

Stephan Otto, Svenja Mareike Kühn

Zwischen Tradition und Innovation - Eine Analyse zentraler Prüfungsaufgaben im Fach Deutsch im nationalen Vergleich

Zusammenfassung: Zentrale Abschlussprüfungen sind seit der letzten Dekade sowohl für das Abitur als auch für den Mittleren Schulabschluss in allen Bundesländern (mit Ausnahme von Rheinland-Pfalz) implementiert (vgl. Klein et al. 2009; Kühn 2013). Gleichwohl liegen bislang nur wenige empirische und insbesondere fachspezifische Befunde zu den Rahmenbedingungen und Wirkungen zentraler Abschlussprüfungen auf schulische und unterrichtliche Prozesse und Ergebnisse vor.

Der vorliegende Beitrag greift dieses Desiderat empirischer Forschung auf und untersucht kategoriengeleitet schriftliche Prüfungsaufgaben aus zentralen Prüfungsverfahren zum Erwerb des Mittleren Schulabschlusses und des Abiturs im nationalen Vergleich ($N=159$) im retrospektiven Längsschnittdesign (2007-2011) für das Kernfach Deutsch.

Schlüsselworte: Zentrale Abschlussprüfungen, Fach Deutsch, Aufgabenkultur

Between tradition and innovation - An analysis of statewide exit exams in the core subject German in a nationwide comparison.

Abstract: State-wide exit exams are implemented in nearly all federal states of Germany by now, except for Rhineland-Palatinate. There are very few empirical studies for the time dealing with the different framework conditions in the federal states and the possible effects of statewide exit examination on teaching.

In the core subject German there are just normative suggestions that statewide exit exams would establish a hidden canon of tasks and contents but empirical evidence for these assumptions are still a desideratum.

This contribution takes up the desideratum and analyses exams for the *Mittlerer Schulabschluss* and the *Abitur* ($N=159$) of all federal states in a longitudinal analysis (2007-2011) to examine the normative suggestions about statewide exit examinations.

Keywords: statewide exit exams, subject German, performance tasks

1. Einführung

Im Kontext der Diskurse über die Entwicklung schulischer Qualität und die Sicherung der Gleichwertigkeit schulischer Zertifikate haben in der letzten Dekade alle Bundesländer (mit Ausnahme von Rheinland-Pfalz) zentrale Prüfungen zum Erwerb von Schulabschlüssen am Ende der Sekundarstufen I und II eingeführt. Im Rahmen zentraler Abschlussprüfungen bearbeiten alle Schülerinnen und Schüler eines Bundeslandes in bestimmten Fächern zum gleichen Zeitpunkt am Ende des jeweiligen Bildungsganges die gleichen schriftlichen, landesweit einheitlich vorgegebenen Aufgaben bzw. von der Kurslehrkraft oder durch die Prüflinge selbst ausgewählte Aufgaben aus einem identischen Aufgabenpool. Betrachtet man jedoch die länderspezifischen Prüfungsverfahren, so wird deutlich, dass sich diese in ihrer jeweiligen Ausgestaltung von der Prüfungsvorbereitung über die Prüfungsdurchführung bis zum Umgang mit Prüfungsergebnissen zwischen den Ländern erheblich unterscheiden (vgl. Klein et al. 2009; Kühn 2013).

Neben diesen strukturellen Veränderungen stehen auch fachinhaltliche und fachdidaktische Innovationen im Fokus des Interesses. So sehen die Befürworter zentraler Prüfungen diese Form der Prüfungsorganisation u.a. als probates Steuerungsinstrument an, um flächendeckend fachinhaltliche Neuerungen und innovative Aufgabenformate einführen zu können. Diesen positiven Steuerungswirkungen steht allerdings die Befürchtung gegenüber, zentrale Prüfungen seien gerade hinsichtlich der Implementation von Innovationen hinderlich, da sie lediglich auf einen begrenzten Pool von Inhalten und Aufgaben zurückgriffen, die immer wieder zur Anwendung kämen (Bekes 2010; Kühn 2010). Gleichwohl hatte diese Diskussion in der Vergangenheit primär normativen Charakter und ist erst in jüngerer Zeit in den Fokus empirischer Forschung gerückt.

Die Durchsicht der Forschungsliteratur zeigt, dass bislang erst wenige entsprechende Analysen von Aufgabenstellungen im Rahmen zentraler Prüfungsverfahren für das Kernfach Deutsch vorliegen, die allerdings auf die Abiturprüfung beschränkt sind (Kammler/Noack 2010) und nicht die zentralen Prüfungen zum Erwerb des Mittleren Schulabschluss mit einbeziehen. Empirische Forschungsarbeiten zu dieser Thematik sind bis dato auf die mathematisch-naturwissenschaftlichen Unterrichtsfächer beschränkt (Kühn/Drüke-Noe 2013; Kühn 2010; Roppelt et al. 2008).

2. Theoretischer Hintergrund

Die nahezu flächendeckende Einführung zentraler Abschlussprüfungen zog vielfältige Diskussionen innerhalb der germanistischen Fachdidaktik nach sich (vgl. z. B. *Der Deutschunterricht*-Schwerpunktheft 1/2010 ‚Zentralabitur‘). Insgesamt wird angenommen, dass sich die deutlichsten Wirkungen zentraler Prüfungen auf der Ebene des Unterrichts entfalten. Im Schwerpunkt drehen sich die Kontroversen insbesondere um die Annahme einer thematischen Engführung des Unterrichts hinsichtlich der prüfungsrelevanten Themengebiete (Abraham 2010; Bekes 2010; Leonord 2009), eine vergleichsweise geringere Schülerorientierung im Unterricht (ebd.) sowie die Marginalisierung fachinhaltlicher (z. B. Gegenwartsliteratur) und fachdidaktischer (z. B. innovative Aufgabenformate) Neuerungen. Dieser s.g. *backwash*-Effekt (Cheng/Curtis 2012; Prodromou 1995) zentraler Prüfungen ist insbesondere im anglo-amerikanischen Raum schon vielfach untersucht worden. In den Studien konnten u.a. eine starke Einengung der Inhalte, das bloße Einüben prüfungsrelevanter Aufgabenformate und sogar korrumpierende Maßnahmen für das Erreichen möglichst guter Testergebnisse nachgewiesen werden (z. B. Nicols/Berliner 2007). Allerdings sind diese internationalen Befunde auf Grund der Strukturdifferenzen der unterschiedlichen Prüfsysteme nicht ohne weiteres auf den bundesdeutschen Kontext übertragbar (Klein/van Ackeren 2012). Im nationalen Kontext ist die Forschung zu den unterrichtsbezogenen Wirkungen zentraler Abschlussprüfungen bislang auf das Abitur (und nicht auf Prüfungsverfahren am Ende der Sekundarstufe I) konzentriert (Block et al. 2011; Kahnert et al. 2011; Maag Merki 2012) – gleichwohl fehlen insbesondere fachspezifische, vertiefende Analysen.

