

Klein, Esther D.; Kühn, Svenja M.; Ackeren, Isabell van; Block, Rainer
**Wie zentral sind zentrale Prüfungen? Abschlussprüfungen am Ende der
Sekundarstufe II im nationalen und internationalen Vergleich**

Zeitschrift für Pädagogik 55 (2009) 4, S. 596-621



Quellenangabe/ Reference:

Klein, Esther D.; Kühn, Svenja M.; Ackeren, Isabell van; Block, Rainer: Wie zentral sind zentrale Prüfungen? Abschlussprüfungen am Ende der Sekundarstufe II im nationalen und internationalen Vergleich - In: Zeitschrift für Pädagogik 55 (2009) 4, S. 596-621 - URN: urn:nbn:de:0111-opus-42654 - DOI: 10.25656/01:4265

<https://nbn-resolving.org/urn:nbn:de:0111-opus-42654>

<https://doi.org/10.25656/01:4265>

in Kooperation mit / in cooperation with:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Inhaltsverzeichnis

Thementeil: Werte – Die Rückkehr zur Erziehung

Hartmut von Hentig

Das Ethos der Erziehung. Was ist in ihr elementar? 509

Gertrud Nunner-Winkler

Prozesse moralischen Lernens und Entlernens 528

Matthias Gronemeyer

Wie kommt das Böse in den Menschen – und wie wieder heraus? Ein
philosophischer Kommentar zu Serkan A. und Spyridon L. 549

Margit Stein

Werteerziehungsansätze an weiterführenden Schulen in ihrem Zusammenhang
mit strukturellen Schulbedingungen 562

Thomas Gensicke

Jugendlicher Zeitgeist und Wertewandel 580

Allgemeiner Teil

Esther Dominique Klein/Svenja Mareike Kühn/Isabell van Ackeren/Rainer Block

Wie zentral sind zentrale Prüfungen? – Abschlussprüfungen am Ende der
Sekundarstufe II im nationalen und internationalen Vergleich 596

Rolf Strietholt/Ewald Terhart

Referendare beurteilen. Eine explorative Analyse von Beurteilungsinstrumenten
in der Zweiten Phase der Lehrerbildung 622

Besprechungen

Marcelo Caruso

Colin Brock/Lila Zia Levers (Hrsg.): Aspects of Education in the Middle East and North Africa 646

Heinz-Elmar Tenorth

Wolfgang Brezinka: Pädagogik in Österreich. Die Geschichte des Faches an den Universitäten vom 18. bis zum 21. Jahrhundert 648

Wolfgang Jantzen

Dagmar Hänsel: Die NS-Zeit als Gewinn für Hilfsschullehrer.

Dagmar Hänsel: Karl Tornow als Wegbereiter der sonderpädagogischen Profession. Die Grundlegung des Bestehenden in der NS-Zeit 650

Hanno Schmitt

Sieglinde Luise Ellger-Rüttgardt: Geschichte der Sonderpädagogik. Eine Einführung 653

Melanie Fabel-Lamla

Fritz Reheis: Bildung contra Turboschule! Ein Plädoyer.

Wiebke Lohfeld (Hrsg.): Gute Schulen in schlechter Gesellschaft.

Hans-Peter Gerstner/Martin Wetz: Einführung in die Theorie der Schule.

Gregor Lang-Wojtasik: Schule in der Weltgesellschaft. Herausforderungen und Perspektiven einer Schultheorie jenseits der Moderne 655

Cristina Allemann-Ghionda

Pedro Rosselló: La teoría de las corrientes educativas y otros ensayos

Christel Adick: Vergleichende Erziehungswissenschaft

Doris Edelmann: Pädagogische Professionalität im transnationalen Raum 660

Dokumentation

Pädagogische Neuerscheinungen 665

Allgemeiner Teil

Esther Dominique Klein/Svenja Mareike Kühn/Isabell van Ackeren/Rainer Block

Wie zentral sind zentrale Prüfungen?

Abschlussprüfungen am Ende der Sekundarstufe II im nationalen und internationalen Vergleich

Zusammenfassung: Das Zentralabitur ist in fast allen Ländern der BRD als Instrument der Standardsicherung eingeführt, gleichwohl wurde es mit teils deutlichen organisatorischen Unterschieden implementiert. Die Verfahren werden dargestellt und kategoriengeleitet verglichen sowie in einen Vergleich von 15 OECD-Mitgliedstaaten mit Zentralprüfungen am Ende der Sekundarstufe II eingeordnet. Diese Referenzperspektive offenbart die Bedeutung des Standardisierungsgrades von Prüfungen gegenüber dem Konzept der „Zentralisierung“ im Hinblick auf differenzierende Analysen der Steuerungswirkungen dieses Reformelementes auf schulische und unterrichtliche Prozesse.

Die deutschen Länder (mit Ausnahme von Rheinland-Pfalz) schließen mit der Einführung zentraler Prüfungen am Ende der Sekundarstufen I und II an internationale Entwicklungen an. Die Standardisierung des Prüfungsverfahrens wird als eine wesentliche bildungspolitische Maßnahme zur Sicherung von Qualitätsstandards angesehen. Dies betrifft sowohl die erreichten Kompetenzen als auch die Qualität des vorgelagerten Unterrichts, indem – so die Erwartung – Stoffgebiete aufgrund der auch für die Lehrkräfte unbekanntem Themenstellung hinreichend breit abgedeckt werden, innovative Curricula und Aufgabenformate schneller in der Breite durchgesetzt werden, die Leistungsbereitschaft von Schüler/innen und Lehrkräften im Sinne extrinsischer Motivierung erhöht und diagnostische Kompetenzen sowie die Anwendung eines kriterialen Bewertungsmaßstabes eher gefördert werden könnten (vgl. zur Anreizfunktion aus bildungsökonomischer Sicht z.B. Bishop/Wößmann 2004). Auch aus der Perspektive der die Abiturienten aufnehmenden Institutionen des Tertiärbereichs kann eine zentralisierte Prüfungsorganisation zur besseren Vergleichbarkeit von Institutionen und einer stärker leistungsorientierten individuellen Selektion im Berechtigungswesen beitragen (vgl. z.B. das Bildungsbarometer der Initiative Neue Soziale Marktwirtschaft und des Instituts der Deutschen Wirtschaft Köln; www.insm-bildungsmonitor.de).

Solchen positiven Steuerungserwartungen steht die Formulierung unerwünschter Nebenwirkungen gegenüber: Angenommen werden eine thematische Engführung des Unterrichts in der Erwartung bestimmter Aufgabenstellungen, die Förderung reproduzier-

tiven Lernens sowie die Marginalisierung verständnisorientierter und kreativer Arbeitsformen, die kaum von zentraler Stelle mit der erforderlichen Tiefe der Aufgabenstellung abfragbar seien; weiterhin fehlende Möglichkeiten, aktuelle Themen, lokale Bedingungen sowie eigene und schülerbezogene Interessen zu berücksichtigen (auch im Sinne einer deprofessionalisierenden Wirkung) sowie das Außerachtlassen unterschiedlicher Bedingungen der Leistungserbringung, etwa hinsichtlich der Frage, ob ein bestimmtes Thema bereits in allen Facetten im Unterricht besprochen wurde (vgl. zu den normativen Diskursen z.B. den Überblick bei van Ackeren 2007).

Bemerkenswert ist, dass sich diese Erwartungen kaum auf empirische Forschung, welche die Komplexität schulischer Bildungsprozesse berücksichtigt, stützen können. Eine Durchsicht der Forschungsbefunde (z.B. Bishop 1998; Baumert/Watermann 2000; Jürges/Schneider/Büchel 2003; Wößmann 2005, 2007; Fuchs/Wößmann 2007; Hanushek/Wößmann 2007), die insbesondere auf Reanalysen von TIMSS- und PISA-Datensätzen beruhen, zeigt tendenziell Vorteile zentraler Prüfungen für das Erreichen hoher Standards. Über die Studien hinweg erweist sich das Bild jedoch als inkonsistent und variiert fach- und altersgruppenspezifisch (vgl. zusammenfassend z.B. Maag Merki 2008). Die Wirkungen der Prüfungsorganisation auf schulische und unterrichtliche Prozesse – die vermutlich komplexer sind, als es ein vergleichsweise einfacher bildungsökonomischer Produktionsfunktionsansatz abzubilden vermag – wurde bislang nicht differenziert analysiert (vgl. auch Schümer/Weiß 2008). Im Rekurs auf die Educational Governance-Theorie (z.B. Altrichter/Brüsemeister/Wissinger 2007) kann davon ausgegangen werden, dass Steuerungshandeln auf der Bildungssystemebene auf unterschiedliche institutionelle und individuelle Handlungsbedingungen hin adaptiert wird. Die Adaption wird von selbstreflexiven und situativen Konstellationen multifaktoriell beeinflusst, so dass sich Auftragsplan (z.B. die Steuerung von Unterricht und Lernerträgen durch zentrale Prüfungen) und Realisierung durch die schulischen Akteure substantziell unterscheiden können.

Unterschiedliche Wirkungen von Steuerungshandeln auf der Systemebene in die Einzelschule und den Unterricht hinein werden – so die hier formulierte Annahme – auch durch unterschiedliche organisatorische Ausdifferenzierungen der Prüfungen in den Ländern bedingt. Die These basiert auf der Tatsache, dass sich hinter dem vordergründig einheitlichen Etikett „zentrale Prüfung“ verschiedene Konzepte verbergen. Auf dieses Problem der fehlenden Differenzierung des „Prüfungscharakters“ verweisen aktuell auch Schümer und Weiß (2008). Vor diesem Hintergrund warnen sie vor generellen Aussagen über die Leistungsrelevanz von Zentralprüfungen allein auf der Basis statistischer Daten und ohne die spezifischen Rahmenbedingungen zu berücksichtigen (vgl. ebd., S. 25).

Die Unterschiede der Prüfungsorganisation liegen augenscheinlich im Ausmaß der Zentralität des Prüfungsverfahrens als Grad der Konzentration struktureller und steuernder Elemente des Prüfungssystems auf die Bildungsadministration als zentrale Einheit. Die innerdeutsche Diskussion um das Zentralabitur ist fokussiert auf die Fragestellung, welche Instanz die Prüfungsaufgaben auswählt und stellt: der Lehrer oder das Ministerium, dezentral oder zentral organisiert. Dabei rücken relevante andere Aspekte

der Prüfungsorganisation, wie z.B. die Durchführung der Prüfung oder das Korrekturverfahren, eher in den Hintergrund. Das im internationalen Kontext verbreitete Gegensatzpaar „(schul-)extern“ versus „(schul-)intern“ ist augenscheinlich besser zur Beschreibung geeignet, wenn es gilt, alle Aspekte der Prüfungsorganisation – von der Vorbereitung der Prüfung bis hin zum Umgang mit den Prüfungsergebnissen – angemessen zu berücksichtigen. Wie sollte man z.B. einen Korrektor bezeichnen, der von einer anderen Schule als der Prüfling stammt, denn als (schul-)externen Korrektor? Es handelt sich weder um einen zentralen Korrektor noch um einen dezentralen Korrektor, denn letzterer ließe sich bei diesem Begriffspaar nicht von einem schulinternen Korrektor unterscheiden, was sich unter dem Gesichtspunkt der gewünschten Standardisierung des Prüfungsverfahrens jedoch als durchaus relevant erweist. Dabei wird hier unter Standardisierung die vorherige inhaltliche Festlegung bestimmter Abläufe und Verfahren verstanden, so dass möglichst nach den gleichen Maßstäben entschieden und gehandelt wird; dies betrifft die Gestaltung der Rahmenbedingungen der Prüfungen ebenso wie die Ausgestaltung der Prüfungsvorbereitungsphase, die Aufgabenstellung und den Umgang mit den Leistungen. Die im Beitragstitel formulierte Frage müsste demnach wie folgt präzisiert werden: Welchen Standardisierungsgrad erreichen die extern und intern organisierten Prüfungselemente des s.g. Zentralabiturs der deutschen Bundesländer aus einer internationalen Referenzperspektive?

