ 44	≥ 20
Gesamtzahlen der Klausuren für alle Leistungs- und Basis-kurse, sofern Sport als Basis-fach belegt wird	≥ 45	≥ 32	≥ 42	22	≥ 34	≥ 38	≥ 42	≥ 41	≥ 39	≥ 54	≥ 72	50–72	≥ 40	≥ 20

Anmerkungen: Vereinfachte Darstellung für 14 Bundesländer, in der verschiedene Besonderheiten und Optionen für einzelne Fächer, Profile, Kombinationen, Zugangsvoraussetzungen, individuelle Belegungswünsche und spezifische schulische Angebote nicht berücksichtigt wurden. Die Vorgaben für Hamburg und Schleswig-Holstein wurden nicht in die Tabelle aufgenommen, da die Vielfalt möglicher Belegungen und Leistungsnachweise nicht mehr darstellbar war. Die Angaben für Rheinland-Pfalz beziehen sich auf fünf Schulhalbjahre.

Tab. 2: Zahl der Klausuren in der zweijährigen Qualifikationsphase am Gymnasium.

verwendet. In einigen Fällen finden sich Hinweise auf die mit den schriftlichen Leistungsnachweisen verbundenen Anforderungen, wonach beispielsweise eine ‚Verknüpfung‘ (BB, MV) oder ein ‚Transfer‘ (MV, SN) erlernter Inhalte geleistet werden soll und auch ‚Grundlagenwissen aus zurückliegenden Lernabschnitten‘ (SA) zur Anwendung kommen kann. Häufig, so auch in der hessischen Verordnung VOGSV, wird formuliert, es sei die Lern-, Leistungs- oder Kompetenzentwicklung mit zu beachten (ebenso NI, RP) und pädagogische Aspekte seien bedeutsam (ebenso BW, RP):

Hierbei ist zu beachten, dass Leistungsbewertung ein pädagogischer Prozess ist, der im Dienste der individuellen Leistungserziehung steht und der sich nicht nur auf das Ergebnis punktueller Leistungsfeststellungen, sondern auf den gesamten Verlauf der Lernentwicklung der Schülerin oder des Schülers bezieht. (§ 26 VOGSV)

Fast überall kann oder soll die sprachliche Richtigkeit und die Art der Darstellung bei schriftlichen Leistungen gewertet werden, wobei es über diese Gemeinsamkeit hinaus deutliche Differenzen zwischen den Bundesländern gibt. Während beispielsweise in der hessischen Verordnung eine exakte Formel für einen Fehlerindex und dessen Anwendung umfassend dargestellt wird, findet sich in der Berliner Verordnung eine gegenteilige Aussage, wonach die „Verwendung eines Fehlerquotienten bei der Beurteilung der Rechtschreibleistung [...] unzulässig“ (§ 19 Abs. 6 Sek I-VO) sei. Auch bezüglich anderer Kriterien zeigen sich deutliche Unterschiede zwischen den Bundesländern in Inhalt und Präzision. Während einige Länder (z.B. HB) sich auf ein Mindestmaß an Regelungen beschränken, werden in anderen konkrete methodische Hinweise für den Unterricht gegeben (z.B. BB). Zugleich finden sich vielfach explizite Hinweise zum pädagogischen Freiraum und zur pädagogischen Verantwortung der Lehrperson (z.B. in der NVO von BW und im Schulgesetz von RP): „Der Lehrer als Erzieher benötigt zur Verwirklichung seiner Aufgaben einen pädagogischen Freiraum, bei der Leistungsbeurteilung einen pädagogischen Beurteilungsspielraum“ (Vorbemerkungen, NVO). „Leistungsfeststellung und Leistungsbeurteilung werden gemäß § 25 Abs. 1 SchulG durch die pädagogische Verantwortung und die Freiheit der Lehrkraft bestimmt“ (§ 50 Abs. 1 SchulO RP). Schließlich finden sich zur zentralen Frage der Vergleichbarkeit von zu erbringenden Leistungen und entsprechenden Noten in einigen Bundesländern Regelungen, welche pro Schuljahr in mindestens einem Kernfach oder in jedem Fach eine klassenübergreifende Klassenarbeit im Sinne einer schulinternen Vergleichsarbeit mit abgestimmten bzw. gleichen Anforderungen und Kriterien vorsehen bzw. vorschreiben (z.B. SL, ST).