Für das Fach Deutsch ist – neben der Erfassung unterrichtlicher Effekte – insbesondere die konkrete Ausgestaltung der Aufgaben in zentralen Prüfungsverfahren relevant, da angenommen wird, über die Aufgabenstellungen zeitnah und großflächig (fach-) didaktische und -methodische sowie curriculare Neuerungen im Unterricht implementieren zu können. In diesem Kontext ist insbesondere die Frage nach einem literarischen Kanon ein kontrovers diskutiertes Thema innerhalb der germanistischen Fachdidaktik, welchem durch die Implementierung zentraler Abschlussprüfungen neue Aktualität verliehen wurde. Nachdem vor allem nach 1968 ein literarischer Kanon als ‚Herrschaftsinstrument‘ kritisiert und sich vehement

gegen eine verbindliche Vorgabe lesenswerter Literatur gewehrt wurde (Schmidt 2009; Erlinger 2002) erhielt diese Diskussion mit der Einführung zentraler Prüfungen und verbindlicher Lektürevorgaben erneut Auftrieb (Kammler/Noack 2010). Ob zentrale Prüfungen tatsächlich zu einer Veränderung eines „faktischen“ bzw. eines „heimlichen“ Kanons geführt haben, ist bislang noch nicht systematisch und umfassend erhoben worden. Einschätzungen bezogen auf einzelne Bundesländer, wie Bekes (2010) und Freudenberg (2010) sie für Nordrhein-Westfalen und Thüringen vornehmen, betonen, dass es zumindest teilweise zu einer Rekanonisierung bestimmter Autoren und Epochen kommt, während andere bestimmte literarische Strömungen, wie etwa die Gegenwartsdramatik, kaum noch Berücksichtigung fänden (Bekes 2010). Die Zusammenschau von Kammler und Noack (2010), welche die Lektürevorgaben der Abiturprüfungen der Jahre 2008 bis 2010 bundesweit erfasst, macht deutlich, dass durchaus eine Fokussierung auf bestimmte Autoren und Epochen im Zentralabitur nachweisbar ist. Der Befund wird allerdings dahingehend relativiert, dass trotz einer Fokussierung auf klassische Epochen und Autoren deutscher Literatur die seit den 1990er Jahren stattfindende Öffnung des Schulkanons für aktuelle Gegenwartsliteratur auch nach Einführung des Zentralabiturs nachweisbar ist (ebd., 2010).

Neben der Kanondiskussion werden zudem Aspekte hinsichtlich des Einsatzes nichtfiktionaler Texte im Kontext zentraler Prüfungen kritisch diskutiert. So rückten spätestens seit der ersten PISA-Erhebung Sachtexte stärker in den Fokus des Interesses, da bei PISA „informatives Lesen als Basiskompetenz für die Teilhabe des Individuums in modernen Gesellschaften“ (Maiwald 2004) erhoben werden sollte und erstmals an der absoluten Dominanz literarischer Texte im Deutschunterricht gerüttelt wurde (ebd.). Becker-Mrotzeck und Kepser (2010) merken gleichwohl an, dass trotz dieser Trendwende seit PISA pragmatische Texte innerhalb zentraler Prüfungen noch immer eine untergeordnete Rolle gegenüber literarischen Texten einnehmen. Eine solche Tendenz wird vor allem deshalb von den Autoren kritisch gesehen, weil dies Konsequenzen für nachfolgende Bildungseinrichtungen nach sich ziehe, da etwa Studierende mangelnde Kompetenzen beim Verfassen wissenschaftlicher Texte aufwiesen (Becker-Mrotzeck/Kepser 2010). Dass der verstärkte Einbezug nichtliterarischer Texte durchaus gefordert wird, spiegelt sich auch in den bundesweit gültigen Vorgaben wider, da etwa die Einheitlichen Prüfungsanforderungen für die Abiturprüfung (im Folgenden EPA) fordern, dass „Pragmatische Texte [...] verstärkt einbezogen [werden]“ (KMK EPA 2002, 4).

Eine weitere Kontroverse innerhalb der Deutschdidaktik im Kontext zentraler Prüfungen besteht hinsichtlich der Aufgabekultur. Dabei gilt es zu klären, inwieweit Elemente einer postulierten ‚neuen‘ Aufgabekultur (Köster 2003; Schmellentin 2012) in zentralen Prüfungen implementiert sind. Hierzu lassen sich insbesondere die Forderung nach einer Kombination von geschlossenen, halboffenen und offenen Aufgabenformaten (Groeben 2005) sowie die Interpretation von nichtkontinuierlichem Material (z. B. Bilder, Abbildungen) zählen (Franke 2005). Vor dem Hintergrund erwarteter normierender Steuerungswirkungen der zentralen Prüfungsaufgaben auf den vorgelagerten Unterricht (Büchter 2007) gilt es im Kontext der Prüfungsaufgaben zu prüfen, ob ausschließlich auf traditionelle Aufgabentypen im Sinne des klassischen Interpretationsaufsatzes zurückgegriffen wird oder ob sich auch innovative Aufgabenformate finden lassen, wie es innerhalb der bundesweit gültigen Vorgaben (EPA, Bildungsstandards für den Mittleren Schulabschluss) sowie aus Sicht der Fachdidaktik (Fix 2006) gefordert wird. Ebenso gilt es zu klären, ob weitere innovative Elemente, wie etwa der Einbezug eines im vorgelagerten Unterricht erstellten Portfolios, welche die gerade für das Fach Deutsch bestehende Bewertungsproblematik versuchen einzudämmen, Eingang in die zentralen Prüfungen finden (Abraham 2010).

Neben den jeweils spezifischen Vorteilen, die den einzelnen aufgabenbezogenen Innovationen zugesprochen werden, wird innovativen Aufgaben im Kontext einer veränderten Aufgabenkultur insgesamt ein großes Potential zur Qualitätssteigerung des Unterrichts, der Lernkultur und der Leistungsmessung zugeschrieben (Eikenbusch 2001). Bezogen auf die Abiturprüfung in den naturwissenschaftlichen Unterrichtsfächern konnte Kühn (2010) allerdings nachweisen, dass hier weitgehend auf die traditionelle Aufgabenkultur zurückgegriffen wird und innovative Aspekte der Aufgabengestaltung nahezu keine Berücksichtigung finden. Inwieweit diese Befunde auf das Fach Deutsch übertragbar sind und inwiefern sich Differenzen zwischen der Aufgabengestaltung der Prüfungsverfahren zum Erwerb des Mittleren Schulabschlusses und des Abiturs zeigen, stellt bis dato ein bedeutendes Desiderat fachdidaktischer Forschung dar.

Ob und inwieweit zentrale Abschlussprüfungen im Kernfach Deutsch den Anforderungen der bundesweit gültigen Vorgaben für die Sekundarstufe I und II (Bildungsstandards, Einheitliche Prüfungsanforderungen) gerecht werden, indem fachdidaktische Neuerungen umgesetzt werden, ist bislang nicht empirisch überprüft worden.

3. Ziele, Fragen und Design des Forschungsprojekts

3.1 Ziele und Fragestellungen

Das nachfolgend dokumentierte Forschungsprojekt greift die (an dieser Stelle nur knapp skizzierten) Desiderata empirischer Bildungsforschung auf und betrachtet Prüfungen zum Erwerb des Mittleren Schulabschlusses und des Abiturs aus allen zentral prüfenden Bundesländern im retrospektiven Längsschnittdesign (2007-2011) vor dem Hintergrund der bundesweit gültigen Prüfungsstandards (Bildungsstandards, EPA), um zu prüfen, ob und inwieweit die schriftlichen Prüfungsaufgaben innovative inhaltliche und formale Elemente (jenseits von ‚Klassikerkanon‘ und dem traditionellen Interpretationsaufsatz) aufweisen. Dabei gilt insbesondere dem Vergleich beider Abschlussprüfungsverfahren ein besonderes Interesse, um mit Blick auf den Mittleren Schulabschluss zu prüfen, ob hier möglicherweise eine Orientierung an der traditionellen Abituraufgabenkultur stattfindet oder ob mehr Wagnis bei der Aufgabenauswahl eingegangen wird, da dieses Zertifikat innerhalb der Öffentlichkeit als nicht so prestigeträchtig erachtet wird wie die Abiturprüfung (Kühn 2013).

Im Kontext der benannten Forschungsdesiderata ergeben sich folgende Forschungsfragen:

- (1) Inwieweit wird im Kontext zentraler Abschlussprüfungen auf einen bestimmten Grundbestand literarischer Texte (Stichwort: ‚Kanon‘) zurückgegriffen?
- (2) Werden die bundesweit gültigen Vorgaben (EPA, Bildungsstandards) hinsichtlich der fachdidaktischen Forderungen nach einem Einbezug innovativer Aufgabenmerkmale umgesetzt?
- (3) Zeigen sich Unterschiede in der Ausgestaltung zwischen den Prüfungen für den Mittleren Schulabschluss und denen für das Abitur und variieren die Befunde zwischen den einzelnen Prüfungsjahrgängen und Bundesländern?