Es wird deutlich, dass es einer differenzierten Perspektive auf die Prüfungsstrukturen bedarf, die unterschiedliche Steuerungswirkungen erwarten lassen (vgl. Schümer/Weiß 2008 sowie das nachfolgende Kapitel) und insofern in entsprechenden Forschungsdesigns Berücksichtigung finden sollten. Dazu ist eine systematische vergleichende Bestandsaufnahme der Organisation von Prüfungsverfahren im Ländervergleich notwendig, die sowohl für den innerdeutschen als auch für den internationalen Vergleich bislang nicht existiert.¹ Die nachfolgend dokumentierte Vergleichsstudie greift dieses Desiderat auf und konzentriert sich auf die allgemein bildende Sekundarstufe II.² Der Blick richtet sich einerseits auf die Bundesländer, andererseits auf ausgewählte OECD- Staaten, um die vielfältigen Typen und Varianten zentraler Prüfungsverfahren aufzuzeigen und die deutschen Entwicklungen verorten und in der Diskussion beurteilen zu können.

1. Länderauswahl und Methodisches Vorgehen

Die nationale Vergleichsstudie berücksichtigt – mit Ausnahme des dezentral prüfenden Landes Rheinland-Pfalz – alle Länder und greift auf die im schuljahr 2007/08 gültigen Verfahren zentraler Abiturprüfungen an Schulformen mit allgemein bildender Oberstufe

1 Die vom Konsortium Bildungsberichterstattung (2006) geleistete Bestandsaufnahme liefert eine knappe tabellarische Übersicht für das Jahr 2005/06.

2 Die Studie ist im Kontext des DFG-Projekts „Bedingungen und Wirkungen dezentraler und zentraler Abschlussprüfungen im naturwissenschaftlichen Unterricht“ im Rahmen der Forschergruppe & Graduiertenkolleg Naturwissenschaftlicher Unterricht in Kooperation zwischen den Universitäten Mainz und Duisburg-Essen entstanden.

zurück. Im Rahmen der internationalen Vergleichsstudie werden die Prüfungsverfahren am Ende der Sekundarstufe II (ISCED 3A) aufgegriffen. Für die Länderauswahl wurden folgende Kriterien berücksichtigt:

- *vollständig oder anteilig zentrale Prüfungsorganisation*: Da der Standardisierungsgrad der Prüfungen insbesondere vom Anteil zentraler Prüfungselemente abhängt, wird der Begriff der „Zentralprüfung“ vergleichsweise breit ausgelegt, um das Spektrum der Prüfungspraxis zu dokumentieren. Insofern werden alle Prüfungsformen einbezogen, bei denen die Schüler/-innen Prüfungsaufgaben bearbeiten, die von einer schulexternen Stelle entwickelt wurden.
- *möglichst vergleichbare wirtschaftliche Rahmenbedingungen*: Die Betrachtung von Staaten vergleichsweise ähnlicher Wirtschafts- und Wohlstandslage bietet eine Basis für die Generierung von Hypothesen zu möglichen Zusammenhängen von Prüfungsverfahren und schulischen Lernerträgen im Rahmen weiterer Analysen.
- *keine föderales Organisationsprinzip*: Eine differenzierte Analyse unterschiedlicher Prüfungsverfahren innerhalb föderaler Systeme kann im internationalen Vergleich an dieser Stelle nicht geleistet werden.

Unter Berücksichtigung dieser Prinzipien wurden 15 OECD-Staaten ausgewählt: Dänemark, Finnland, Frankreich, Irland, Italien, Luxemburg, Neuseeland, die Niederlande, Norwegen, Polen, Portugal, Schottland, die Slowakei, Ungarn und das Vereinigte Königreich.

Tabelle 1 gibt einen Überblick über die Art der Prüfungsverfahren am Ende der Sekundarstufe II in den OECD-Staaten und allen Ländern der BRD. Die dort aufgeführten Länder spiegeln zugleich die Bandbreite der Möglichkeiten, den Übergang zwischen Sekundär- und Tertiärbereich zu gestalten, die sich über folgende zwei Grundmodelle beschreiben und jeweils weiter ausdifferenzieren lassen:

- Schulen attestieren ihren Absolventen eine öffentlich definierte Leistung, worauf sich die abnehmenden Systeme (Berufsausbildung, Hochschulen, Beschäftigungssystem) bei der Aufnahme verlassen. In Ausdifferenzierung dieses Grundmodells können die Zertifikate entweder auf der Basis grundsätzlich zentral organisierter Prüfungen oder einer lediglich als Hochschulzugangsberechtigung verpflichtenden zentral abzulegenden Prüfung oder einer dezentral organisierten Prüfung oder aber einer kumulativen Prüfung im Verlauf des Schuljahrs erworben werden.
- Das abnehmende System stellt in selbst entwickelten Verfahren in den vorgelagerten Systemen erworbene Qualifikationen fest und macht den Zugang zu ihm von der Erbringung der als erforderlich erachteten Leistung abhängig. Auch an dieser Stelle können dezentrale, durch die jeweilige Hochschule organisierte Prüfungen und zentrale, landesweite Verfahren unterschieden werden.

Die Ergebnisse der nachfolgend berichteten nationalen Studie zu den zentralen Prüfungsverfahren stützen sich auf eine Analyse der länderspezifischen Abiturprüfungsverordnungen sowie auf im Jahr 2007/2008 durchgeführte Experteninterviews, u.a. mit Mitarbeitern in den Ministerien sowie mit Mitgliedern aus Aufgabenkommissionen. Die

ausschließlich zentrale Abschlussprüfungen	teilweise zentrale Abschlussprüfungen	dezentrale Abschlussprüfungen	keine Abschlussprüfungen	föderale Organisation mit teils unterschiedlichen Verfahren
Dänemark Finnland Frankreich Irland Luxemburg Polen ¹ Portugal Schottland Vereinigtes Königreich Baden-Württemberg Bayern Hessen Mecklenburg-Vorpommern Niedersachsen Nordrhein-Westfalen Saarland Sachsen Sachsen-Anhalt Schleswig-Holstein Thüringen	Italien Neuseeland Niederlande Norwegen Slowakei Ungarn Berlin Brandenburg Bremen Hamburg	Island Österreich Schweiz Tschechien ² Rheinland-Pfalz	Griechenland ⁴ Japan ⁵ Korea ⁵ Schweden ⁶ Spanien ⁷ Türkei ⁶	Belgien ³ Australien Kanada USA ⁸
<p>1 Zentrale Prüfungen nur für Hochschulzugangsberechtigung verpflichtend</p> <p>2 erstmals Zentralprüfungen nach dem Schuljahr 2007/08 (Sek. II)</p> <p>3 Schulen können nach eigenem Ermessen Prüfungen durchführen; in Flandern Hochschuleingangsprüfungen für bestimmte Studiengänge</p> <p>4 teilweise zentrale Prüfungen am Ende der Sekundarstufe II, die jedoch den Status von Hochschuleingangsprüfungen haben</p> <p>5 zentrale Prüfungen nach jedem Schulhalbjahr, jedoch keine speziellen Abschlussprüfungen</p> <p>6 zentrale Hochschuleingangsprüfungen</p> <p>7 Universitäten organisieren Hochschuleingangsprüfungen nach staatlichen Rahmenvorgaben</p> <p>8 bundesstaatenübergreifende, privat organisierte Eignungstests (z.B. SAT) sind Hauptkriterium für Hochschulzugang</p>				

Tab. 1: Übersicht über zentrale Abschlussprüfungen am Ende der Sekundarstufe II in den OECD-Staaten und den Ländern der Bundesrepublik

internationale Studie basiert auf der Darstellung der Bildungssysteme der europäischen Staaten auf der Bildungsplattform „Eurydice“ der Europäischen Kommission, den Internetpräsenzen der Ministerien und Prüfungskommissionen der jeweiligen Staaten sowie auf den entsprechenden Prüfungsverordnungen. Die Recherche wurde durch Expertenbefragungen bei den zuständigen Behörden ergänzt.

In Parallelisierung des nationalen sowie des internationalen Vergleichs beginnen die Analysen jeweils mit einer knappen allgemeinen Skizze der Rahmenbedingungen der Prüfungsverfahren und betrachten anschließend die einzelnen Prüfungselemente, die – dem chronologischen Verlauf der Prüfung entsprechend – in die Kategorien „Prüfungsvorbereitung“, „Prüfungsdurchführung“, „Prüfungskorrektur“ und „Umgang mit Prüfungsergebnissen“ untergliedert sind. In diesen Bereichen gibt es Unterschiede im Grad der Standardisierung, die wiederum differenzierte Steuerungswirkungen für die Qualität von Bildungsprozessen und -ergebnissen erwarten lassen.