3.4 Diskussion

Der vorliegende Beitrag untersucht mit Blick auf die hohe Relevanz schulischer Noten für Übergangsentscheidungen und Auswahlprozesse sowie aufgrund bislang fehlender Auseinandersetzungen mit diesbezüglichen rechtlichen Vorgaben mittels Dokumenten-

analyse aktuelle länderspezifische Regelungen zur Benotung schriftlicher und mündlicher Leistungen an Gymnasien. Dies geschieht auch vor dem Hintergrund empirisch bestätigter Hinweise auf Zusammenhänge zwischen rechtlichen Vorgaben und Benotungsprozessen auf der Mikroebene (Hübner et al. 2020, Neumann et al. 2011). Die erhaltenen Ergebnisse werden im Folgenden entlang der Fragestellungen diskutiert und in übergeordnete Diskussionen eingeordnet.

Welche Leistungen bilden die Grundlage der Fachnoten?

Unbestritten bilden in allen Bundesländern verschiedene Leistungen, zu denen immer schriftliche und mündliche Leistungen gehören, die Grundlage der Fachnoten. Deutliche Differenzen zwischen den Ländern zeigen sich zum einen darin, wie die verschiedenen Leistungen voneinander abgegrenzt, gruppiert und aufeinander bezogen werden. Klare Unterschiede werden zum anderen bei der Frage sichtbar, wie ausführlich und präzise und mit welchem Vokabular die jeweiligen Leistungen benannt werden. Diese Differenzen bedeuten nicht nur, dass es Bundesländer mit mehr oder weniger Vorgaben für die Notengebung gibt, sondern sie bewirken auch, dass einfach erscheinende Fragen wie beispielsweise jene nach der Gewichtung schriftlicher und mündlicher Leistungen im Ländervergleich kaum beantwortbar sind. Sie erschweren außerdem Anschlüsse an erziehungswissenschaftliche Diskussionen, in denen Begriffsdefinitionen und -abgrenzungen von hoher Relevanz sind und es nicht beliebig ist, ob z. B. der Begriff der Kompetenz verwendet wird und ob, wie in Sachsen, der Begriff der Mitarbeit in eine Reihe mit den Begriffen ‚Betragen, Fleiß [...] und Ordnung‘ gestellt wird. Insgesamt erweisen sich die Regelungen als nur schwer vergleichbar, da sie auf verschiedenen Ebenen des Bildungssystems mit unterschiedlichen Reichweiten und Auflösungsgraden entwickelt und dann auf spezifische Weise sukzessive ergänzt und aktualisiert wurden.

Auch weitere Fragen wie jene, ob alle erbrachten Leistungen in einem Fach oder nur spezifische Leistungen zum Nachweis des jeweiligen Leistungsstandes Grundlage der Fachnoten sind, werden in den Ländern verschieden und zum Teil auch gar nicht geregelt. Dabei können sie weit reichende Folgen nicht nur für die Notengebung, sondern auch für die Gestaltung des Unterrichts – in diesem Fall als Lern- oder Leistungssituation (Baumert, 1999) – mit sich bringen, was einmal mehr die Relevanz rechtlicher Regelungen für das Handeln auf der Mikroebene verdeutlicht (vgl. Ditton, 2017; Fend, 2008).

Wie werden die verschiedenen Leistungen bei der Bildung der Fachnoten gewichtet?

Die Gewichtung schriftlicher Leistungen im Sinne von Klassenarbeiten und Klausuren ist in den Bundesländern unterschiedlich geregelt, wobei spezifische Vorgaben häufig fehlen und somit Freiräume entstehen können. Bedeutsam erscheint neben den numerischen Angaben zur Gewichtung auch die Frage, ob Gewichtungen exakt vorgeschrieben, mit Zusätzen wie „in der Regel“ versehen oder direkt an die Einzelschule verwiesen werden: Exakte Angaben zur Gewichtung sehen keine Spielräume vor, offene Vorschriften fordern Schulen bzw. Lehrpersonen auf, selbst geeignete Lösungen zu ent-

wickeln. Aus der Perspektive einzelner Lernender kann schließlich die Frage der Gewichtung schriftlicher Leistungen sehr relevant sein, wenn sie in diesem Bereich im Vergleich zu anderen Bereichen deutlich höhere oder geringere Leistungen zeigen.