3.2 Entwicklung eines Kategoriensystems

Um empirisch abgesicherte Antworten auf die zuvor vorgestellten Forschungsfragen geben zu können, wurde ein Kategoriensystem zur Analyse von Prüfungsaufgaben im Fach Deutsch entwickelt (Brosius et al. 2009). Dabei galt es zu berücksichtigen, dass sowohl die konkreten Aufgabenstellungen als auch die jeweiligen Textgrundlagen aus insgesamt 159 Abschlussprüfungen untersucht werden müssen.

Der erste Teil der kategoriengeleiteten Analyse bezieht sich auf die Frage nach einer möglichen Kanonisierung durch zentrale Abschlussprüfungen. Hier soll neben einer möglichen Fokussierung auf bestimmte Textsorten, Gattungen und Epochen insbesondere der Frage nach einer „faktischen Kanonisierung“ (Abraham 2008) bestimmter Autoren und Werke durch zentrale Abschlussprüfungen nachgegangen werden. Eine Durchsicht der länder-spezifischen Vorgaben für die zentralen Prüfungen im Zuge der Entwicklung des Kategoriensystems hat deutlich gemacht, dass in den Ländern Lektürevorgaben für das Abitur (alle Länder bis auf Bayern, Sachsen-Anhalt und Thüringen), umfangreiche Lektürelisten für die Sekundarstufen (Baden-Württemberg und Sachsen-Anhalt) sowie verbindliche Ganzschriften für die Prüfung zum Mittleren Schulabschluss (Baden-Württemberg, Saarland, Sachsen) zu finden sind. Diese länderinternen Vorgaben für Ganzschriften, die durch eine Analyse prüfungsrelevanter Dokumente der Bundesländer (z. B. Prüfungsordnungen) ermittelt werden konnten, wurden miteinander verglichen, um hieraus eine bundesweite Übersicht von Autoren und Werken zu erstellen, die mindestens in den Vorgaben von zwei Ländern benannt worden sind. Die auf diese Weise erstellten Autor- und Werklisten dienen dann als Grundlage für die Kategorien zu Werken und Autoren, um entscheiden zu können, wie häufig sich Autoren (insgesamt 63 Autoren; u.a. Brecht, Dürrenmatt, Grass, Goethe) und Werke (insgesamt 81 Werke; u.a. Leben des Galilei, Faust I, Woyzeck) in den Prüfungen nachweisen lassen.

Grundlage dieses ersten Analyseaspekts sind damit die jeweiligen Textgrundlagen, wobei sowohl in der Prüfung vorliegende Texte bzw. Textauszüge als auch die im vorgelagerten Unterricht behandelten Texte (z. B. Ganzschriften), die zur Bearbeitung der Prüfungsaufgabe mit einbezogen werden müssen, für die Analyse berücksichtigt werden. Kurze Zitate, die sich nicht auf ein im Vorfeld behandeltes Text beziehen und Einzelsätze etc., die als Impuls zur Beantwortung einer Aufgabe dienen sollen, werden hingegen nicht als Textgrundlage berücksichtigt. Insgesamt wurden im Rahmen der hier dokumentierten Studie 525 Textgrundlagen analysiert.

Der zweite Teil des Kategoriensystems wurde entwickelt, um die konkreten Aufgabenstellungen der Prüfungen angemessen analysieren zu können. Um erfassen zu können, was traditionelle bzw. innovative Aufgaben im Fach Deutsch sind, galt es zunächst vor dem Hintergrund fachdidaktischer Diskussionen zu klären, welche aufgabenbezogenen Innovationen im Fach Deutsch diskutiert werden und welche Aufgaben schon lange im Fach Deutsch zur Anwendung kommen. Diesbezüglich zeigt sich, dass vor allem verschiedene Aufsatzformen, denen bestimmte Textmuster zugrunde liegen, als traditionelle Aufgaben des Deutschunterrichts eingeschätzt werden (Fix 2006). Traditionelle Aufsatzformen zeichnen sich dadurch aus, dass sie innerhalb des Deutschunterrichtes schon lange zur Anwendung kommen und bis heute eine wichtige Rolle in der Schule spielen (Becker-Mrotzeck/Böttcher 2006). Definiert wird ein solcher Aufsatz etwa als „ein vom Schüler in einer vorgeschriebenen Zeit geschriebener Text, der bestimmten Anforderungen in Bezug auf Form, Stil und Inhalt unterliegt“ (Fix 2006, 90). Neben der fiktionalen Erzählung, die mittlerweile im Kontext des kreativen Schreibens schon zu den innovativen Schreibaufgaben gerechnet wird, lassen sich zu meist vier verschiedenen Textsorten ausmachen, die als traditionelle Aufsatzformen be-

zeichnet werden: Der Bericht, die Schilderung, die Beschreibung sowie die Abhandlung bzw. die Erörterung (Ludwig 1996, 240). Eine solche Unterscheidung mag zwar auf den ersten Blick unvollständig erscheinen, da die wohl prominentesten Aufsatzformen der Sekundarstufe, die Analyse und Interpretation, nicht genannt werden. Hier bietet etwa die Erklärung von Fix (2006) Aufschluss, der erklärt, dass komplexere schulische Schreibaufgaben, wie etwa die Textanalyse auf mehrere Elemente der zuerst genannten Aufsatzarten zurückgreife, da sie aus beschreibenden, berichteten sowie argumentativen Elementen bestehe (ebd. 105). Somit zählen auch diese Schreibaufgaben zu den traditionellen Aufgabenformen des Deutschunterrichts. Solchen traditionellen Aufsatzformen nach Definition des Kategoriensystems liegt auch eine traditionelle Materialgrundlage zugrunde: entweder ein fiktionaler bzw. nichtfiktionaler Lesetext, ein Verweis auf einen im Vorfeld behandelten Text oder ein kurzer textueller Stimulus (z. B. Zitat, ein kontroverser Standpunkt etc.).

Eine Vielzahl fachdidaktischer Innovationen hat sich gerade in Abgrenzung zu traditionellen Aufsatzformen und deren textuellen Materialgrundlagen entwickelt, da aufgrund der starken Normierung durch zugrundeliegende Textmuster der eigene Gestaltungsspielraum des Schreibenden bedeutend eingeschränkt und eine solche Textproduktion lebensweltlich nicht relevant sei (Fix 2006). Ausgehend von dieser Kritik etablierten sich sowohl hinsichtlich der Materialgrundlagen als auch bezogen auf die zu produzierenden Textsorten weitere Formate. Nach Durchsicht aller vorliegenden Prüfungsaufgaben können die Anwendung medial vermittelter Texte (z. B. Film- oder Hörspielsequenzen, die über eine Kombination von Seh- und Höreindruck bzw. den Höreindruck rezipiert werden müssen), der Einbezug nicht-kontinuierlicher Materialgrundlagen (Grafiken, Bilder etc.) sowie der Gebrauch eines Portfolios zu den materialbezogenen Innovationen gezählt werden (Franke 2005). Bei Portfolios handelt es sich um die erst in wenigen Ländern etablierte Praxis, eine im Vorfeld durch den Prüfling zum Schwerpunktthema erstellte Materialsammlung aus eigenständig formulierten sowie selbstständig gesammelten Texten in der Prüfung zuzulassen, die den Vorteil bieten kann, einen Rückschluss auf nicht in der Prüfung abgerufene Wissensbestände zuzulassen (Abraham 2010). Zu den innovativen Schreibaufgaben zählen das kreative Schreiben, , gestaltende Aufgaben (Beantwortung einer Aufgabe in nicht textueller Form, z. B. in Form einer mind-map), Aufgaben mit Bezug zur Lebenswelt der Schülerinnen und Schüler sowie kombinierte Texterschließungsformen. Diese zeichnen sich dadurch aus, dass die Interpretation einer Textgrundlage oder die Erörterung einer strittigen Frage nicht im Sinne der traditionellen Form verfasst wird, sondern in Form eines gestaltenden Textes. Eine typische Aufgabe dieser Art wäre etwa das Verfassen eines Tagebucheintrages (Abraham/Kepser 2009).