In der Kategorie *Rahmenbedingungen* werden Aspekte der Prüfungstradition sowie Hintergründe bzw. Motive für die Einführung zentraler Prüfungen dargestellt. Das Herausarbeiten spezifischer nationaler „Steuerungsphilosophien“ (z.B. die Implementierung von Zentralprüfungen als Element von Bildungsmärkten, vgl. z.B. van Ackeren 2003) kann die möglicherweise grundlegende Bedeutung für die Adaption dieses Instrumentes durch die beteiligten Akteure aufzeigen. Gleichwohl bedarf es vertiefter qualitativer Studien zu dieser Fragestellung über diesen Beitrag hinaus. Die Kategorie *Prüfungsvorbereitung* greift u.a. die Anzahl der Prüfungsfächer, die Reichweite der zentralen Prüfungen (Anteil der zentral organisierten Prüfungen am gesamten Prüfungsverfahren) und die Aufgabenentwicklung auf. Diese Aspekte sind ebenfalls Gradmesser der Vereinheitlichung der Prüfungsorganisation. In diesem Zusammenhang sollte die angenommene steuernde Wirkung auf vorgelagerte Arbeitsprozesse umso größer sein, je mehr Prüfungsfächer zentral geprüft werden und je größer das Gewicht der zentral geprüften Leistungen an der Gesamtqualifikation ist. Ein weiteres Beispiel für mögliche differenzielle Steuerungseffekte ist die Zusammensetzung der Aufgabenkommissionen. Die Einbindung von Fachdidaktikern könnte z.B. eher dazu beitragen, curriculare Neuerungen durchzusetzen, als allein mit Lehrkräften besetzte Kommissionen. Die Kategorie *Prüfungsdurchführung* schließt alle organisatorischen Gesichtspunkte ein, die zur eigentlichen Prüfung gehören (Zeitpunkt, Ort, Aufsicht), die Anzahl und Dauer der schriftlichen Prüfungen sowie Möglichkeiten der Aufgabenauswahl. Auch in dieser Hinsicht ergeben sich wieder Möglichkeiten, den Standardisierungsgrad zu erhöhen oder auch zu lockern, wenn z.B. Auswahlmöglichkeiten bei der Bearbeitung von Aufgaben gegeben sind. Die *Prüfungskorrektur* umfasst Aspekte des Korrekturverfahrens, Lösungs- und Bewertungsvorgaben und die Anonymität der Prüflinge; damit sind Maßnahmen verbunden, welche die Einhaltung der Gütekriterien der Leistungsmessung erhöhen können. Schließlich werden in der Kategorie *Umgang mit Prüfungsergebnissen* sowohl die Bedeutung der Ergebnisse für die Abschlussnote als auch die Frage nach einer Ergebnismeldung an die Schulen bzw. die Öffentlichkeit betrachtet.

Die Darstellung orientiert sich an einem in der Vergleichenden Erziehungswissenschaft häufig angewandten Ablauf von Analyseschritten. So wurden zunächst die landesspezifischen Strukturen und Verfahren im Rahmen von Länderberichten beschrieben (*Deskription*). Diese können aufgrund ihres Umfangs an dieser Stelle jedoch nicht abgebildet werden. Auf der Grundlage der oben beschriebenen Analysekategorien werden die Länderbefunde nach einheitlichen Kriterien nebeneinander gestellt und analysiert (*Juxtaposition*) sowie einer vergleichenden Betrachtung unterzogen (*Komparation*).

2. Das Zentralabitur im nationalen Vergleich

Die KMK gibt mit der „Vereinbarung über die Abiturprüfung der gymnasialen Oberstufe in der Sekundarstufe II“ (KMK 2007) sowie den fachspezifischen Einheitlichen Prüfungsanforderungen (EPA)³ in der Abiturprüfung bundesweit gültige Rahmenvorgaben vor, die jedoch länderspezifisch unterschiedlich ausgestaltet werden können, so dass sich in der gegenwärtigen Praxis zum Teil erhebliche Differenzen in der Prüfungsgestaltung zeigen. Gemeinsam ist dem zentralen Prüfungsverfahren über die Länder hinweg, dass alle Schüler/-innen eines Landes in bestimmten Fächern zur gleichen Zeit die gleichen Aufgabenstellungen bzw. vom Kurslehrer oder selbst ausgewählte Aufgaben aus einem Aufgabenpool bearbeiten. Der Beitrag bezieht sich auf die schriftliche Abiturprüfung, da die mündlichen Prüfungen immer als dezentrale Prüfungen mit schulinternen Aufgabenstellungen konzipiert sind.⁴ Für länderspezifische Zuordnungen der Ausführungen wird aus Platzgründen auch auf Tabelle 2 verwiesen.

2.1 Rahmenbedingungen

Nach zuvor ausschließlich dezentral organisierten Prüfungsverfahren lassen sich seit 1945 drei Implementationszeiträume des Zentralabiturs festmachen: Traditionell zentral prüfen – größtenteils durch den Einfluss der Alliierten bedingt – das Saarland (seit 1945), Bayern (seit 1946) und Baden-Württemberg (seit 1946 bzw. 1952). Nach der Wiedervereinigung 1990 knüpften Sachsen, Sachsen-Anhalt, Thüringen und Mecklenburg-Vorpommern an die Zentralprüfungstradition der ehemaligen DDR (dort seit 1959) an. Die Mehrheit der Länder hat sich erst nach der Vorlage der PISA-Befunde auf diesen Weg gemacht (vgl. ausführlicher van Ackeren 2007).

2.2 Prüfungsvorbereitung

Prüfungsfächer und Reichweite. Im Rahmen der Abiturprüfung sind vier (in der Mehrheit der Länder) oder fünf Prüfungen abzulegen. Dabei unterscheiden sich sowohl die Anzahl der schriftlichen Prüfungen (vgl. 2.3) als auch die Prüfungsfächer, was auf strukturelle Unterschiede der gymnasialen Oberstufen zurückzuführen ist: In Systemen mit verpflichtenden Kernfächern und einer begrenzten Anzahl an Profulfächern legen die Schüler/-innen die schriftlichen Prüfungen in bestimmten Fächern ab (z.B. Baden-Württemberg). In den übrigen Systemen können die Prüfungen nach Maßgabe der Verordnungen und einzelschulischen Angebote in allen wählbaren Fächern des Kursangebots stattfinden.

3 Vgl. www.kmk.org/schul/home.htm; dort sind alle EPA (seit 2002) veröffentlicht.

4 In einigen Ländern kann ein externer Prüfer an der Prüfung teilnehmen.

In zehn von fünfzehn Ländern finden alle schriftlichen Prüfungen zentral statt. In der Gruppe der verbleibenden Länder erfolgen sie ausschließlich in bestimmten Fächern bzw. Fächergruppen, während die übrigen weiterhin dezentral geprüft werden (Berlin, Brandenburg, Hamburg). In den anderen beiden Ländern (Bremen, Schleswig-Holstein) wird zudem nur auf bestimmten Anforderungsniveaus zentral geprüft. Für die fünf Länder zeigt sich, dass Deutsch, Englisch, Französisch und Mathematik immer zentral geprüft werden. Hinzu kommen – je nach Land – weitere Sprachen sowie die Naturwissenschaften und ausgewählte Fächer des gesellschaftswissenschaftlichen Aufgabenfeldes.

Inhaltliche Vorgaben. Die inhaltliche Grundlage der Prüfungen bilden die bundesweit gültigen EPA sowie die landesspezifischen Lehrpläne (bzw. Kerncurricula, Rahmenrichtlinien). Die Mehrheit der Länder benennt weiterhin s.g. Schwerpunktthemen, also bestimmte Themen des Lehrplans, die potenziell Gegenstand der Prüfung sein können. Während einzelne Länder die Themen bereits mehr als zwei Jahre im Voraus benennen, veröffentlicht die Mehrheit der Länder diese ein bis zwei Jahre vorher, Mecklenburg-Vorpommern erst ein Jahr vorher (vgl. Lavorgna 2007). In einigen Ländern, die bestimmte verbindliche Lehrplaninhalte benennen, gelten diese zeitlich unbegrenzt.

Aufgabenentwicklung. Kern der organisatorischen Prüfungsvorbereitung ist die Entwicklung der Abituraufgaben. Dieses Verfahren wird von der obersten Schulaufsichtsbehörde koordiniert, wobei sich zwei Modelle unterscheiden lassen:

- Zum einen ein *top down*-Modell, bei dem die Aufgaben durch Kommissionen an den jeweiligen Landesinstituten entwickelt werden, in denen von der obersten Behörde beauftragte erfahrene Lehrkräfte sowie – bundeslandabhängig – Fachreferenten des Landesinstituts bzw. der Schulaufsichtsbehörde zusammenarbeiten (lediglich Sachsen-Anhalt ergänzt die Kommissionen um Fachdidaktiker der Universitäten);
- zum anderen ein *bottom up*-Modell, bei dem von der obersten Aufsichtsbehörde bestimmte erfahrene Lehrkräfte im Vorfeld der Prüfungen mit der Erstellung von Vorschlägen beauftragt werden und eine Kommission daraus Aufgaben auswählt bzw. Modifikationen vornimmt.

2.3 Durchführung der Prüfung

Die Abiturprüfung findet immer zum Ende der Qualifikationsphase zu landesweit einheitlichen Terminen in den Schulen statt. Sie wird von schulinternen Lehrkräften, i.d.R. die jeweiligen Kurslehrer, beaufsichtigt.

Anzahl der Prüfungen. Von den insgesamt vier oder fünf abzulegenden Teilprüfungen sind drei oder vier Prüfungen schriftlich konzipiert. In zwölf der fünfzehn Länder werden die Kandidat/-innen in drei Fächern schriftlich geprüft, lediglich in Baden-Württemberg, Mecklenburg-Vorpommern und Sachsen-Anhalt müssen sie vier Prüfungen ablegen. Dabei prüft Baden-Württemberg alle vier Fächer auf erhöhtem Anforderungsniveau, während Mecklenburg-Vorpommern und Sachsen-Anhalt jeweils zwei

Fächer auf grundlegendem und erhöhtem Niveau prüfen.⁵ Alle anderen Länder prüfen zwei Fächer auf erhöhtem und ein Fach auf grundlegendem Niveau.

Aufgabenauswahl. Immer entscheidet die oberste Schulaufsichtsbehörde, welche Aufgaben zentral gestellt werden. In der Regel wird den Schulen ein landesweit einheitlicher Aufgabenpool vorgegeben, aus dem entweder die zuständige Kurslehrkraft und/oder die Prüflinge selbst eine Auswahl treffen müssen, mit folgenden Varianten: Drei Länder bieten grundsätzlich eine Auswahlgelegenheit für Lehrkräfte und Prüflinge an; vier Länder ermöglichen – fachunabhängig – immer allen Schüler/-innen Wahlmöglichkeiten. Die Mehrheit der Länder trifft fachspezifische Regelungen, z.T. ist auch keine Auswahlmöglichkeit gegeben.

Zeitlicher Umfang der Prüfungen. Fünf Länder sehen auch für die Prüfungsdauer fachspezifische Regelungen vor, die übrigen Länder geben für die schriftlichen Arbeiten in allen Fächern die gleiche Bearbeitungszeit vor. Die schriftlichen Prüfungen in Kursen mit erhöhtem Anforderungsniveau dauern in der Mehrheit der Länder 300 Minuten; lediglich Thüringen, NRW, Brandenburg und Hessen weichen mit 30 bis 60 Minuten weniger Zeit von dieser Regelung ab. Auch bei den Prüfungen mit grundlegendem Anforderungsniveau gibt es Differenzen von bis zu 60 Minuten. Schließlich zeigen die Anzahl der Aufgaben, aus denen ausgewählt werden muss, die zugelassenen Hilfsmittel und die Gestaltung der Aufgaben zum Teil erhebliche Unterschiede – sowohl zwischen den Ländern als auch aus fachspezifischer Sicht. Unter anderem zeigt sich in einigen Ländern, dass die in den EPA geforderten Unterschiede zwischen Kursen mit grundlegendem und erhöhtem Anforderungsniveau teils quantitativer und nicht qualitativer Art sind, indem mehr Aufgaben zu bearbeiten sind.⁶ Zudem wird eine Spannweite bezüglich der Offenheit von Aufgaben deutlich – von kleinschrittigen, reproduktiven Aufgaben bis hin zu offeneren Aufgabenstellungen, die mehrere Lösungswege zulassen. Auch fachspezifische Arbeitsweisen, z.B. Formen kreativen Schreibens oder experimentellen Arbeitens, werden offensichtlich eher in einer Minderheit der Länder geprüft.