Welche Anforderungen und Kriterien sollen bei der Bildung der Noten gelten?

Die Ausführungen zu den Anforderungen und Kriterien verbleiben überwiegend auf einem sehr allgemeinen Niveau und es finden sich nur wenige Präzisierungen. Exakt werden hingegen in fast allen Ländern die Anzahlen der zu schreibenden Klassenarbeiten bzw. Klausuren pro Fach und Klassenstufe benannt. Die Dokumentenanalyse kann zeigen, dass die Anforderungen im Sinne der vorgeschriebenen Zahl an Klassenarbeiten und Klausuren enorm zwischen den Ländern variieren. Damit berühren die rechtlichen Regelungen nicht nur die Frage, mit wie vielen Klassenarbeiten und Klausuren sich Lernende in einem Schuljahr konfrontiert sehen, sondern auch das unterrichtliche Handeln und die Korrekturleistungen der Lehrpersonen. Auch die Frage, ob und inwieweit Rechtschreibung und Darstellung in die Fachnote eingehen, wird in den Ländern unterschiedlich beantwortet, wodurch in einigen Bundesländern Zeugnisnoten anders als in anderen Ländern zustande kommen können oder müssen. Damit werden Lernende mit bestimmten Stärken bzw. Schwächen je nach Bundesland systematisch bevorteilt oder benachteiligt.

Mehrere Bundesländer formulieren auch Vorgaben für schulinterne Vergleichsarbeiten, wodurch sie wichtige schulinterne Abstimmungs-, Reflexions- und Entwicklungsprozesse sowie die Professionalisierung des Personals anstoßen und eine Vereinheitlichung bei der Notengebung befördern können. Auf der anderen Seite wird auch vielfach darauf verwiesen, dass es sich bei der Leistungsbeurteilung um eine pädagogische Aufgabe einer Lehrperson handelt und diese eines pädagogischen Freiraums bedarf. Interessant erscheint an dieser Stelle die Frage, die weitere Studien erfordert, inwiefern Lehrpersonen die rechtlichen Regelungen und Freiräume ihres Bundeslandes zur Notengebung kennen und umsetzen, wie sich also das Zusammenspiel zwischen Vorgaben auf der Makroebene und dem Lehrpersonenhandeln auf der Mikroebene beschreiben lässt und ob es hier Unterschiede zwischen den Ländern gibt.

Zusammenfassende Überlegungen

Innerhalb der einzelnen Bundesländer vermögen die untersuchten rechtlichen Regelungen eine vereinheitlichende Funktion auszuüben, vor allem im Falle konkret gefasster Ausführungen. In einigen Bundesländern sind die Regelungen jedoch sehr offen gehalten, wodurch Freiräume entstehen. Gestärkt wird die vereinheitlichende Funktion, wenn Hinweise zu klassenübergreifenden Klassenarbeiten im Sinne schulinterner Vergleichsarbeiten Teil der Regelungen sind, was häufig der Fall ist. Klare Handlungsanweisungen für Herausforderungen im Rahmen hoch anspruchsvoller Beurteilungsprozesse (vgl. z. B. Terhart, 2011) sind hingegen nicht Teil der rechtlichen Regelungen und können von diesen mit Blick auf die große Vielfalt an Schulen auch nicht erwartet werden. Sie können jedoch mit entsprechenden Hinweisen, welche Fragen von der Einzelschule zu regeln sind, Schulentwicklung befördern. Ebenso erlauben und verlangen (begrenzt)