Für die Aufgabenanalyse stellte sich in diesem Kontext die Problematik, dass innerhalb der germanistischen Fachdidaktik keine allgemeingültige Aufgabenterminologie nachweisbar ist. Anknüpfend an den in bisherigen Forschungsarbeiten zur Aufgabenanalyse etablierten „engen Aufgabenbegriff“ (Jordan et al. 2006; Neubrand 2002) wird daher im Rahmen dieser Studie die Analyseeinheit der *Handlungsaufforderung* gewählt. Diese kann definiert werden als kleinste zusammenhängende inhaltliche Aufforderung an den Bearbeiter einer Aufgabe, die eine Sinneinheit bildet. Handlungsaufforderungen treten üblicherweise entweder als Verben in der Imperativform (z. B. *Analysieren Sie [...]*; *Interpretieren Sie [...]*; *Kreuzen Sie [...]* an), als Fragen (z. B. *Warum ist die Rede im Konjunktiv gestaltet?*), als Substantivierung (z. B. *Textinterpretation*) oder als implizite Arbeitsanweisung (z. B. *Aufgabe: Texte beschreiben*) auf. Insgesamt wurden im Rahmen dieser Studie 2872 Handlungsaufforderungen analysiert.

3.3 Methodisches Vorgehen

Die Entwicklung des Kategoriensystems orientierte sich an inhaltsanalytischen Verfahren (Mayring 2000) und etablierten Forschungsarbeiten zum Umgang mit Kategoriensystemen (z. B. Seidel 2003). Zur Sicherung der Qualität des Kategoriensystems wurden verschiedene Maßnahmen durchgeführt; dazu gehörte insbesondere ein intensives, mehrstufiges Training der an der Aufgabenanalyse beteiligten Beurteiler (Langer/Schulz von Thun 2007). Um festzustellen, inwieweit diese hinreichend trainiert sind, wurde die Beurteilerübereinstimmung berechnet (Wirtz & Caspar 2002), die nach Abschluss des Trainings für alle Kategorien Werte von $\kappa \geq .75$ und eine prozentuale Übereinstimmung zwischen beiden Beurteilern von wenigstens 90 Prozent aufwies. Damit ist das entwickelte Kategoriensystem ein verlässliches Erhebungsinstrument zur Analyse von Prüfungsaufgaben und deren Textgrundlagen zum Erwerb des Mittleren Schulabschlusses und des Abiturs im Fach Deutsch. Die abschließende Analyse wurde von zwei Beurteilern durchgeführt.

4. Empirische Befunde

4.1 (K)ein Kanon in zentralen Abschlussprüfungen

Die empirische Analyse der zentralen Prüfungsaufgaben konnte in einem ersten Untersuchungsschritt aufzeigen, dass innerhalb der Bearbeitungsgrundlagen ($N=525$) fiktionale Texte (59,0%) gegenüber nichtfiktionalen Bearbeitungsgrundlagen (27,6%) überrepräsentiert sind (die fehlenden Prozentwerte entfallen auf analysierte Prüfungen, denen keine textuelle Bearbeitungsgrundlage – etwa nichtkontinuierliches Material in Form von Bildern oder Grafiken – zu Grunde lag).

Dieser Befund verdeutlicht, dass die in den bundesweiten Vorgaben geforderte ausgewogene Verteilung der beiden Textarten oder gar der in den EPA angesprochene verstärkte Einbezug nichtfiktionaler Texte (KMK, EPA 2002) keinesfalls realisiert wurde und eine eindeutige Fokussierung auf fiktionale Texte nachweisbar ist.

Auch die differenzierte Betrachtung der literarischen Texte ($N=305$) nach den jeweiligen Gattungen machte deutlich, dass eine Präferenz für bestimmte Gattungen nachweisbar ist. So zeigte sich, dass epische Texte am häufigsten nachweisbar waren (46,9%), gefolgt von lyrischen Texten (33,8%). Die Dramatik nahm gegenüber diesen Textgattungen eine untergeordnete Rolle ein (19,3%).

Bezüglich der nur marginalen Bedeutung dramatischer Texte zeigten sich deutliche Unterschiede zwischen den verschiedenen Abschlussprüfungen, weil ein nur geringer Einbezug der Dramatik bei den Mittleren Schulabschlussprüfungen stärker ausgeprägt ist als bei den Abiturprüfungen, in denen Lyrik und Dramatik beinahe gleich häufig nachweisbar sind.

Tab.1: Häufigkeiten literarischer Gattungen – Abitur ($N=175$) und MSA ($N=130$)

Gattung	absolute Häufigkeit		relative Häufigkeit	
	MSA	Abitur	MSA	Abitur
Epik	70	73	53,8%	41,7%
Lyrik	50	53	38,5%	28,0%
Dramatik	10	49	7,7%	30,3%

Auch in einem Großteil der untersuchten Bundesländer war eine solche Verteilung nachweisbar, sodass die von Seiten der Fachdidaktik monierte Vernachlässigung dramatischer Texte (Kammler/Noack 2010) sich bundesweit zu bestätigen scheint.

Einen weiteren Untersuchungsaspekt stellte die Verteilung der Epochen der literarischen Texte dar, wobei sich hier wiederum zeigte, dass einzelne Epochen bevorzugt einbezogen werden. Insgesamt ist die Literatur der Jahrhundertwende (16,1%) sowie die Gegenwartsliteratur (15,4%) eindeutig am häufigsten nachweisbar, wobei sich dieser Befund zwischen den Schulabschlüssen unterscheidet. Für das Abitur bzw. den Mittleren Schulabschluss bilden jeweils andere Epochen einen Schwerpunkt, der bei den MSA-Prüfungen eher auf moderner Literatur liegt und bei den Abiturprüfungen eher auf weiter zurückliegenden literarischen Epochen. Der Befund von Kammler und Noack (2010), dass ältere literarische Epochen beinahe überhaupt nicht zur Anwendung kommen, bestätigt sich auch bei dieser Untersuchung, da die Literatur der Antike (0,4%), des Mittelalters (0 %) und des Barock (1,9%) die am wenigsten thematisierten Epochen darstellen.

Im Vergleich der einzelnen Bundesländer ließ sich nachweisen, dass diese bis auf wenige Ausnahmen auf ein recht breites Spektrum unterschiedlicher literarischer Epochen zurückgriffen. Die bisherigen Befunde zu den Textgrundlagen zentraler Prüfungen konnten also auf allen Untersuchungsebenen aufzeigen, dass grundsätzlich bestimmte Inhalte innerhalb der Prüfungen verstärkt einbezogen werden, während anderen eine untergeordnete Bedeutung zukommt.

Der letzte Untersuchungsaspekt widmete sich der strittigen Frage, ob sich für zentrale Prüfungen ein bundesweiter Grundbestand an Autoren und Werken nachweisen lässt. Dabei zeigten sich deutliche Unterschiede zwischen Abitur- und MSA-Prüfung: In den Abiturprüfungen wird wesentlich häufiger auf Autoren zurückgegriffen, die als kanonisch charakterisiert wurden als in den Prüfungen zum Erwerb des Mittleren Schulabschlusses.

Tab. 2: Häufigkeiten (nicht)kanonischer Autoren – Abitur (N=175) und MSA (N=130)

	absolute Häufigkeit		relative Häufigkeit	
	MSA	Abitur	MSA	Abitur
nichtkanonischer Autor	91	65	70,0%	37,1%
kanonischer Autor	39	110	30,0%	62,9%

Betrachtet man diese als kanonisch geltenden Autoren insgesamt sowohl für die Mittleren Schulabschluss als auch für die Abiturprüfungen, zeigt sich, dass klassischen Autoren der deutschen Literatur par excellence, wie beispielsweise Goethe, Schiller, Brecht, Eichendorff und Büchner, die quantitativ größte Bedeutung haben. Mit Rückbezug zu den literarischen Epochen, in denen die Schaffensphase dieser Autoren lag, zeigt sich, dass die am häufigsten nachweisbaren Autoren überwiegend dem 18. und 19. Jahrhundert entstammen und lediglich Bertolt Brecht als Autor des 20. Jahrhunderts nachweisbar ist. Autoren der Gegenwartsliteratur (u.a. Bernhard Schlink, Daniel Kehlmann und Jurek Becker mit je drei Textgrundlagen) lassen sich dagegen wesentlich seltener nachweisen (vgl. Abbildung 1).