2.4 Prüfungskorrektur

Erwartungshorizonte und Bewertungsmuster. Generell wird ein Erwartungshorizont bzw. Bewertungsmuster vorgegeben, wobei alternative Lösungen zulässig sind. In Be-

5 Der Fachunterricht wird nach Maßgabe der EPA auf unterschiedlichen Anspruchsebenen erteilt. Dabei repräsentiert Unterricht mit grundlegendem Anforderungsniveau das Lernniveau der gymnasialen Oberstufe unter dem Aspekt einer wissenschaftspropädeutischen Bildung. Unterricht mit erhöhtem Anforderungsniveau repräsentiert das Lernniveau der gymnasialen Oberstufe unter dem Aspekt einer wissenschaftspropädeutischen Bildung, die exemplarisch vertieft wird. Neben der traditionellen Untergliederung in Grund- (grundlegendes Anforderungsniveau) und Leistungskurse (erhöhtes Anforderungsniveau), existieren in fünf Bundesländern Systeme mit verpflichtenden Kernfächern und einer begrenzten Anzahl an Profulfächern (erhöhtes Anforderungsniveau) sowie zusätzlichen Kursen auf grundlegendem Anforderungsniveau.

6 Auf der Grundlage einer ersten internen Analyse von Prüfungsaufgaben im Ländervergleich im Rahmen eines Dissertationsvorhabens.

zug auf die Bewertungsgrundlagen zeigen sich drei Varianten: Zum einen werden für bestimmte Teilaufgaben detaillierte Punkte vergeben; aus einer dazugehörigen Verrechnungstabelle ergibt sich dann eine Endnote (z.B. Mathematik, Naturwissenschaften). Eine weitere Möglichkeit sind die Vorgabe von inhaltlichen Schwerpunkten sowie groben Punkteverteilungen für bestimmte Teilbereiche (z.B. Fremdsprachen: Inhalt, Sprachrichtigkeit und Ausdrucksvermögen) oder inhaltliche Vorgaben ohne die Angabe von Verrechnungspunkten (teilweise – mit länderspezifischen Unterschieden – z.B. in Deutsch). Hier spiegeln sich offensichtlich fachkulturelle Unterschiede.

Korrekturverfahren. Die Erstkorrektur erfolgt immer durch den jeweiligen Kurslehrer. Der Zweitkorrektor hingegen kann entweder ein weiterer Fachlehrer der Schule oder ein schulexterner Fachlehrer sein (jeweils in sechs Ländern). Die Verordnungen in drei weiteren Ländern sehen entweder einen schulinternen oder externen Zweitkorrektor vor.

Für den Fall einer starken Abweichung der Erst- und Zweitkorrektur ist eine Drittkorrektur möglich. Der Schwellenwert hierfür liegt je nach Land bei drei, mehr als drei oder mehr als vier Punkten Differenz. Die Mehrheit der Länder sieht jedoch keine Punkterege lung vor – hier entscheidet grundsätzlich der Prüfungsausschussvorsitzende (oftmals die Schulleitung, alternativ ein Mitglied der Aufsichtsbehörde oder einer anderen Schule) über die endgültige Benotung der Arbeit, der bei Bedarf die Einschätzung eines Drittgutachters einholen kann.

Anonymität der Prüflinge. Die Anonymität der Kandidat/-innen stellt ein weiteres Unterscheidungskriterium dar. In der Mehrheit der Länder ist das Verfahren vollständig personalisiert, demzufolge jedem Korrektor der Name des Prüflings und der Schule bekannt ist (neun Länder). Zwei Länder (Hamburg und Sachsen-Anhalt) wenden hingegen ein komplett anonymisiertes Verfahren mit Schülerkennungen an. Die übrigen Länder kennen eine Mischform der Verfahren über die verschiedenen Korrekturinstanzen hinweg.

2.5 Umgang mit Prüfungsergebnissen

Anteil an der Gesamtqualifikation. Insgesamt ist festzuhalten, dass die Bedeutung der zentralen Prüfungen für die Gesamtqualifikation und die Abiturdurchschnittsnote formal vergleichsweise gering ist, wenngleich davon auszugehen ist, dass die zentrale Prüfung eine ungleich höhere normierende Wirkung auf den vorgelagerten Unterricht hat. In vielen Ländern beträgt der Anteil der zentralen schriftlichen Abiturprüfung etwa 20 Prozent; in anderen liegt er – teilweise bedingt durch spezifische Fächerwahlen oder die Struktur der gymnasialen Oberstufe – bei etwa 7 bis ca. 15 Prozent.

Rückmeldung und Ergebnisverwendung. Nach Abschluss des Prüfungsverfahrens müssen die Schulen ihre Ergebnisse (i.d.R. Klausurergebnisse und Abiturdurchschnittsnoten) an die zuständige Behörde übermitteln. Von dort werden den Schulen ihre Ergebnisse zumeist im Vergleich zum Landesdurchschnitt zurückgemeldet, teilweise erhalten sie auch einen Landesbericht (mit unterschiedlich detaillierten Ergebnissen). Eine ver-

	BW	BY	BE	BB	HB	HH	HE	MV	NI1	NW	SL	SN	ST	SH ¹	TH
Rahmenbedingungen															
Zeitpunkt der Einführung	1946	1946	2007	2005	2007	2005	2007	1991	2006	2007	1945	1993	1993	2008	1990
Vorbereitung der schriftlichen Prüfung															
Anzahl der Prüfungsfächer insgesamt (schriftl. u. mündl.)	5	4	5	4-5	4	4	5	5	4	4	4	4	5	4	4
Möglichkeiten (Fächer, in denen die schriftliche Abiturprüfung abgelegt werden kann)															
In allen Fächern ²		x	x	x		x	x		x	x	x	x		x	x
Nur in ausgewählten Fächern ³	x				x			x					x		
Reichweite (Anteil der zentralen Prüfungen)															
Alle schriftlichen Prüfungen zentral	x	x					x	x	x	x	x	x	x		x
Nur bestimmte Fächer/Fächergruppen			x ⁴	x ⁵		x ⁶									
Nur bestimmte Fächer + bestimmtes Anforderungsniveau					x ⁷									x ⁸	
Grundlage der Prüfungen															
EPA + Lehrpläne		x									x	x	x		x
EPA + Lehrpläne + Schwerpunktthemen	x		x	x	x	x	x	x	x	x				x	
Zeitpunkt der Bekanntgabe von Schwerpunktthemen	~2 J.		~2 J.	1.5J.	2.5J.	~2 J.	~2 J.	~1 J.	~2 J.	2-3J.				> 2J.	
Entwicklung der Prüfungsaufgaben															
Top down-Modell		x	x		x	x	x	x				x	x		x
Bottom up-Modell	x			x					x	x	x			x	
Durchführung der Prüfung															
Organisatorisches: Zeit, Ort, Aufsicht															
Zeitpunkt: Ende der Qualifikationsphase	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Ort: jeweilige Einzelschule	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Aufsicht: Kurslehrer / schulinterne Lehrkräfte	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Anzahl der schriftlichen Prüfungen															
Anzahl der schriftlichen Prüfungen insgesamt	4	3	3	3	3	3	3	4	3	3	3	3	4	3	3
davon: Kurse mit erhöhtem Anforderungsniveau	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2
davon: Kurse mit grundlegendem Anforderungsniveau		1	1	1	1	1	1	2	1	1	1	1	2	1	1
Auswahlmöglichkeiten bzgl. der Aufgaben (aus identischem Aufgabenpool)															
Auswahlmöglichkeit für Schüler								x	x				x		x
Auswahlmöglichkeit für Lehrkräfte oder Schüler			x	x										x ⁹	

	BW	BY	BE	BB	HB	HH	HE	MV	NI1	NW	SL	SN	ST	SH ¹	TH
Fachspezifische Regelungen	x	x				x	x			x	x	x			
Dauer der schriftlichen Prüfungen (in Minuten) ¹⁰															
Kurse mit erhöhtem Anforderungsniveau			300	240		300	240		300 ¹¹	255	300		300	300	270 ¹²
Kurse mit grundlegendem Anforderungsniveau			240	180		240	180		220	180	210		210	240	21012
Fachspezifische Regelungen	x	x			x			x				x			
Korrektur															
Korrektoren															
Erstkorrektor: Kurslehrer	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Zweitkorrektor: Fachlehrer der gleichen Schule		x						x			x			x	x
Zweitkorrektor: Schulexterner Fachlehrer	x		x		x	x				x		x	x		
Drittkorrektor: Fachlehrer der gleichen Schule		x	x ²	x		x		x						x	x
Drittkorrektor: Schulexterner Fachlehrer	x						x			x		x			
Erforderliche Punktdifferenz (für Drittkorrektoreinsatz)															
3 Punkte Differenz (zwischen Erst- und Zweitkorrektor)	x														
> 3 Punkte Differenz (zwischen Erst- und Zweitkorrektor)						x						x			
≥ 4 Punkte Differenz (zwischen Erst- und Zweitkorrektor)				x						x					
keine Punktregelung (PA-Vorsitzender entscheidet)		x	x		x		x	x	x		x		x	x	x
Vorsitzender des Prüfungsausschusses (Regelfall, Ausnahmen möglich)															
Intern		x			x	x		x	x	x		x		x	x
Extern															
Extern (Mitglied der Schulaufsichtsbehörde)	x						x				x				
Variabel			x										x		
Anonymität der Prüflinge im Korrekturverfahren															
vollständig personalisiertes Prüfungsverfahren		x	x		x		x	x	x	x		x	x	x	
vollständig anonymisiertes Verfahren						x									
Anonymität der Prüflinge im Korrekturverfahren															
vollständig personalisiertes Prüfungsverfahren		x	x		x		x	x	x	x		x	x	x	
vollständig anonymisiertes Verfahren						x						x			
Mischform	x			x							x				x

Tab. 2: Innerdeutscher Vergleich (für das Schuljahr 2007/2008)

(Fortsetzung) Tab. 2

	BW	BY	BE	BB	HB	HH	HE	MV	NI1	NW	SL	SN	ST	SH ¹	TH
Lösungs- und Bewertungsvorgaben ³															
Erwartungshorizonte	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Hinweise zu Korrektur und Bewertung	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Prüfungsergebnisse															
Anteil zentraler Prüfungen an der Gesamtabiturnote ⁴	21%	21%	20%	23%	23%	21%	16%	21%	21%	21%	21%	21%	21%	14,5%	21%
Rückmeldung der Ergebnisse an die Schulen (intern)															
mind. Landesdurchschnitt (ggf. zusätzliche Quoten)	x	x	x	x	x	x	x	x		x	x	x	x	x	x
Landesbericht (unterschiedlich detaillierte Ergebnisse)			x	ggf.	x				x				x		
Veröffentlichung der Ergebnisse nach außen															
Schulporträts (Veröff. einzelstaatlicher Ergebnisse)			x	x						x		x		x	
Öffentliches (Teil-)Ranking			x												
Landesbericht (unterschiedlich detaillierte Ergebnisse)							x			x					x