offene Regelungen eine professionelle Ausgestaltung konkreter Beurteilungsprozesse durch professionelle Lehrpersonen vor Ort. Dabei sollte die unstrittige Notwendigkeit der so genannten pädagogischen Freiheit, die Lehrkräften den für ihre Arbeit notwendigen Gestaltungsspielraum sichern soll, nicht falsch verstanden werden: „Sie ist ihm nicht um seiner selbst, sondern um seiner Funktion, seines Amtes willen gewährleistet. Es handelt sich also nicht um eine personale, sondern um eine auf den Schulzweck, auf die Bildungsinteressen der Schüler bezogene Freiheit.“ (Avenarius 2002, S. 13; vgl. Wiezorek & Reimer, 2021). Lehrkräfte stehen somit vor der großen Herausforderung, im Spannungsfeld zwischen individuellem Blick und gesellschaftlicher Aufgabe zu agieren, also pädagogisch begründet und zugleich im mehr oder weniger offenen Rahmen rechtlicher Vorgaben zu benoten. Wünschenswert erscheint daher und mit Blick auf den Befund, dass sich Lehrkräfte häufig nicht angemessen vorbereitet auf die Beurteilung von Schülerinnen:leistungen sehen (vgl. Terhart, 2011), eine stärkere Verankerung der Thematik in der Lehrer:innenbildung.

Im Ländervergleich zeigen sich bei den Regelungen zur Bildung von Fachnoten (enorme) quantitative und qualitative Differenzen, aber auch unterschiedliche Begrifflichkeiten und Auflösungsgrade. Letztere mögen für rechtliche Dokumente nicht ungewöhnlich sein, erschweren aber Ländervergleiche, Überlegungen zu Vereinheitlichungen und Anschlüsse an erziehungswissenschaftliche Diskussionen und konnten im Rahmen dieses Beitrags nur in Ansätzen abgebildet werden. Perspektivisch wären weitere Dokumentenanalysen aufschlussreich, die eine systematische Differenzierung nach Fächern und Schulstufen vornehmen. Da die notwendigen Informationen für diese Differenzierungen nicht oder nur unzureichend genau in den analysierten Dokumenten vorliegen und teilweise schulspezifisch variierend umgesetzt werden, müssten hierzu weitere Dokumente identifiziert und in die Analyse aufgenommen werden.

Die Befunde der Dokumentenanalyse bestärken die vom Bundesverfassungsgericht benannte „eingeschränkte länderübergreifende Vergleichbarkeit“ (BVerfG, 2017) schulischer Noten. Sie lassen deutlich werden, dass die Vorschläge der KMK (2023) zur Vereinheitlichung der Zahl der Klausuren in der gymnasialen Oberstufe eine bedeutsame Veränderung mit Blick auf die Vergleichbarkeit der zu erbringenden Leistungen darstellen. Auch die systematische Bevorteilung oder Benachteiligung von Lernenden mit bestimmten Stärken und Schwächen in den verschiedenen Bundesländern ist ein wichtiger Befund. Aufgrund der hohen Relevanz von Noten im Rahmen von Auswahlprozessen und Zulassungsentscheidungen erscheint es wichtig, ihre Vergleichbarkeit weiter in den Blick zu nehmen. Der hier gewählte Fokus auf Differenzen in den länderspezifischen Regelungen, der einerseits Unterschiede zwischen, auf der anderen Seite aber auch große Gestaltungsspielräume innerhalb von Ländern aufzeigen konnte, stellt einen bedeutsamen Schritt hierzu dar.