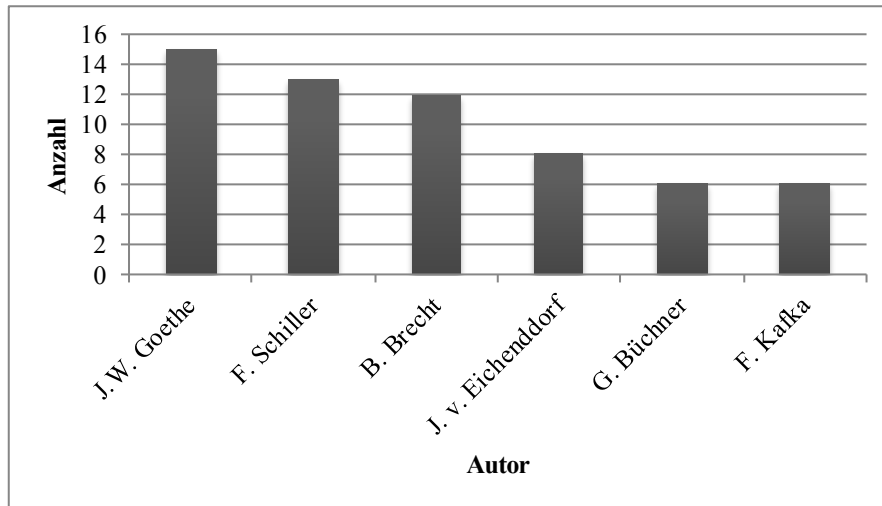


Abb. 1: Autorenkanon in zentralen Abschlussprüfungen (MSA und Abitur) (N=149)

Innerhalb der als kanonisch kategorisierten Autoren zeigt sich also eindeutig, dass hier starke Tendenzen zu einem „Klassiker-Kanon“ der deutschen Literatur nachweisbar sind.

Die Untersuchung der literarischen Textgrundlagen konnte schließlich aufzeigen, dass auch diesbezüglich ein bestimmter Grundbestand an Werken nachweisbar ist, die vermehrt in den zentralen Abschlussprüfungen Anwendung fanden (Abbildung 2). So ließen sich insgesamt 88 Werke kanonischer Autoren in der zugrunde gelegten Übersicht nachweisen, wobei Büchners *Woyzeck* (5 Textgrundlagen) der am häufigsten vorkommende Text in den zentralen Abschlussprüfungen ist, gefolgt von Horváths „Geschichten aus dem Wiener Wald“ und Schillers „Don Carlos“ sowie „Die Räuber“ (je 4 Textgrundlagen). Dies bestätigt den Befund der Zusammenschau von Kammler und Noack (2010), die ebenfalls nachwiesen, dass Büchners *Woyzeck* der am häufigsten vorkommende Text im bundesweiten Vergleich zentraler Abiturprüfungen ist. Interessanterweise lässt sich der Einbezug dieses Werkes allerdings ausschließlich in den Abiturprüfungen nachweisen, nicht in denen zum Erwerb des Mittleren Schulabschlusses.

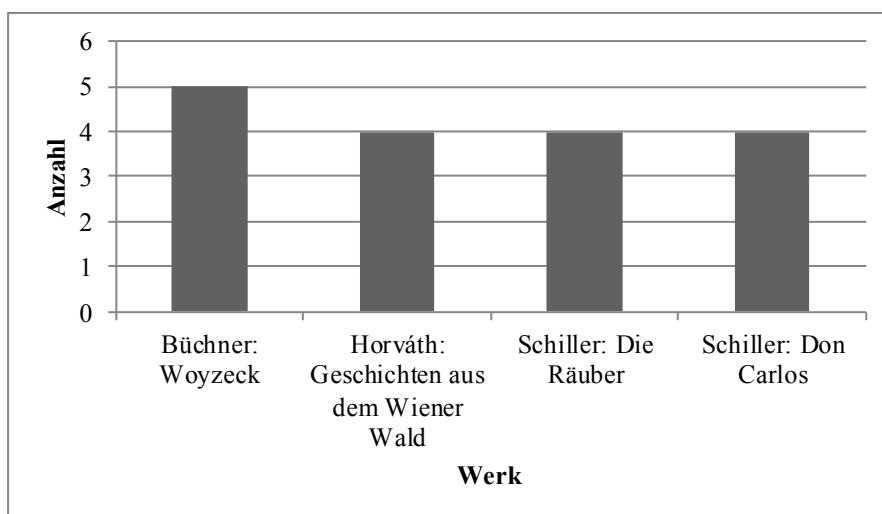


Abb. 2: Werkkanon in zentralen Abschlussprüfungen gesamt (MSA und Abitur) (N=88)

Es wird deutlich, dass die am häufigsten nachweisbaren kanonischen Werke älteren Epochen der deutschen Literatur zuzuordnen sind und sich unter diesen Werken kein Werk der Gegenwartsliteratur finden lässt.

Die häufigsten kanonischen Texte der Gegenwartsliteratur, in diesem Fall Daniel Kehlmanns „Die Vermessung der Welt“ sowie Bernhard Schlinks „Der Vorleser“, lassen sich jeweils dreimal nachweisen. Interessant ist bezüglich der kanonischen Werke ebenfalls, dass im Falle von Goethe zu einem Großteil auf nicht-kanonische Werke des Autors zurückgegriffen wird. Seine mehrfach nachweisbaren kanonischen Werke (Faust I: 3 Textgrundlagen, Die Leiden des jungen Werthers/Iphigenie auf Tauris: je 2 Textgrundlagen) machen insgesamt nicht einmal die Hälfte aller Textgrundlagen dieses Autors aus.

4.2 Aufgabenbezogene Innovationen in zentralen Abschlussprüfungen

Die Analyse der Prüfungsaufgaben macht insgesamt deutlich, dass innovative Schreibaufgaben und Materialgrundlagen gegenüber traditionellen Formen nur eine marginale Bedeutung in zentralen Prüfungen haben. Traditionelle Aufsatzformen sowie Lesetexte als Materialgrundlagen zur Beantwortung der Prüfungsaufgaben sind grundsätzlich gegenüber den anderen genannten Innovationen überrepräsentiert.

Tab. 3: Aufsatzformen in zentralen Abschlussprüfungen (N=941)

	absolute Häufigkeit	relative Häufigkeit
traditioneller Aufsatz	756	80,3%
nicht-traditionelle Aufsatzform	185	19,7%

Interessante Differenzen zeigen sich diesbezüglich zwischen den Prüfungen für den Mittleren Schulabschluss und den Abiturprüfungen, da in den Prüfungen für den Mittleren Schulabschluss wesentlich häufiger Innovationen zur Anwendung kommen als in den zentralen Abiturprüfungen. So verlangen in den Abiturprüfungen unter 10% aller analysierten Aufgaben *nicht* die Produktion eines traditionellen Aufsatzes, während in den MSA-Prüfungen etwa ein Viertel der Aufgaben (25,3%) als innovative Aufsatzform kategorisiert wurde. Auch die Anwendung und Kombination von geschlossenen, halboffenen und offenen Aufgabenformaten, die als besonders geeignet für Leistungsüberprüfungen angesehen wird, ist ein Aspekt der sich *ausschließlich* in den MSA-Prüfungen nachweisen lässt.