1 Regelungen gültig bis 2011; ab 2011 Kernfächer (insgesamt 5 Prüfungsfächer)

2 gem. Oberstufenverordnung und vorliegende EPA

3 bedingt durch die Struktur der gymnasialen Oberstufe

4 Grund- und Leistungskurse in D, M, E, F, Rus, Span, Ital, Lat, Griech, Türk, Poln

5 Grund- und Leistungskurse Bio, Che, D, F, E, Ek, Ge, M, Phy, Politische Bildung

6 Grund- und Leistungskurse D, E, F, Sp., Lat, Gemeinschaftskunde, Mathematik, Biologie

7 Leistungskurse D, E, F, Sp, M, Bio, Che, Phy sowie die Grundkurse dieser Fächer, zusätzlich GK Geschichte und Politik

8 Leistungskurse D, E, F, Sp, Dän, Lat, Griech., M, Bio, Che, Phy (alle Schüler nehmen mindestens an einer zentralen Prüfung teil); M und E gemeinsam mit HH

9 Für Lehrer immer; für Schüler: Fachspezifische Regelungen (z. B. Deutsch ja, Mathe nein)

10 Reine Bearbeitungszeit; hinzu kommen ggf. Auswahl- und Einlesezeit, Verlängerungen auch bei Experimenten oder fachpraktischen Aufgaben (max. 60 Minuten); in Berlin ist die Auswahlzeit bereits inklusive

11 Ausnahme: Leistungskurs Sport: 240 Minuten

12 Ausnahme: Leistungskurs Deutsch = 330 Minuten; Grundkurs Deutsch: 270 Minuten

13 Regelungen gültig bis 2011; ab 2011 Kernfächer (insgesamt 5 Prüfungsfächer)

14 PA-Vorsitzender setzt Endnote fest (variabel); er kann ggf. von einem weiteren Fachlehrer (schulintern oder -extern) ein Drittgutachten anfordern

15 Länderspezifisches Vokabular

16 Annahme: So viele schriftliche Prüfungen wie möglich

pflichtende Veröffentlichung einzelschulischer Prüfungsdaten ist zumeist nicht vorgesehen. Jedoch existieren in einigen Ländern Schulporträts, in denen die Schulen ihr Ergebnis freiwillig im Vergleich zum Landesdurchschnitt eintragen können. Vier Länder bieten auch der Öffentlichkeit einen (unterschiedlich) detaillierten Landesbericht. Einzig Berlin veröffentlicht eine Bezirksliste, in der die besten Schulen sowie die erfolgreichsten Abiturient/-innen namentlich erwähnt werden (vgl. für eine Länderübersicht Lavorgna 2007).

3. Zentralisierte Prüfungsverfahren im internationalen Vergleich

In internationaler Perspektive zeigt sich ebenfalls ein differenziertes Bild mit einer Vielzahl von mehr oder weniger stark zentralisierten Prüfungssystemen (vgl. Tab. 3).

Der Vergleich der Prüfungsverfahren ist analog zum innerdeutschen Vergleich in die Kategorien der Rahmenbedingungen, Vorbereitung, Durchführung, Korrektur und des Umgangs mit Prüfungsergebnissen unterteilt. In der Analyse wird, sofern möglich, für die einzelnen Prüfungselemente jeweils beschrieben, welche Varianten von einer Mehrzahl der Länder genutzt werden. Die exemplarische Beschreibung weiterer Varianten skizziert das Spektrum der Möglichkeiten zwischen eher dezentralen und zentralen, standardisierten und weniger standardisierten Prüfungsverfahren.

3.1 Rahmenbedingungen

Ein Teil der Länder blickt auf eine längere Tradition zentraler Abschlussprüfungen zurück. Hierzu gehört z.B. Irland, wo bereits 1879 zentrale Prüfungen eingeführt wurden, um einheitliche Standards zu gewährleisten. Im ebenfalls traditionell zentral prüfenden Frankreich ist dieses Verfahren vor allem in dem Versuch begründet Chancengleichheit herzustellen. In England hingegen sind die zentralen Prüfungen in eine marktorientierte Schulentwicklung eingebunden, die u.a. über die Publikation von Prüfungsergebnissen ein Bemühen der Schulen um Qualitätsentwicklung intendiert. In den Niederlanden wurden zentrale Prüfungen im Zuge des Autonomiediskurses in den 1970er-Jahren vor allem als Gegengewicht zur hohen Autonomie der Schulen eingeführt.⁷

In anderen Ländern ist das zentrale Prüfungsverfahren dagegen noch vergleichsweise jung. Insgesamt lässt sich in den letzten Jahren ein Trend hin zu zentralen Tests und Prüfungen im Kontext einer Steuerungserwartung hinsichtlich des Bildungssystems dokumentieren. Zwischen 2000 und 2005 wurden schullaufbahnbegleitende Tests in sieben europäischen Staaten eingeführt (vgl. Europäische Kommission u.a. 2005), was in einigen Fällen auch die Einführung oder den Ausbau zentraler Prüfungsverfahren beinhaltete (z.B. in Ungarn und Polen).

⁷ Vgl. für den Vergleich von England, Frankreich und den Niederlanden van Ackeren 2003.

3.2 Prüfungsvorbereitung

Verwaltungsebene. Im Vereinigten Königreich gibt es bei einem insgesamt landeseinheitlichen Prüfungsverfahren für die *A-Level*-Prüfungen hinsichtlich der Aufgabenstellung regionale Differenzierungen: England, Wales und Nordirland haben jeweils eigene Prüfungsbehörden; die Aufgaben werden unabhängig voneinander in den *Awarding Bodies* als nichtstaatliche Non-Profit-Organisationen erstellt. Die Schulen sind regional nicht an bestimmte *Awarding Bodies* gebunden, sondern können frei wählen. Hiervon abzugrenzen sind Länder, in denen sich die Prüfungsverfahren als Ganzes auf regionaler Ebene unterscheiden (bspw. Deutschland). In den meisten hier herangezogenen Länderbeispielen legen alle Schüler/-innen – abgesehen von Auswahlmöglichkeiten (s.u.) – landesweit die gleichen Prüfungen ab.

Reichweite der Prüfungen. Die allgemeine Tendenz geht dahin, dass fast ausschließlich schriftliche Prüfungen zentral geprüft werden. In Irland und Dänemark werden auch mündliche Prüfungen zentral durchgeführt. Daneben gibt es einige weitere Länder, in denen zumindest für die Bewertung der mündlichen Prüfung ein externer Gutachter hinzugezogen wird (z.B. Norwegen, Polen).

Hinsichtlich der schriftlichen Prüfungen zeigen sich deutliche Unterschiede: In einem Großteil der Länder werden diese grundsätzlich zentral durchgeführt; allerdings wird oft nur ein Teil der Fächer schriftlich geprüft. In der Slowakei wird ein einziges Fach (Slowakische Sprache und Literatur), im englischsprachigen Raum werden alle Fächer zentral schriftlich geprüft. Allerdings ist in diesen Ländern (vor allem im Vereinigten Königreich mit einem Minimum von sechs Fächern) die Zahl der Fächer von vornherein deutlich stärker begrenzt als in anderen Ländern. Mit bis zu zwölf zentralen Pflichtprüfungen dürften die Schüler/-innen in Luxemburg den höchsten Anteil zentraler Prüfungen haben.

Prüfungsfächer. Gemeinsam haben fast alle Länder, dass mindestens die Landessprache und wenige weitere Kernfächer zentral geprüft werden; zu diesen zählen vor allem Fremdsprachen und – allerdings nicht flächendeckend – weitere geistes- sowie naturwissenschaftliche Fächer. Darüber hinaus hängt die Festlegung der Prüfungsfächer auch von der Organisation der Sekundarstufe II im jeweiligen Land ab. Vereinfacht lässt sich zwischen Ländern differenzieren, in denen die Schüler/-innen sich in der Sekundarstufe II auf eine bestimmte Fachrichtung (Naturwissenschaften, Sozial- und Wirtschaftswissenschaften, künstlerische Fächer etc.) spezialisieren, und den Ländern, bei denen die Spezialisierung auf wenige Profulfächer begrenzt ist; beide Formen sind etwa gleich stark vertreten.

In Ländern mit Spezialisierung werden in der Regel die Spezialisierungs- und Kernfächer geprüft und es gibt nur wenige Pflichtprüfungen (s.o.), die alle Schüler/-innen unabhängig von der Spezialisierung schreiben müssen. Darüber hinaus wird den Prüflingen in fast allen Ländern mindestens in einer Prüfung die Wahl des Faches überlassen, um zumindest teilweise unterschiedlichen Interessen Rechnung zu tragen.

Länder, in denen sich die Schüler/-innen nicht auf ein Fachgebiet festlegen, prüfen tendenziell eher allgemeine Kernfächer (s.o.) und ein bis zwei zusätzliche Wahl- oder

Profilfächer. Besonders offen ist die Wahl der Fächer im Vereinigten Königreich. Hier werden zwar alle belegten Fächer geprüft, da es aber keinerlei Belegpflicht – auch keine Kernfächer – gibt, wählen die Schüler/-innen zumindest indirekt ihre Prüfungsfächer selbst. Im Gegensatz dazu steht Finnland mit der striktesten Festlegung von Prüfungsfächern: Insgesamt werden vier Fächer geprüft. Die Prüfung in Finnisch ist generell vorgegeben; darüber hinaus können drei weitere Fächer aus einem Pool von vier vorgegebenen Prüfungsfächern bzw. Fächergruppen (zweite Landessprache, Fremdsprachen, Mathematik sowie eine Prüfung aus der Fächergruppe „*General Studies*“) ausgewählt werden. Zwar ist es möglich, sich in zusätzlichen Fächern prüfen zu lassen; diese Prüfungen werden jedoch auf einem gesonderten Zeugnis zusammen mit den Jahresleistungen angegeben. Dies liegt vor allem in einer möglichst gerechten Selektion beim Hochschulzugang begründet, der i.d.R. an das Prüfungsergebnis gebunden ist.

Aufgabenentwicklung. In den untersuchten Ländern werden die Prüfungen zumeist durch staatliche Einrichtungen organisiert, in der Regel durch dem Bildungsministerium untergeordnete Instanzen. Einzelne Länder verfügen über mehrere verschiedene Institutionen, die in einem nicht hierarchisch strukturierten Mehrebenensystem angeordnet und für jeweils spezifische Aspekte der Prüfungen zuständig sind (z.B. Portugal). Ein alternatives Konzept, in dem nicht staatliche Institutionen für die Konzeption und Durchführung der Prüfungen zuständig sind, verfolgen die Niederlande mit dem CITO-Institut und das UK mit seinen *Awarding Bodies*. Dies spricht für den Anspruch, die Prüfungen von einer unabhängigen Einrichtung durchführen zu lassen.