Literatur

- Autor:innengruppe Bildungsberichterstattung (2022). *Bildung in Deutschland 2022*. Bielefeld: wbv. <https://doi.org/10.3278/6001820hw>
- Avenarius, H. (2002). Welche Rechte und Pflichten haben Lehrkräfte, Schulleitungen und Schulaufsicht bei der Qualitätsentwicklung der Schulen und bei der Sicherung gleicher Lebenschancen? *Zeitschrift für Bildungsverwaltung (ZBV)*, 17(2), 5–17.
- Avenarius, H. (2019). Die Bedeutung des Schulrechts für die Lehrerbildung. *Schulverwaltung: Fachzeitschrift für Schulentwicklung und Schulmanagement. Hessen, Rheinland-Pfalz*, 24(4), 108–111. <https://doi.org/10.25656/01:17608>
- Baumert, J. (1999). Befunde der internationalen Leistungsvergleiche zum mathematisch-naturwissenschaftlichen Unterricht und fachdidaktische Konsequenzen. In Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg.), *Weiterentwicklung des mathematisch-naturwissenschaftlichen Unterrichts* (S. 11–20). Waiblingen: Druckhaus Waiblingen.
- Baumert, J., Trautwein, U., & Artelt, C. (2003). Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann, & M. Weiß (Hrsg.), *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-97590-4_11
- Beutel, S.-I., & Pant, H. A. (2020). *Lernen ohne Noten: Alternative Konzepte der Leistungsbeurteilung*. Stuttgart: Kohlhammer.
- Birkel, P. (2005). Beurteilungsübereinstimmung bei Mathematikarbeiten? *Journal für Mathematikdidaktik*, 28(1), 28–51. <https://doi.org/10.1007/BF03339005>
- Bonefeld, M., & Dickhäuser, O. (2018). (Biased) Grading of students' performance: Students' names, performance level, and implicit attitudes. *Frontiers in Psychology*, 9, 481. <https://doi.org/10.3389/fpsyg.2018.00481>
- Bowers, A. J. (2011). What's in a Grade? The Multidimensional Nature of What Teacher Assigned Grades Assess in High School. *Educational Research & Evaluation*, 17(3), 141–159. <https://doi.org/10.1080/13803611.2011.597112>
- Bräu, K., & Fuhrmann, I. (2015). Die soziale Konstruktion von Leistung und Leistungsbewertung. In K. Bräu & C. Schlickum (Hrsg.), *Soziale Konstruktionen im Unterricht* (S. 49–64). Opladen: Barbara Budrich. <https://doi.org/10.2307/j.ctvdf0brd>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Ditton, H. (2017). Zum Wirkungszusammenhang der schulischen Handlungsebenen. In U. Steffens, K. Maag Merki, & H. Fend (Hrsg.), *Schulgestaltung* (S. 59–84). Münster: Waxmann.
- Falkenberg, K., Vogt, B., & Waldow, F. (2017). Ständig geprüft oder kontinuierlich unterstützt? Schulische Leistungsbeurteilung in Schweden zwischen formativem Anspruch und summativer Notwendigkeit. *Zeitschrift für Pädagogik*, 63(3), 317–333. <https://doi.org/10.25656/01:18542>
- Fend, H. (2008). *Neue Theorie der Schule* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Füssel, H.-P. (2020). Recht – ein blinder Fleck in der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 114–122). Bad Heilbrunn: Klinkhardt.
- Goodman, A. L. (2020). Can Group Oral Exams and Team Assignments Help Create a Supportive Student Community in a Biochemistry Course for Nonmajors? *Journal of Chemical Education*, 97, 3441–3445. <https://doi.org/10.1021/acs.jchemed.0c00815>
- Harks, B., Rakoczy, K., Klieme, E., Hattie, J., & Besser, M. (2014). Indirekte und moderierte Effekte von schriftlicher Rückmeldung auf Leistung und Motivation. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen* (S. 163–194). Münster/New York: Waxmann.