Innerhalb der als nicht-traditionelle Aufsatzform kategorisierten Aufgaben ließ sich der Einbezug verschiedener innovativer Merkmale nachweisen. Aufgaben, die kombinierte Texterschließungsformen erfordern (34,1%), einen Bezug zur Lebenswelt der Schülerinnen und Schüler aufweisen (26,5%) oder kreatives Schreiben verlangen (25,9%), stellen hierbei den Großteil der Innovationen dar, während etwa der Einbezug eines Portfolios (2,2%), sowie medial vermittelte Texte (1,1%) und nichtkontinuierliches Material (11,9%) eher marginale Bedeutung haben.¹

¹ Angaben über 100% hängen damit zusammen, dass einzelne Innovationen gleichzeitig zur Anwendung kommen können (etwa Portfolio und Lebensweltbezug)

Die differenzierte Betrachtung von traditionellen und innovativen Aufgabenmerkmalen in den einzelnen Bundesländern weist darauf hin, dass die unterschiedlich lange zentrale Prüfungstradition der Länder² nicht zwangsläufig Rückschlüsse auf den Einbezug von Innovationen in den einzelnen Ländern zulässt. So lässt sich zwar in mehreren traditionell zentral prüfenden Ländern (etwa in Bayern und Baden-Württemberg) die als innovativ eingestufte Kombination verschiedener Aufgabenformate nicht nachweisen, allerdings greifen auch diese Länder eine Vielzahl der anderen untersuchten Innovationen auf. Darüber hinaus ist Bayern eines der wenigen Länder, welches medial vermittelte Texte für die Abiturprüfung mit einbezieht und in Baden-Württemberg lässt sich der Einbezug eines Portfolios für die zentrale Prüfung zum Erwerb des Mittleren Schulabschlusses nachweisen. Trotz der insgesamt großen Bandbreite von einbezogenen Innovationen zeigt sich, dass nur wenige dieser Impulse, wie etwa das kreative Schreiben und kombinierte Texterschließungsformen, bundesweit etabliert sind. Für viele der weiteren nachgewiesenen Neuerungen zeigt sich das zuvor beschriebene Bild: ihr Einsatz bleibt auf wenige Länder beschränkt. Beim Vergleich des Verhältnisses traditioneller und innovativer Aufgaben in den einzelnen Prüfungsjahren zeigt sich insgesamt eine ausgeglichene Tendenz, da sich nicht nachweisen lässt, dass innovative Aufgaben von Jahr zu Jahr mehr bzw. weniger eingesetzt werden.

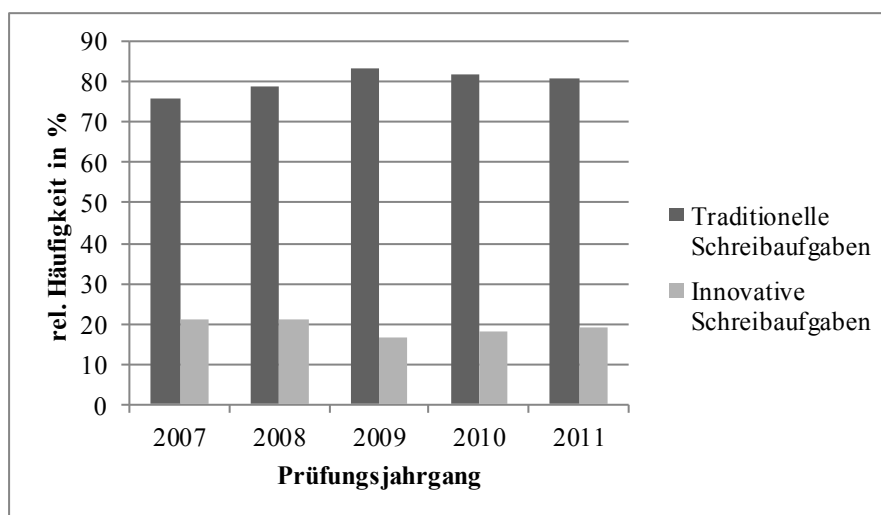


Abb. 3: Innovative und traditionelle Schreibaufgaben in zentralen Abschlussprüfungen (N=941)

² Klein et al. (2009) stellen heraus, dass sich in Deutschland insgesamt drei Implementationszeiträume nachweisen lassen: Traditionell zentral prüfen seit der Nachkriegszeit Bayern, Baden-Württemberg sowie das Saarland. Der Großteil der ostdeutschen Länder (Sachsen, Sachsen-Anhalt, Thüringen und Mecklenburg-Vorpommern) knüpfte an die zentrale Prüfungstradition der DDR nach der Wiedervereinigung 1990 an. Die übrigen Länder hielten doch bis Mitte der 2000er Jahre an der dezentralen Prüfungsform fest und stellten im Zuge der Reformen ab der Jahrtausendwende auf eine zentrale Prüfungsorganisation um.

5. Diskussion der Ergebnisse

In Bezug auf den Titel dieses Beitrages „Zwischen Tradition und Innovation“ lässt sich zusammenfassend festhalten, dass traditionelle Aufgaben und Inhalte nach wie vor ein starkes Gewicht innerhalb der Abschlussprüfungen haben, fachdidaktische Innovationen und Anstöße aber durchaus berücksichtigt werden. Die Prüfungen zum Mittleren Schulabschluss weisen auf allen Untersuchungsebenen ein größeres Maß an Innovationen auf als die Abiturprüfungen, die insgesamt ein wesentlich begrenzteres Spektrum an Inhalten und Aufgabenformaten zeigen. Diese Befunde können als vorsichtige Bestätigung der eingangs formulierte These, dass die Aufgabenentwickler bei diesen Prüfungen ein größeres Wagnis hinsichtlich der Implementation inhaltlicher und formaler Neuerungen bei der Aufgabengestaltung an den Tag legen, gewertet werden – gleichwohl müsste diese Vermutung in weiterführenden Studien vertieft werden (z. B. im Rahmen von Interviews mit den an der Aufgabenentwicklung beteiligten Lehrkräften). Innovative Aufgaben machen insgesamt nur einen geringen Anteil der Prüfungsaufgaben aus, wobei auf wenige innovative Aufgabenformate (kreatives Schreiben, kombinierte Texterschließungsformen, Einbezug der Lebenswelt) ein Großteil der Aufgaben entfällt, während weitere innovative Aufgabenkonzepte (Portfolios, medial vermittelte Texte, nichtkontinuierliches Material) weiterhin eine Randerscheinung darstellen. Vor allem der geringe Einbezug innovativer Materialgrundlagen, wie des Portfolios oder nicht textförmiger Materialgrundlagen, zeigt deutlich auf, dass traditionelle Materialgrundlagen in Form eines Lesetextes die Regel in zentralen Prüfungen darstellen. Hinsichtlich der innovativen Aufgaben zeigte sich deutlich, dass zwar alle Länder auf innovative Aufgaben zurückgreifen, jedoch nur wenige aufgabenbezogene Innovationen bundesweit (etwa das kreative Schreiben sowie kombinierte Texterschließungsformen) etabliert sind. Auch im Vergleich der einzelnen Prüfungsjahrgänge zeigt sich keine signifikante Veränderung hinsichtlich des Einbezuges innovativer Aufgaben: ihr Anteil liegt grundsätzlich bei etwa 20%. Dieser Befund bestätigt vorherige Befunde zur hohen zeitlichen Konstanz der Aufgabengestaltung in zentralen Prüfungen (Kühn 2010; Kühn/Drüke-Noe 2013). Durch die kategoriengeleitete Analyse der Prüfungsaufgaben konnte letztendlich aufgezeigt werden, dass die Forderungen der bundesweit gültigen Vorgaben sowie der Fachdidaktik Deutsch durchaus Eingang in die zentralen Abschlussprüfungen finden. Elemente der geforderten neuen Aufgabenkultur im Deutschunterricht (Köster 2003) lassen sich auch in den zentralen Abschlussprüfungen nachweisen. Insbesondere mit Blick auf die normierende Wirkung der Prüfungen auf die Unterrichtsgestaltung wäre allerdings ein verstärkter Einbezug dieser Aufgaben durchaus wünschenswert, damit innovative Aufgaben und Inhalte auch im Unterrichtsgeschehen nicht nur eine untergeordnete Rolle gegenüber traditionellen Aufgaben spielen.