Die Entwicklung der Prüfungsaufgaben wird in den meisten Ländern vom untersten Glied der Instanzenkette vorgenommen. Mit der Konzeption der Aufgaben werden fast ausschließlich spezielle Kommissionen betraut. Eine Ausnahme zum top-down-Verfahren bildet Schottland, wo erfahrene Lehrer durch die *Scottish Qualifications Authority* mit der Aufgabenentwicklung beauftragt werden, die dann aber vor der Drucklegung mehrere Überprüfungsphasen durchlaufen.

In den anderen Ländern, in denen vornehmlich oben erwähnte Kommissionen eingesetzt werden, variiert die Zusammensetzung der Kommissionen im Ländervergleich teils erheblich. Sie können fast vollständig mit Lehrkräften (z.B. Frankreich) oder fast ausschließlich mit spezialisierten Universitätsangehörigen besetzt sein (z.B. Portugal). In Irland sind die Kommissionen Expertenteams, die aus Fachlehrern sowie jeweils einem *Subject Expert* aus dem Hochschulbereich bestehen. Die entworfenen Aufgaben durchlaufen in fast allen Staaten mehrere Schritte, in denen sie auf inhaltliche und formale Richtigkeit überprüft werden. In einigen Ländern werden sie unter Prüfungsbedingungen getestet (bspw. in Frankreich).

3.3 Durchführung

Ort und Aufsicht. Zur Gewährleistung der Vergleichbarkeit werden die Prüfungen in einigen Ländern auch unter gleichen lokalen Bedingungen organisiert. Im Bezug auf den Prüfungsort sind drei Varianten bekannt. Demnach finden die Prüfungen in den jeweili-

gen Klassen- oder Kursräumen des Prüfungsfaches statt (z.B. Portugal) oder sie werden in speziell ausgewählten Schultrakten mit gut zu überblickenden Räumen und Fluren organisiert (z.B. in Ungarn). Zu einer dritten Kategorie gehören insbesondere englischsprachige Länder: Hier finden die Prüfungen in s.g. *Examination Centres* statt. Diese können zwar auch Schulen sein, müssen aber als Prüfungszentrum genehmigt werden.

Hinsichtlich der Prüfungsaufsicht lassen sich mind. zwei verschiedene Verfahren ausmachen. Im ersten Verfahren sind die Schulen selbst für die Auswahl und Bereitstellung der aufsichtführenden Personen verantwortlich, müssen sich dabei aber nach bestimmten Vorgaben, z.B. zur Anzahl der Aufsichtspersonen und zu deren Fachexpertise, richten. In Portugal werden pro Prüfung zwei fachfremde Aufsichtführende verlangt; außerdem muss ein Fachlehrer zur Klärung von Verständnisfragen anwesend sein. In einem anderen Verfahren wird die Aufsicht von Beauftragten der jeweils für die Prüfung verantwortlichen Behörden übernommen (z.B. *Superintendents* in Irland).

Aufgabenauswahl. Während der Recherchen zum internationalen Vergleich gab es keinerlei Hinweise darauf, dass es in einem der hier betrachteten Länder die Möglichkeit für die zuständigen Lehrkräfte gibt, aus einem Aufgabenpool Aufgaben auszuwählen. Gleichwohl gibt es in fast allen Ländern Unterschiede in den Aufgaben, die die Schüler/-innen tatsächlich während der Prüfung bearbeiten. Dies sind entweder thematische Unterschiede oder aber Unterschiede im Anforderungsniveau. Zwar haben die Prüflinge der hier betrachteten Länder nicht in jedem Fach die Möglichkeit, zwischen zwei oder mehreren thematisch unterschiedlichen Aufgaben zu wählen; tendenziell wird ihnen aber in den meisten Staaten zumindest in sprachlich-literarischen Fächern eine Auswahlmöglichkeit geboten. Eine alternative Variante ist z.B. in Schottland zu finden: Hier bestehen die meisten Aufgaben in sprachlich-literarischen Fächern aus einem Pflichtteil sowie einem Wahlpflichtteil, in dem zwei Alternativen angeboten werden. In den Niederlanden werden keine thematischen Alternativen gewährt.

3.4 Prüfungskorrektur

Korrekturverfahren. Die Korrekturmodi der verschiedenen Staaten unterscheiden sich bereits hinsichtlich der Zahl der Korrekturinstanzen: So gibt es Staaten, die nur einstufig korrigieren (vor allem in den englischsprachigen Ländern), andere Staaten setzen auf Erst- und Zweitkorrektur. In einem Teil der Länder wird eine zusätzliche Drittkorrektur durchgeführt, wenn sich die Erst- und Zweitbeurteilung stark unterscheiden oder wenn gegen die Beurteilung Einspruch eingelegt wird (z.B. in Dänemark). In Luxemburg wird grundsätzlich eine Drittkorrektur durchgeführt.

Ein vollständig internes Korrekturverfahren gibt es in keinem der untersuchten Staaten und auch eine interne Erstkorrektur ist unüblich. Lediglich in Finnland wird die Erstkorrektur durch den Klassenlehrer und die Zweitkorrektur durch Mitarbeiter der staatlichen Prüfungskommission durchgeführt. In den Niederlanden wird ein Großteil der Aufgaben extern elektronisch korrigiert; offene Antwortformate werden vom Kurslehrer bewertet, die Zweitkorrektur dieser Aufgaben erfolgt durch einen Fachlehrer ei-

ner anderen Schule. Ebenfalls interne Anteile hat der Korrekturmodus in Italien: im einstufigen Verfahren ist eine Prüfungskommission für die Korrektur zuständig, die zu jeweils 50 Prozent aus Lehrern der Einzelschule und externen Mitgliedern besteht. In allen anderen Staaten liegt die Korrektur gänzlich außerhalb der Einzelschule. In einigen Staaten wird sie regional durch Prüfungskommissionen (z.B. in Polen) oder durch Fachlehrer anderer Schulen (z.B. Frankreich) vorgenommen. In anderen Nationen wird die Korrektur durch staatliche Einrichtungen durchgeführt: In Norwegen werden die Korrekturteams mit speziell für diese Aufgabe ausgewählten Lehrkräften oder Experten besetzt.

In Irland ebenso wie in Neuseeland, Schottland und dem Vereinigten Königreich durchlaufen die Korrektoren („marker“) ein *Marker Training*, in dem sie auf ihre Aufgabe vorbereitet werden. In Staaten mit zentralstaatlichen Korrekturverfahren wird darüber hinaus meist keine Zweitkorrektur durchgeführt (zur Sicherung des Korrekturstandards vgl. den Abschnitt zu Bewertungsvorgaben). Insgesamt ist folgendes Muster erkennbar: Wird die Korrektur durch eine zentrale Stelle durchgeführt, ist sie in der Regel einstufig; wird sie dagegen dezentral durchgeführt, erfolgt dies meist in mehreren Schritten.

Anonymität der Prüflinge. Das Verhältnis zwischen Ländern mit personalisiertem bzw. anonymisiertem Prüfungsverfahren ist relativ ausgeglichen. Allerdings steht die Anonymität zumindest einseitig mit der Ausgestaltung des Korrekturverfahrens in Zusammenhang: In den Staaten, in denen das Korrekturverfahren zumindest teilweise intern vorgenommen wird, ist dem internen Korrektor in der Regel auch die Identität der Prüflinge bekannt.

Bewertungsvorgaben. Um vergleichbare Korrekturen sicherzustellen, werden in allen Ländern Bewertungsvorgaben festgesetzt; überdies beinhalten die Korrekturverfahren zusätzliche Mechanismen, die Einheitlichkeit gewährleisten sollen. Insgesamt liegen zu inhaltlichen Aspekten der Bewertungsvorgaben nur für wenige Länder Informationen vor; eine Tendenz ist jedoch auszumachen: Die Bewertungsvorgaben orientieren sich in der Regel am Aufgabenformat. Enthalten die Prüfungen eher geschlossene Antwortformate (Multiple Choice, Kurzantworten), sind die Vorgaben in der Regel strenger als bei offenen Antwortformaten; für letztere werden vornehmlich mögliche Aspekte angegeben, die von den Prüflingen genannt werden sollten. Darüber hinaus ist die Punkteverteilung für die jeweiligen Teilaufgaben zumeist vorgegeben.

Besonders eng sind die Bewertungsvorgaben oftmals in den englischsprachigen Ländern; hier ist meist genau verzeichnet, wie Aufgaben bearbeitet werden müssen und wo mögliche Alternativantworten erlaubt sind. Gleichwohl werden die kriterial formulierten Bewertungsvorgaben in Schottland erst nach der Sichtung von Antworten endgültig festgelegt. Auch in Frankreich wird die Bewertung den durchschnittlichen Leistungen der Schüler/-innen angepasst. In England findet ein *Standardisation Meeting* statt, nachdem alle Prüfungen korrigiert wurden. Erst nachdem die durchschnittlichen Leistungen der Schüler/-innen bekannt sind, wird festgesetzt, welche Note für welches Punktespektrum vergeben wird. Hierzu werden die aktuellen Prüfungsleistungen u.a. mit Leistungen vorangegangener Jahrgänge verglichen. Somit hat man in einigen Län-

dem mit der Anwendungsmöglichkeit der sozialen Bezugsnorm ein Korrektiv eingeführt, das es ermöglicht, die hochstandardisierten, extern durchgeführten Bewertungsverfahren den tatsächlich gezeigten Schülerleistungen anzupassen und darüber auch den Anteil der Schulabgänger ohne Schulabschluss zu kontrollieren.

3.5 Prüfungsergebnisse

Anteil an der Gesamtqualifikation. In den meisten Ländern besteht die Endnote nur teilweise aus den in den Prüfungen erreichten Teilnoten. Meist werden die Ergebnisse formativer oder interner summativer Evaluationen zumindest gesondert angegeben (z.B. Finnland, Ungarn) oder in die Berechnung der Endnote mit einbezogen. Ist Letzteres der Fall, muss hier nach der Gewichtung der einzelnen Faktoren differenziert werden; hier lässt sich allerdings keine dominierende Variante feststellen. In Portugal z.B. kommt den Noten der Jahresleistungen mit einer siebenfachen Wertung deutlich höheres Gewicht zu als den Noten aus der zentralen Abschlussprüfung, die nur dreifach gewichtet wird. Gleichgewichtig sind die Ergebnisse der *Centraale Examens* und der *School-examens* in den Niederlanden. In Luxemburg dagegen kommt den Ergebnissen der zentralen Prüfungen mehr Gewicht zu: Hier zählen die Jahresleistungen nur einfach, die zentralen Prüfungsergebnisse dagegen zweifach. Einzig in Irland besteht die Gesamtqualifikation ausschließlich aus den Ergebnissen der zentralen *Leaving Certificate Examinations*.