- Helbig, M., & Nikolai, R. (2015). *Die Unvergleichbaren. Der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949*. Bad Heilbrunn: Klinkhardt. <https://doi.org/10.25656/01:11095>
- Herzmann, P., & Liegmann, A. B. (2020). Mündliche Prüfungen im Kontext des Forschenden Lernens. *Zeitschrift für Pädagogik*, 66(5), 727–745. <https://doi.org/10.25656/01:24581>
- Hoffmann, L., Schröter, P., & Stanat, P. (2022). Jüngere Entwicklungen bei Abitur und Abiturprüfungen in Deutschland. In L. Hoffmann, P. Schröter, A. Groß, S. M. Schmid-Kühn & P. Stanat (Hrsg.), *Das unvergleichliche Abitur. Entwicklungen – Herausforderungen – Empirische Analysen* (S. 39–62). Bielefeld: wbv. <https://doi.org/10.3278/9783763972494>
- Holder, K., & Kessels, U. (2017). Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective. *Social Psychology of Education*, 20(3), 471–490. <https://doi.org/10.1007/s11218-017-9384-z>
- Hosenfeld, I., & Schrader, F.-W. (2006). *Schulische Leistung*. Münster: Waxmann.
- Hübner, N., Jansen, M., Stanat, P., Bohl, T., & Wagner, W. (2024). Alles eine Frage des Bundeslandes? Eine mehrbenenanalytische Betrachtung der eingeschränkten Vergleichbarkeit von Schulnoten. *Zeitschrift Für Erziehungswissenschaft*, 27, 517–549. <https://doi.org/10.1007/s11618-024-01216-9>
- Hübner, N., Spengler, M., Nagengast, B., Borghans, L., Schils, T., & Trautwein, U. (2022). When academic achievement (also) reflects personality: Using the personality-achievement saturation hypothesis (PASH) to explain differential associations between achievement measures and personality traits. *Journal of Educational Psychology*, 114(2), 326–345. <https://doi.org/10.1037/edu0000571>
- Hübner, N., Wagner, W., Hochweber, J., Neumann, M., & Nagengast, B. (2020). Comparing apples and oranges: Curricular intensification reforms can change the meaning of students' grades! *Journal of Educational Psychology*, 112(1), 204–220. <https://doi.org/10.1037/edu0000351>
- Hugo, J., & Heinrich, M. (2021). Zu den Eigenlogiken rechts- und erziehungswissenschaftlicher Perspektiven auf Bildung – die Kommission Bildungsorganisation, Bildungsplanung und Bildungsrecht (KBBB) auf der Suche nach ihrem dritten „B“. *Erziehungswissenschaft*, 32(62), 11–21. <https://doi.org/10.3224/ezw.v31i1.04>
- Ingenkamp, K. (Hrsg.) (1971). *Die Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz.
- Jürgens, E., & Sacher, W. (2008). *Leistungserziehung und Pädagogische Diagnostik in der Schule*. Stuttgart: Kohlhammer.
- Kalthoff, H. (2000). „Wunderbar, richtig“. Zur Praxis mündlichen Bewertens im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 3(3), 429–446. <https://doi.org/10.1007/s11618-000-0042-3>
- Klieme, E. (2003). Benotungsmaßstäbe an Schulen: Pädagogische Praxis und institutionelle Bedingungen. In H. Döbert, B. von Kopp, R. Martini, & M. Weiß (Hrsg.), *Bildung vor neuen Herausforderungen* (S. 195–210). Neuwied: Luchterhand.
- KMK (2022). *Standards für die Lehrerbildung: Bildungswissenschaften* (Beschluss der KMK vom 16. 12. 2004 i. d. F. vom 07. 10. 2022). Berlin. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf [10.09.2024].
- KMK (2023). *Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung* (Beschluss der KMK vom 07.07.1972 i. d. F. vom 16.03.2023). Berlin. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1972/1972_07_07-VB-gymnasiale-Oberstufe-Abiturpruefung.pdf [10.09.2024].
- KMK (2024). *Abiturnoten im Ländervergleich*. <https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html> [10.09.2024].