Für die Textgrundlagen zeigte sich allerdings wesentlich deutlicher, dass sich von Seiten der Fachdidaktik angestoßene sowie von den bundesweit gültigen Vorgaben (EPA, BS MSA) vorgegebene Veränderungen *kaum* in den Prüfungsaufgaben nachweisen lassen. So finden sich beispielsweise nur wenige nicht-literarische Texte in den Prüfungen (27,6%) und mit beinahe 60% machen literarische Textgrundlagen weiterhin den Großteil der Bearbeitungsgrundlagen aus. In Bezug auf die literarischen Texte zeigt sich des Weiteren, dass in zentralen Prüfungen lediglich auf einen bestimmten Kernbestand an Literatur zurückgegriffen wird. Es wird deutlich, dass klassische Autoren und Werke der deutschen Literatur gegenüber der Gegenwartsliteratur eine wesentlich stärkere Bedeutung innerhalb der Prüfungen haben. Der Befund einer „Rekanonisierung“ (Abraham 2010) scheint sich also zu bestätigen. Auch der von Seiten der Fachdidaktik beklagte zu geringe Einbezug von dramatischen Texten bestätigte sich durch die kategoriengeleitete Analyse, ebenso die Marginalisierung bestimmter literarischer Epochen, in diesem Falle vor allem der Literatur der Zeit von der Antike bis zum

Barock, die weniger als 3% aller Textgrundlagen ausmacht. Dies ist vor dem Hintergrund zahlreicher Publikationen, die etwa die Chancen dieser literarischen Epochen für kompetenzorientiertes Lernen betonen (Miedema 2011) bedenkenswert. Des Weiteren gilt es nun kritisch zu diskutieren, inwieweit diese Einengung des literarischen Spektrums auch negative Auswirkungen auf den vorgelagerten Unterricht hinsichtlich eines möglichen teaching to the test Effekts hat, da jüngste Studien bestätigen, dass durch eine Fokussierung auf Abitur-schwerpunkte sich längerfristig eine eher problematische Unterrichtsgestaltung, die auf das Einüben der Prüfungsinhalte zielt, etablieren kann (Maag Merki 2012). Demgegenüber stehen allerdings normative Annahmen hinsichtlich eines Kanons, die entgegen der Entwicklungen seit 1968 einen literarischen Kanon nicht mehr als aufoktroiertes Herrschaftsinstrument sehen, sondern durchaus dessen Vorteile etwa für die Schulorganisation begrüßen (Vogelsaenger 2012). Wie also letztendlich eine literarische Kanonisierung bestimmter Autoren und Werke der deutschen Literatur im Kontext zentraler Abschlussprüfungen zu bewerten sein wird, kann zum jetzigen Zeitpunkt noch nicht abschließend beurteilt werden.

Es zeigt sich letztendlich also, dass in zentralen Abschlussprüfungen Innovationen zum Teil aufgenommen werden, insgesamt aber noch Optimierungsbedarf besteht. Damit zentrale Abschlussprüfungen tatsächlich ihren intendierten Steuerungswirkungen gerecht werden können, wäre also ein vermehrter Einbezug innovativer Aufgaben und Inhalte durchaus begrüßenswert, da diese somit auch stärker in den vorgelagerten Unterricht Eingang finden würden. Insgesamt muss bei der Interpretation der Ergebnisse auch immer berücksichtigt werden, dass es sich hierbei lediglich um eine Analyse weniger Prüfungsjahrgänge handelt und generellen Aussagen hierdurch Grenzen gesetzt sind. In weiteren Forschungsarbeiten wäre deshalb dieser begrenzte Fokus auszuweiten und auch danach zu fragen, wie zentrale Prüfungen tatsächlich den Unterricht sowie die Gestaltung von Klassenarbeiten bzw. Klausuren beeinflussen. Es wäre in diesem Kontext auch von besonderem Interesse, ob sich hier vor dem Hintergrund des unterschiedlichen Innovationsgrades der Prüfungen für das Abitur bzw. den Mittleren Schulabschluss Unterschiede in der Unterrichtsgestaltung zwischen Sekundarstufe I und II zeigen. Auch vor dem Hintergrund der jüngst veröffentlichten Bildungsstandards für die Abiturprüfungen, die ab dem Schuljahr 2016/2017 die EPA als Grundlage für die Gestaltung der Prüfungen ablösen sollen, stellt sich die Frage, ob und inwieweit hier Veränderungen auftreten werden.

Literatur

- Abraham, U. (2008): Zentral + Föederal = Katastrophal? Ein Länderbericht zum Zentralabitur. In: Didaktik Deutsch, 25 (2008), 52-77.
- Abraham, U. (2010): Was sollen unsere Abiturienten in Zukunft können und wie lässt sich das überprüfen? In: Der Deutschunterricht, 62, H. 1 (2010), 52-61.
- Abraham, U./Kepser M. (2009): Literaturdidaktik Deutsch. Eine Einführung. 3. Auflage. Berlin.
- Becker-Mrotzeck, M./Böttcher, I. (2006): Schreibkompetenz entwickeln und beurteilen. Praxishandbuch für die Sekundarstufe I und II. Berlin.
- Becker-Mrotzeck, M./Kepser, M. (2010): Sprach-, kultur- und medienwissenschaftliche Themen im Zentralabitur. In: Der Deutschunterricht, 62, H. 1 (2010), 14-18.
- Bekes, P. (2010): Berichte und Einschätzungen zum Zentralabitur aus fünf Bundesländern. Nordrhein-Westfalen. In: Der Deutschunterricht, 62, H.1 (2010), 71-86.

- Block, R./Klein, E. D./Ackeren, I. van/Kühn, S. M. (2011): Leistungseffekte des Zentralabiturs? Eine kritische Auseinandersetzung mit bildungsökonomischen Interpretationen zu den Effekten der Prüfungsorganisation auf Basis von PISA E 2003-Daten. In: *bildungsforschung*, 8, H. 1 (2011), 215-238.
- Böhme, K./Bremerich-Vos, A. (2008): Expertise MSA/P10 Deutsch. Bezug der Prüfungsaufgaben in Berlin und Brandenburg 2006/2007 zu den kompetenzorientierten Anforderungen der länderübergreifenden Bildungsstandards für das Fach Deutsch. Verfügbar unter: http://www.isqbb.de/uploads/media/ISQ_Expertise__Deutsch_MSA_P10.pdf (Zugriff am 19.11.13).
- Brosius, H. B./Koschel, F./Haas, A. (2009): *Methoden der empirischen Kommunikationsforschung*. Wiesbaden.
- Bühler, K. (1999): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. 3. Auflage. Stuttgart.
- Büchter, A. (2007): Zentrale Prüfungen am Ende der Klasse 10 – Wirkungen auf Unterricht und Leistungsbewertung. In: *Beiträge zum Mathematikunterricht*, 503-506.
- Cheng, L./Curtis, A. (2012). Test Impact and Washback: Implications for Teaching and Learning. In: Coombe, C./O'Sullivan, B./Davidson, P./Stoyhoff, S. (Eds.): *Cambridge Guide to Second Language Assessment*. Cambridge, 89-95.
- Eikenbusch, G. (2001): *Qualität im Deutschunterricht der Sekundarstufe I und II*. Berlin.
- Erlinger, H. D. (2002): Kanonfragen für die Medienerziehung im Deutschunterricht. In: *Text+Kritik: Literarische Kanonbildung*. München, 23-40.
- Fix, M. (2006): *Texte schreiben. Schreibprozesse im Deutschunterricht*. Paderborn/München/Wien/Zürich.
- Franke, C. (2005): Viel Lärm um Nichts? – Vom Text zum Diagramm und zurück. In: *Deutschunterricht*, 58, H. 5 (2005), 10-15.
- Freudenberg, R. (2010): Berichte und Einschätzungen zum Zentralabitur aus fünf Bundesländern. Thüringen. In: *Der Deutschunterricht*, 62, H.1 (2010), 71-86.
- Groeben, N. (2006): Auf dem Weg zu einer deutsch-didaktischen Unterrichtsforschung? In: Stückrath, J./Strobel, R. (Hrsg.): *Deutschunterricht empirisch. Beiträge zur Überprüfbarkeit von Lernfortschritten im Sprach- Literatur und Medienunterricht*. Baltmannsweiler, 7-33.
- Jordan, A./Ross, N./Krauss, S./Baumert, J./Blum, W./Neubrandt, M./Löwen, K./Brunner, M./Kunter, M. (2006): *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt*. Berlin.
- Kahnert, J./Lorenz, R./Eickelmann, B. (2012): Zentralabitur NRW – Das Verfahren zur Qualitätssicherung. In: *Friedrich Jahresheft „Schule vermessen“*, 54-56.
- Kammler, C./Noack, B. (2010): Literaturgeschichte und Kanon im Zentralabitur 2008-2010. In: *Der Deutschunterricht*, 62, H. 1 (2010), 5-13.
- Klein, E. D./van Ackeren, I. (2012). Challenges and Problems for Research in the Field of Statewide Exams. A Stock Taking of Differing Procedures and Standardization Levels. In: *Studies in Educational Evaluation*, 37, H. 4 (2012), 180-188.
- Klein, E. D./Kühn, S. M./van Ackeren, I./Block, R. (2009): Wie zentral sind zentrale Prüfungen? Abschlussprüfungen am Ende der Sekundarstufe II im nationalen und internationalen Vergleich. In: *Zeitschrift für Pädagogik*, 55, H. 4 (2009), 596-621.
- Köster, J. (2003): Konstruieren statt entdecken - Impulse aus der PISA-Studie für die deutsche Aufgabenkultur. In: *Didaktik Deutsch*, 9 (2003), 4-17.
- Kühn, S.M./Drücke-Noe, C. (2013). Qualität und Vergleichbarkeit durch Bildungsstandards und zentrale Prüfungen? – Ein bundesweiter Vergleich von Prüfungsanforderungen im