Rückmeldung und Ergebnisverwendung. In England, den Niederlanden, Finnland und Frankreich – um einige Beispiele zu nennen – werden nicht nur Gesamtstatistiken, sondern auch Prüfungsergebnisse einzelner Schulen publik gemacht. In den englischen *Performance Tables* werden die Prüfungsergebnisse mit anderen Schulen sowohl regional als auch national verglichen. Auch in den Niederlanden werden die Prüfungsergebnisse einzelner Schulen veröffentlicht; dies ist durch den Druck der Presse in den 1990er Jahren entstanden. Jede Schule hat eine *Kwaliteitskaart*, die Informationen zur Schule, auch deren Abschneiden in der externen Evaluation sowie in zentralen Abschlussprüfungen, enthält.

Exemplarisch für die Staaten, in denen Ergebnisse von Einzelschulen nicht systematisch veröffentlicht werden, kann auf Irland verwiesen werden: Hier wird für eine Fächerauswahl der *Chief Examiners Report* veröffentlicht. Darin werden die Ergebnisse der Prüfungen u.a. mit den Resultaten vergangener Jahre verglichen. Darüberhinaus werden Probleme beschrieben, die bei einer großen Zahl der Schüler/-innen auftraten und Empfehlungen für Lehrer ausgesprochen. Ein ähnliches Dokument stellt beispielsweise der *Principal Assessor Report* in Schottland dar.

4. Vergleich der nationalen und internationalen Prüfungsverfahren und Diskussion

Der innerdeutsche Vergleich der Prüfungsverfahren im Zentralabitur zeigt – ungeachtet der berichteten Heterogenität – dass sich bestimmte Prüfungselemente in ihrer Ausgestaltung als weitgehend konsensfähig unter den Bundesländern erwiesen haben: Mindestens drei Viertel der Länder mit Zentralabitur prüft drei schriftliche Abiturfächer am Ende der Qualifikationsphase zu einem landeseinheitlichen Zeitpunkt unter Aufsicht der Kurslehrkraft, wobei derselbe Kurslehrer als Erstkorrektor auftritt und zur Benotung auf zentral erstellte Erwartungshorizonte und Korrekturhinweise zurückgreift. Die Ergebnisse werden den Schulen letztlich zumindest intern mit Landesvergleichswerten zurückgemeldet.

Gleichwohl überwiegen die Unterschiede. Diese zeigen sich bei – unter Standardisierungsgesichtspunkten – so zentralen Sachverhalten wie beispielsweise der Frage

- ob in allen oder nur in ausgewählten Fächern die schriftliche Abiturprüfung abgelegt wird,
- ob alle schriftlichen Prüfungen zentral erfolgen oder in Abhängigkeit von bestimmten Fächern bzw. Fächergruppen und Anforderungsniveaus,
- ob als Grundlage der Prüfungen die so genannten EPAs und Lehrpläne mit oder ohne Berücksichtigung von Schwerpunktthemen dienen,
- zu welchem Zeitpunkt vor der Prüfung entsprechende Schwerpunktthemen bekannt gegeben werden,
- aus welcher Personengruppe die Zweitkorrektoren stammen (schulintern oder -extern),
- ab welcher Bewertungsdifferenz zwischen Erst- und Zweitkorrektor ein Drittkorrektor hinzugezogen wird oder
- welchen Anteil die zentralen Prüfungselemente an der Gesamtbewertung ausmachen.

In der Zusammenschau lassen sich keine klaren Ländertypen identifizieren bzw. solche, die sich über alle beschriebenen Prüfungselemente hinweg durch einen hohen oder niedrigen Normierungsgrad auszeichnen.

Auffällig ist, dass die weitgehend konsensfähigen Prüfungselemente in den Bundesländern im berichteten internationalen Maßstab einen vergleichsweise nur geringen bis mittleren Standardisierungsgrad der zentralen Prüfungsorganisation repräsentieren. So stellen vier bis sechs schriftliche Prüfungsfächer eher den Durchschnitt der hier untersuchten OECD-Länder dar. Insbesondere im angelsächsischen Raum werden die Zentralprüfungen auch in behördlich genehmigten Prüfungszentren unter Aufsicht von behördlichem Personal durchgeführt. Erstkorrekturen werden in den OECD-Ländern überwiegend von einem schulexternen Gutachter durchgeführt. Die vollständige Anonymisierung der Prüflinge im Korrekturverfahren ist beispielsweise in Ländern wie Luxemburg, Frankreich und Irland realisiert. In Wales, Nordirland, England und Schottland wird auf eine Anonymisierung der Prüflinge verzichtet, wohl gerade weil sowohl

Land ⁸	DK	FI	FR	HU	IT	LU	NL	NO	NZ	PL	PT	RI	SC	SK	UK	
Vorbereitung der schriftlichen Prüfung																
Anzahl der Prüfungsfächer insgesamt (schrift. u. mündl.)	10+	4+	var.	var.	2	var.	var.	k.A.	var. ¹	5+	4	alle	alle	1	alle	
Möglichkeiten (Fächer, in denen die schriftliche Abiturprüfung abgelegt werden kann)																
In allen Fächern	X					X			X			X	X		X	
Nur in ausgewählten Fächern		X	X	X	X		X	X		X	X			X		
Reichweite (Anteil der zentralen Prüfungen)																
Alle schriftlichen Prüfungen zentral			X			X	X ²		X	X	X	X	X	X	X	
Nur bestimmte Fächer/Fächergruppen		X			X			X						X		
Nur bestimmtes Anforderungsniveau	X			X												
Entwicklung der Prüfungsaufgaben																
Top down-Modell	X	X	X	X	X	k.A.	X	X	X	X	X	X	X	X	X	
Bottom up-Modell													X			
Durchführung der Prüfung																
Anzahl der schriftlichen Prüfungen	var.	4	var.	var.	2	var.	var.	3	var.	3	4	5+	alle	1	alle	
Auswahlmöglichkeiten bzgl. der Aufgaben (aus identischem Aufgabenpool)																
Auswahlmöglichkeit für Schüler		X														
Fachspezifische Regelungen	k.A.	X	k.A.	k.A.	X	k.A.		X	X	k.A.	X	X	X	X	k.A.	
Korrektur																
Korrekturinstanzen	2	2	2	1	1	3	1-2 ³	1	1	1	1	1	1	1	k.A.	1

Land ⁸	DK	FI	FR	HU	IT	LU	NL	NO	NZ	PL	PT	RI	SC	SK	UK
Erstkorrektur durch															
Kurslehrer		X					X ³								
Anderer Lehrer	X		X		X									k.A.	
Übergeordnete Instanz				X	X			X	X	X	X	X	X		X
Anonymität der Prüflinge	nein	nein	ja	ja	k.A.	ja	nein	k.A.	ja	k.A.	ja	ja	nein	ja	nein
Prüfungsergebnisse															
Berechnung der Gesamtqualifikation															
Nur zentrale Prüfungsergebnisse			X									X			
Prüfungsergebnisse und Jahresleistungen gesondert	X		X							k.A.				X	
Schnitt aus Prüfungsergebnissen und Jahresleistungen	X ⁴				X	X ⁵	X ⁴	X	X ⁴		X ⁶		X		X
Veröffentlichung von Einzelschulergebnissen	k.A.	ja	ja	k.A.	k.A.	k.A.	ja	k.A.	ja	k.A.	k.A.	nein	k.A.	k.A.	ja
Abkürzungen: var.: variable Anzahl an Prüfungsfächern (z.B. in Abhängigkeit des Spezialisierungsgebietes) k.A.: es liegen keine Informationen vor															
1 Alle Fächer werden in nach Kompetenzbereichen getrennten Einzelprüfungen intern und extern geprüft															
2 Alle Prüfungsfächer werden zentral und dezentral geprüft															
3 Multiple-Choice-Fragen werden durch Computer korrigiert; offene Aufgabenformate durch ein zweistufiges Verfahren															
4 Gleichwertige Gewichtung von Prüfungsergebnissen und Jahresleistungen															
5 Prüfungsergebnisse werden stärker gewichtet als Jahresleistungen															
6 Jahresleistungen werden stärker gewichtet als Prüfungsergebnisse															
7 Länder in der Reihenfolge: Dänemark, Finnland, Frankreich, Ungarn, Italien, Luxemburg, Niederlande, Norwegen, Neuseeland, Polen, Portugal, Republik Irland, Schottland, Slowakei, Großbritannien															

Tab. 3: Internationaler Vergleich zentraler Prüfungen (Stand Schuljahr 2006/2007)

Prüfungsort, Aufsichtspersonen, Erst- und Zweitkorrektoren in der Regel schulextern angesiedelt sind (bzw. sein können).

Ein zusammenfassender Vergleich für alle Bereiche der Prüfungsorganisation findet sich in Tabelle 4. Hier ist die Prüfungspraxis, die von der Mehrheit der Länder in der BRD geteilt wird (schattierte Felder) in der Spannweite der internationalen Prüfungspraxis verortet. Die verschiedenen Prüfungselemente sind danach geordnet, ob sie einen eher hohen oder aber niedrigen Standardisierungsgrad repräsentieren.

In der Zusammenschau der Befunde wird im Übrigen auch ein weiterer Trend in einigen Staaten sichtbar, der den internen Beurteilungen mehr Gewicht einräumt: Luxemburg bezieht seit 1994 die Jahresleistungen der Schüler/-innen in die Berechnung der Gesamtqualifikation ein. In Portugal wiegt die interne Beurteilung deutlich mehr als die externe, mit dem Anspruch Chancengleichheit zu wahren, indem individuell heterogene Lernzugänge stärker berücksichtigt werden: Schüler/-innen sollen über einen längeren Zeitraum beurteilt werden, und zwar von Menschen, die sie kennen und einschätzen können. Andere Länder verzichten gänzlich auf zentral oder dezentral gestaltete Abschlussprüfungen (vgl. Tabelle 1) und sind gleichwohl in der Lage, hervorragende Ergebnisse bei internationalen Leistungsvergleichen zu erzielen (vgl. z.B. Flandern).

Es bleibt die Frage zu klären, wie zentral „das Zentralabitur“ in der BRD im internationalen Vergleich tatsächlich ist. Die Antwort fällt eindeutig aus: Die Mehrheit der Bundesländer präferiert – mit Ausnahme der Bewertungsvorgaben für die Benotung – solche Organisationsformen des Zentralabiturs, die einen im internationalen Vergleich nur geringen bis mittleren Standardisierungsgrad repräsentieren. Dies betrifft alle Bereiche der Prüfungsorganisation: Von der Vorbereitung der Prüfung bis hin zum Umgang mit den Prüfungsergebnissen.