- Krämer, S., & Zimmermann, F. (2020). Zum Einfluss von störendem Schülerverhalten im Unterricht auf Leistungsbeurteilungen: Explizite Einschätzungen und experimentelle Befunde. *Zeitschrift für Pädagogische Psychologie*, 34(2), 99–115. <https://doi.org/10.1024/1010-0652/a000250>
- Krelle, M. (2014). *Mündliches Argumentieren in leistungsorientierter Perspektive. Eine empirische Analyse von Unterrichtsdiskussionen in der neunten Jahrgangsstufe*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Krieger, R. (2003). Das sogenannte Mündliche: eine Befragung von Beurteilten und Beurteilern zur Praxis und Problematik der Bewertung mündlicher Leistungen. *Bildung und Erziehung*, 56(1), 75–92. <https://doi.org/10.7788/bue.2003.56.1.75>
- Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (5. Aufl.). Weinheim/Basel: Beltz Juventa.
- Lintorf, K., & Buch, S. R. (2021). Stabile Präferenz oder flexibel am Diagnoseziel orientiert? – Die Bezugsnormwahl angehender Lehrkräfte. *Zeitschrift für Pädagogische Psychologie*, 35(2–3), 107–118. <https://doi.org/10.1024/1010-0652/a000271>
- Maier, U. (2015). *Leistungsdiagnostik in Schule und Unterricht*. Bad Heilbrunn: Klinkhardt. <https://doi.org/10.25656/01:18390>
- Meister, N., & Hollstein, O. (2018). Leistung bewerten. In M. Proske & K. Rabenstein (Hrsg.), *Kompendium Qualitative Unterrichtsforschung* (S. 123–136). Bad Heilbrunn: Klinkhardt.
- Neumann, M., Nagy, G., Trautwein, U., & Lüdtke, O. (2009). Vergleichbarkeit von Abiturleistungen: Leistungs- und Bewertungsunterschiede zwischen Hamburger und Baden-Württemberger Abiturienten und die Rolle zentraler Abiturprüfungen. *Zeitschrift für Erziehungswissenschaft*, 12(4), 691–714. <https://doi.org/10.1007/s11618-009-0099-6>
- Neumann, M., Trautwein, U., & Nagy, G. (2011). Do central examinations lead to greater grading comparability? A study of frame-of-reference effects on the University entrance qualification in Germany. *Studies in Educational Evaluation*, 37(4), 206–217. <https://doi.org/10.1016/j.stueduc.2012.02.002>
- Sacher, W. (2014). *Leistungen entwickeln, überprüfen und beurteilen* (6. Aufl.). Bad Heilbrunn: Klinkhardt.
- Salgo, L. (2021). (Wozu) Brauchen Pädagog*innen Rechtskenntnisse? Ein Zwischenruf. In DGfE (Hrsg.), *Erziehungswissenschaft und Bildungsrecht – Anknüpfungspunkte zur Belebung eines brachliegenden Diskurses* (S. 47–64). Opladen: Barbara Budrich. <https://doi.org/10.3224/ezw.v32i1.05>
- Salheiser, A. (2014). Natürliche Daten: Dokumente. In N. Baur, & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 813–827). Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-531-18939-0_62
- Schröter, P., Söldner, H., Hoffmann, L., Riemenschneider, A., Jost, J., & Wieser, D. (2022). Wie vergleichbar sind die Bewertungen von Abiturarbeiten im Fach Deutsch? In L. Hoffmann, P. Schröter, A. Groß, S. M. Schmid-Kühn & P. Stanat (Hrsg.), *Das unvergleichliche Abitur. Entwicklungen – Herausforderungen – Empirische Analysen* (S. 213–250). Bielefeld: wbv. <https://doi.org/10.3278/9783763972494>
- Schürmann, L., Käner, T., & Ringeisen, T. (2022). Need strength, perceived need support, stress symptomatology, and performance in the context of oral exams: A typological approach. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.992314>
- Sparfeldt, J. R., Rost, D. H., Baumeister, U. M., & Christ, O. (2013). Text anxiety in written and oral examinations. *Learning and Individual Differences*, 24, 198–203. <https://doi.org/10.1016/j.lindif.2012.12.010>
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19(1/2), 85–95. <https://doi.org/10.1024/1010-0652.19.12.85>

- Spinath, B., & Brünken, R. (2016). *Pädagogische Psychologie*. Göttingen: Hogrefe.
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of Teachers' Judgements of Students' Academic Achievement: A Meta-Analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>
- Terhart, E. (2011). Die Beurteilung von Schülern als Aufgabe des Lehrers. In ders., H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 699–717). Münster: Waxmann.
- Trapmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 21(1), 11–27. <https://doi.org/10.1024/1010-0652.21.1.11>
- Weich, M., & Stalder, B. E. (2019). Mathematik, Engagement und Leistung am Gymnasium. Eine Mixed-Method-Studie bei Lehrpersonen sowie Schülerinnen und Schülern an gymnasialen Maturitätsschulen. <https://zenodo.org/records/3581988> [10. 09. 2024].
- Wiezorek, C., & Reimer, F. (2021). Der Integrationsauftrag der Schule – erziehungs- und rechtswissenschaftliche Perspektiven. In DGfE (Hrsg.), *Erziehungswissenschaft und Bildungsrecht – Anknüpfungspunkte zur Belebung eines brachliegenden Diskurses* (S. 23–34). Oppladen: Barbara Budrich. <https://doi.org/10.3224/ezw.v32i1.03>
- Zabarowski, K. U., Meier, M., & Breidenstein, G. (2011). *Leistungsbewertung und Unterricht*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rechtliche Regelungen und Urteile