- Fach Mathematik zum Erwerb des Mittleren Schulabschlusses. In: Zeitschrift für Pädagogik 59, H. 6 (2013), 912-932.
- Kühn, S. M. (2013). Vergleichbarkeit beim Mittleren Schulabschluss? Ein Überblick über die Vielfalt schulstrukturell möglicher Bildungswege und Prüfungsverfahren in den deutschen Ländern. In: Die Deutsche Schule, 105, H. 1 (2013), 87-101.
- Kühn, S. M. (2011): Kategoriensystem zur Analyse zur Analyse von Prüfungsaufgaben zum Erwerb des Mittleren Schulabschlusses im Fach Mathematik. Internes Arbeitspapier, Arbeitsgruppe Bildungsforschung. Universität Duisburg-Essen.
- Kühn, S. M. (2010): Steuerung und Innovation durch Abschlussprüfungen? Wiesbaden.
- Langer, I./Thun, F. S. von (2007): Messung komplexer Merkmale in Psychologie und Pädagogik. Ratingverfahren. Münster.
- Leonord, C. (2009): Abitur nach Partitur? Auswirkungen der Zentralabitureinführung auf die Individualkonzepte von Musiklehrern. Bremen.
- Ludwig, O. (1996): Der Unterricht findet nicht statt: Zur Schreibpraxis der reformierten Oberstufe. In: Peyer, A./Portmann, P. (Hrsg.): Norm, Moral und Didaktik – Die Linguistik und ihre Schmuttelkinder. Tübingen, 221-240.
- Maag Merki, K. (2012): Zentralabitur. Die längsschnittliche Analyse der Wirkung der Einführung zentraler Abiturprüfungen in Deutschland. Wiesbaden.
- Maiwald, K. (2004): Neue Kleider für den Kaiser, oder alte Hüte? Zur schwierigen Verwebung von Textverstehen und Testaufgaben. In: Köster, J./Lütgert, W./Creutzburg, J. (Hrsg.): Aufgabenkultur und Lesekompetenz. Deutschdidaktische Positionen. Frankfurt a. Main, 43-50.
- Mayring, P. (2000): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 7. Auflage. Weinheim/Basel.
- Miedema, N. (2011): Mittelalterliche Texte. In: Praxis Deutsch, 230 (2011), 4-11.
- Neubrand, M. (2002): Einige Hinweise zu mathematisch-didaktisch relevanten Ansätzen von PISA-2000. In: Mitteilungen der Gesellschaft für Didaktik der Mathematik, 74 (2002), 60-64.
- Nicols, S./Berliner, D. C. (2007): Collateral damage. How high-stakes testing corrupts America's schools. Cambridge Mass.
- Prodromou, L. (1995): The backwash effect. From testing to teaching. In: ELT Journal 49, H. 1 (1995), 13-25.
- Roppelt, A./Blum, W./Keller, K./Leiß, D./Drüke-Noe, C. (2008): Expertise MSA/P10 Mathematik. Bezug der Prüfungsaufgaben in Berlin und Brandenburg 2006/2007 zu den kompetenzorientierten Anforderungen der länderübergreifenden Bildungsstandards für das Fach Mathematik. Verfügbar unter:
http://www.isqbb.de/uploads/media/ISQ_Expertise_Mathematik_MSA_P10.pdf (Zugriff am 23.11.13).
- Schmellentin, C. (2012): Kompetenzorientierung im Deutschunterricht: Auswirkungen auf Aufgaben im Bereich Grammatik. In: Keller, S./Bender, U. (Hrsg.): Aufgabekulturen. Fachliche Lernprozesse herausfordern, begleiten, reflektieren. Seelze, 113-124.
- Schmidt, R. (2009): Einleitung. Der literarische Kanon: ein Organ des Willens zur Macht oder des Gewinns an Kompetenzen? In: Saul, N./Schmidt, R. (Hrsg.): Literarische Wertung und Kanonbildung. Würzburg, 9-21.
- Seidel, T. (2003): Videobasierte Kodierverfahren in der IPN Videostudie Physik. Ein methodischer Überblick. In: Seidel, T./Prenzel, M./Duit, R./Lehrke, M. (Hrsg.): Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“. Kiel, 99-111.

Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2002): Einheitliche Prüfungsanforderungen in der Abiturprüfung Deutsch. Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 24.05.2002. Verfügbar unter:

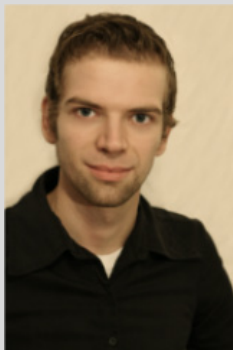
http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/Beschluesse_Veroeffentlichungen/allg_Schulwesen/epa_deutsch.pdf (Zugriff am 23.11.13).

Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2003): Bildungsstandards für den Mittleren Schulabschluss im Fach Deutsch. Beschluss vom 04.12.2003. Verfügbar unter:

http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-Deutsch-MS.pdf (Zugriff am 23.11.13).

Vogelsaenger, W. (2012): Über den Nutzen eines Bildungskanons für die Schulorganisation. In: Bildungskanon heute. Berlin, 48-56.

Wirtz, M./Caspar, M. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen.



Stephan Otto

ist wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Bildungsforschung, Arbeitseinheit Bildungssystem- und Schulentwicklungsforschung, am Institut für Pädagogik der Universität Duisburg-Essen.

Kontakt: stephan.otto@uni-due.de



Dr. Svenja Mareike Kühn

ist Akademische Rätin in der Arbeitsgruppe Bildungsforschung, Arbeitseinheit Bildungssystem- und Schulentwicklungsforschung am Institut für Pädagogik der Universität Duisburg-Essen. Sie ist Vorstandsmitglied im Zentrum für Empirische Bildungsforschung.

Kontakt: svenja.kuehn@uni-due.de
www.uni-due.de/bifo