Das Ergebnis dieser Expertise gibt Anlass zu weiterführenden Fragestellungen. Ist die Organisation des Zentralabiturs in der BRD unter dem Gesichtspunkt der angestrebten Standardsicherung möglicherweise suboptimal ausgestaltet? Bietet ein höherer Grad der Standardisierung der Prüfungselemente wirklich die Gewähr für eine bessere Standardsicherung? Und welches sind die möglichen, auch kontraproduktiven Nebenwirkungen einer verstärkten Standardisierung? Die Beantwortung dieser Fragen bewegt sich jenseits des Horizontes der hier vorgestellten Untersuchung und stellt ein Desiderat für weitere Forschungsbemühungen dar. Die nationale und internationale Forschung ergibt derzeit keine eindeutige und abgesicherte Befundlage zu den Effekten von zentralen Abschlussprüfungsverfahren auf schulische, unterrichtliche und individuelle Arbeitsprozesse und -ergebnisse. Umso weniger existieren aussagekräftige Forschungsbefunde zu den Effekten *unterschiedlicher* Standardisierungsgrade bei zentralen Abschlussprüfungen. Dass eine fortschreitende Standardisierung der Prüfungsorganisation die faktische Sicherung erzielter Leistungsstandards befördere, ist – unter Berücksichtigung des derzeitigen Forschungsstandes – bislang weder zu belegen noch zu widerlegen. Unter steuerungstheoretischen Gesichtspunkten erscheinen uns jedoch folgende Aspekte von großer Bedeutung. So werden zentrale Abschlussprüfungen – auch in der wissenschaftlichen Diskussion – häufig einzig als Steuerungsinstrument der Rechenschaftslegung bzw. externen Evaluation begriffen. Diese Charakterisierung verschlei-

		Minimum	Mittlerer Standardisierungsgrad	Maximum
Vorbereitung der Prüfung	Durchführung der Prüfung	niedriger Standardisierungsgrad	Standardisierungsgrad	hoher Standardisierungsgrad
	Prüfungsart der zentralen Prüfung	* nur schriftliche Prüfung		schriftliche und mündl. Prüfung
	Anzahl der schriftlichen Prüfungsfächer	1 Fach	* 3-6 Fächer	12 Fächer
	Entwicklung der Prüfungsaufgaben	Anforderung bei Lehrkräften	* Entwicklung in Aufgabenkommissionen	Aufgabenpilotierung
	Prüfungstermin	regional unterschiedlich	* regional bzw. bundeslandweit einheitlich	national einheitlich
	Prüfungsort	* Klassen- bzw. Kursräume	schulinterne Prüfungsräume	externes Prüfungszentrum
	Prüfungsaufsicht	* schulinterner Kurs-/ Klassenlehrer	schulinterne, kursexterne Person	schulexterne Person
	Auswahlmöglichkeiten der Schüler/-innen	in allen Prüfungsfächern gegeben	* in Abhängigkeit vom Fach und den Wahlmöglichkeiten der Lehrkräfte	in keinem Prüfungsfach gegeben
	Auswahlmöglichkeiten der Lehrkräfte	* in Abhängigkeit vom Fach und den Wahlmöglichkeiten der Schüler		in keinem Prüfungsfach gegeben
Korrekturverfahren	Korrekturinstanzen	nur Erstkorrektur	* Erst- und Zweitkorrektur, ggf. Drittkorrektur	obligatorische Drittkorrektur
	Erstkorrektor	* schulinterner Kurslehrer		schulexterner Korrektor
	Zweitkorrektor	schulinterner Fachlehrer	* schulintern oder extern	schulexterner Korrektor
	Drittkorrektor	schulinterner Fachlehrer	* schulintern oder extern	schulexterner Korrektor
	Anonymität der Prüflinge (bei Erst- und Zweitkorrektur)	* Verfahren vollständig personalisiert	Mischform	Verfahren vollständig anonymisiert
	Benotung auf der Basis von Bewertungsvorgaben	ex post definiert ¹		* ex ante definiert ²
Umgang mit Prüfungsergebnissen	Veröffentlichung der Prüfungsergebnisse	keine Veröffentlichung einzelschulischer Ergebnisse	* partielle Veröffentlichung einzelschulischer Daten	flächendeckendes öffentliches Schulranking
	Eingang der zentralen Prüfungen in die Gesamtnote	* zentrale Elemente machen weniger als 25 Prozent der Gesamtqualifikation aus	zentrale Elemente machen rund 50 Prozent der Gesamtqualifikation aus	zentrale Elemente machen 100 Prozent der Gesamtqualifikation aus

Anm.: Die mit einem Stern markierten Felder bilden die Prüfungsformen ab, die von der Mehrheit der Bundesländer (Modus) in der BRD im zentralen Prüfungsverfahren Anwendung finden.

¹ Die Bewertungsvorgaben für die Benotung werden in Abhängigkeit von den konkreten Prüfungsergebnissen erst im Nachhinein im Rahmen einer sozialen Bezugsnormierung festgelegt.

² Die Bewertungsvorgaben werden vor den Prüfungen quasi als kriterieller Standard konzipiert.

Tab. 4: Verortung bundesrepublikanischer Prüfungselemente in der Spannweite internationaler Organisationsformen zentraler Abschlussprüfungen am Ende der Sekundarstufe II

aber die möglicherweise weit reichenden direkten Steuerungseffekte einer (wie auch immer im Einzelnen gestalteten) zentralen Prüfungsorganisation

- auf den Bereich der Lernorganisation (z.B. Festlegung von Lernbereichen, unterrichtliche Differenzierung bei Schüler/-innen im gleichen Kurs mit und ohne späterer zentraler Prüfung),
- auf den Bereich der Unterstützungsangebote (z.B. zeitliche Entlastung der Lehrerschaft durch die Administration bei der Erstellung von Prüfungsaufgaben) und
- auf den Bereich der Orientierungsvorgaben für die Schulen (z.B. Schwerpunktthemen, Benotungsvorgaben und Horizonte).

Das Zentralabitur ist der wohl prägnanteste Ausdruck des derzeitig vorherrschenden schuladministrativen Steuerungsmodells „Dezentralisierung bei gleichzeitiger Rezentralisierung“ (vgl. Böttcher 2007). Es scheint strukturell als Steuerungsinstrument in besonderer Weise geeignet, die differenziellen Handlungslogiken der schulischen Akteure und deren potenzielle Verselbstständigung kontrollieren zu können. Damit erweist sich das Zentralabitur als janusköpfige, steuerungspolitische Kehrseite der größeren strukturellen Selbstständigkeit der Einzelschulen und des damit verknüpften Professionalisierungsanspruches. Die intendierten und nicht-intendierten Wirkungen dieses Spannungsverhältnisses sollten sorgfältig (und möglichst zeitnah) beobachtet und analysiert werden.

Literatur

- Ackeren, I. van (2003): Evaluation, Rückmeldung und Schulentwicklung. Erfahrungen mit zentralen Tests, Prüfungen und Inspektionen in England, Frankreich und den Niederlanden. Münster: Waxmann.
- Ackeren, I. van (2007): Zentrale Abschlussprüfungen. Entstehung, Struktur und Steuerungsperspektiven. *Pädagogik* 59, H. 3, S. 12–15.
- Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.) (2007): *Educational Governance. Handlungskoordination und Steuerung im Bildungssystem*. Wiesbaden: Verlag für Sozialwissenschaften.
- Baumert, J./Watermann, R. (2000): Standardsicherung durch die Abiturprüfung. Zentralabitur oder dezentrale Prüfungsorganisation?. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): *TIMSS III. Dritte internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Bd 2. Opladen: Leske + Budrich, S. 341–351.
- Bishop, J. (1998): The effect of curriculum-based external exit exams on student achievement. In: *Journal of Education* 29, H. 2, S. 172–182.
- Bishop, L./Wößmann, L. (2004): Institutional effects in a simple model of educational production. In: *Education Economics* 12, H. 1, S.17–38.
- Böttcher, W. (2007): Zur Funktion staatlicher Inputs in der dezentralisierten und outputorientierten Steuerung. In: Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.): *Educational Governance. Handlungskoordination und Steuerung im Bildungssystem*. Wiesbaden: Verlag für Sozialwissenschaften, S. 185–206.
- Europäische Kommission/Eurydice/Eurostat (2005): *Schlüsselzahlen zum Bildungswesen in Europa 2005*. Luxemburg: Amt Für Amtliche Veröffentlichungen Der Europäischen Gemeinschaften.

- Fuchs, T./Wößmann, L. (2007): What accounts for international differences in student performance? A re-examination using PISA data. In: *Empirical Economics* 32, H. 2, S. 433–464.
- Jürges, H./Schneider, K./Büchel, F. (2003): The effect of central examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. Working Paper No. 939. München: ifo.
- Hanushek, E.A./Wößmann, L. (2007): The role of education quality in economic growth. Washington: World Bank.
- Konsortium Bildungsberichterstattung (2006): Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. Bielefeld: Bertelsmann.
- Lavorgna, C. (2007): Zentrale Abiturprüfungsverfahren im Bundesländervergleich – Eine deskriptiv-vergleichende Bestandsaufnahme. Schriftliche Hausarbeit im Rahmen der Ersten Staatsprüfung. Universität Duisburg-Essen: unveröffentlicht.
- Maag Merki, K. (2008): Die Einführung des Zentralabiturs in Bremen – Eine Fallanalyse. In: *Die Deutsche Schule* 100, H. 3, S. 357–368.
- Schümer, G./Weiß, M. (2008): Bildungsökonomie und Qualität der Schulbildung. Kommentar zur bildungsökonomischen Auswertung von Daten aus internationalen Schulleistungsstudien. Frankfurt: GEW. Verfügbar unter: <http://www.gew.de/Binaries/Binary36403/GEW-Bioekoweb.pdf>, (letztes Abrufdatum 15.09.2008).
- Wößmann, L. (2005): Ursachenkomplexe der PISA-Ergebnisse: Untersuchungen auf Basis der internationalen Mikrodaten. Ifo-Arbeitspapier Nr. 16, München: ifo.
- Wößmann, L. (2007): Fundamental Determinants of School Efficiency Equity: German states as a Microcosm for OECD Countries. CESifo Working Paper Nr. 1981, München: CESifo.

Abstract: In almost all of the German Laender, the centrally standardized „Abitur“ (school leaving certificate and university entrance qualification) has been introduced as an instrument of standard maintenance; however, the differences in its implementation are often quite significant. The procedures are described and compared with regard to specific categories and are then, in turn, compared to 15 OECD countries with centrally standardized exit exams at the end of the upper secondary school level. This reference-guided perspective reveals the significance of the degree of standardization of exams as opposed to the concept of “centralization” with regard to differentiating analyses of the effects of control achieved through this reform element on the level of both school-related and teaching-related processes.

Anschrift der Autoren

Esther Dominique Klein, Johannes Gutenberg-Universität Mainz, Fachbereich Sozialwissenschaften, Medien und Sport, Institut für Erziehungswissenschaft, 55099 Mainz.
E-Mail: dominique.klein@uni-mainz.de

Svenja Mareike Kühn, Universität Duisburg-Essen, Graduiertenkolleg Naturwissenschaftlicher Unterricht, Schützenbahn 70, 45127 Essen.
E-Mail: svenja.kuehn@uni-due.de

Prof. Dr. phil. Isabell van Ackeren, Universität Duisburg-Essen, Fachbereich Bildungswissenschaften, Institut für Pädagogik, 45117 Essen.
E-Mail: isabell.van-ackeren@uni-due.de

Dr. phil. Rainer Block, Johannes Gutenberg-Universität Mainz, Fachbereich Sozialwissenschaften, Medien und Sport, Institut für Erziehungswissenschaft, 55099 Mainz.
E-Mail: blockr@uni-mainz.de