- APO-GOST. Verordnung über den Bildungsgang und die Abiturprüfung in der gymnasialen Oberstufe [Nordrhein-Westfalen] vom 5. Oktober 1998 i. d. F. vom 20. März 2023 (GV. NRW. S. 217). <https://bass.schul-welt.de/9607.htm> [10. 09. 2024].
- BayEUG. Bayerisches Gesetz über das Erziehungs- und Unterrichtswesen in der Fassung der Bekanntmachung vom 31. Mai 2000 (GVBl. S. 414, 632). <https://www.gesetze-bayern.de/Content/Document/BayEUG> [10. 09. 2024].
- NVO. Verordnung des Kultusministeriums über die Notenbildung vom 5. 5. 1983 i. d. F. vom 18. 6. 2020 (GBl. S. 577, 589). <https://www.landesrecht-bw.de/jportal/?quelle=jlink&query=NotBildV+BW&psml=bsbawueprod.psml&max=true&aiz=true> [10. 09. 2024].
- SchulO RP. Schulordnung für die öffentlichen Realschulen plus, Integrierten Gesamtschulen, Gymnasien, Kollegs und Abendgymnasien [Rheinland-Pfalz] vom 12. Juni 2009 i. d. F. vom 6. Dezember 2021 (GVBl. S. 631). <https://landesrecht.rlp.de/bsrp/document/jlr-SchulORP2009pP58/part/X> [10. 09. 2024].
- Sek I-VO. Verordnung über die Schularten und Bildungsgänge der Sekundarstufe I [Berlin] vom 31. März 2010 i. d. F. vom 9. August 2023 (GVBl. S. 302). https://gesetze.berlin.de/perma?a=SekIV_BE [10. 09. 2024].
- VOGSV. Verordnung zur Gestaltung des Schulverhältnisses [Hessen] vom 19. August 2011 i. d. F. vom 18. März 2021 (GVBl. S. 166). <https://www.rv.hessenrecht.hessen.de/bshe/document/hevr-SchulVerhGVHE2011V5Anlage2/part/X> [10. 09. 2024].
- BVerfG. Urteil des Ersten Senats vom 19. Dezember 2017 – 1 BvL 3/14 –, Rn. 1–253. https://www.bverfg.de/e/ls20171219_1bvl000314.html [10. 09. 2024].

Abstract: Due to the great importance of school grades for life and career opportunities on the one hand and diverse academic criticism regarding their meaning and comparability on the other hand, this article compares current legal regulations on grading practices in Germany. In order to do this, written and oral performance at grammar schools are focused on. Employing document analysis, we have investigated and discussed the subject grades that build the basis of school grades in the different federal states, how these are determined and weighted, and which requirements and criteria are legally valid for grade assignment. The findings suggest enormous variations in legal regulations between federal states, a lack of specificity within states, and differing terminologies.

Keywords: Grades, Achievement, Grammar School, Legal Regulations, Federal States

Anschrift der AutorInnen

Prof. Dr. Britta Kohler, Eberhard Karls Universität Tübingen,
Institut für Erziehungswissenschaft,
Münzgasse 26, 72070 Tübingen, Deutschland
E-Mail: britta.kohler@uni-tuebingen.de

Prof. Dr. Thorsten Bohl, Eberhard Karls Universität Tübingen,
Tübingen School of Education,
Wilhelmstraße 31, 72074 Tübingen, Deutschland
E-Mail: thorsten.bohl@uni-tuebingen.de

Jun.-Prof. Dr. Nicolas Hübner, Eberhard Karls Universität Tübingen,
Institut für Erziehungswissenschaft,
Münzgasse 26, 72070 Tübingen, Deutschland
E-Mail: nicolas.huebner@uni-tuebingen.de