

Svenja Mareike Kühn

## **Und was kommt an?**

**Analysen zur Implementation von Innovationen  
in der gymnasialen Oberstufe und im Abitur**

---

### **Zusammenfassung**

*In jüngster Zeit sind zahlreiche Neuerungen für die gymnasiale Oberstufe und die Abiturprüfung implementiert worden. Ob das Innovationspotenzial dieser Reformen jedoch tatsächlich mit den neuen Instrumenten genutzt und in der Praxis aufgegriffen wird, ist bislang kaum empirisch überprüft worden. Der vorliegende Beitrag betrachtet vor diesem Hintergrund im Ländervergleich kategoriengeleitete Klausur- und Abiturprüfungsaufgaben im retrospektiven Längsschnittdesign, um zu erfassen, inwieweit Innovationen hinsichtlich der Abiturprüfungen und gymnasialen Oberstufen tatsächlich implementiert worden sind.*

*Schlüsselwörter: Abitur, gymnasiale Oberstufe, Aufgabe, Innovation, Implementation*

### **Reality Check**

**An Analysis of How Innovations Are Implemented in Upper Secondary Education and the Abitur-Exams**

### **Abstract**

*In recent years, numerous formal and content-related alterations with regard to upper secondary education and the Abitur-examination have been implemented. However, findings concerning their innovative impact and their practical implementation are still pending. The presented study is based on this research gap and analyses Abitur-examination tasks and class exercises from various German Länder in a retrospective longitudinal design. In summary, the results contribute to empirical evidence regarding the factual implementation of innovations in upper secondary education and in the Abitur-examination.*

*Keywords: exit examinations, upper secondary education, tasks, innovation, implementation*

Seit der Einführung der Abiturprüfung im Übergang vom 18. zum 19. Jahrhundert wurden immer wieder intensive Diskurse über Verfahren, Inhalte und letztlich

auch den Stellenwert dieser Abschlussprüfung geführt. So steht sie auch gegenwärtig in der Diskussion um die Qualität schulischer Bildung und Maßnahmen zu ihrer Verbesserung im Fokus des öffentlichen, bildungspolitischen und mittlerweile auch wissenschaftlichen Interesses.

Betrachtet man die zahlreichen seitens der Bildungspolitik initiierten Reformen in Folge der großen Schulleistungsstudien der letzten Jahre, so sind diese auch an der gymnasialen Oberstufe und der Abiturprüfung nicht folgenlos vorbeigegangen. Bundesweit lässt sich beispielsweise ein Trend zur Verkürzung der gymnasialen Schulzeit von neun auf acht Jahre, zu zentralen Abiturprüfungen sowie zur Aufwertung sog. ‚Kernfächer‘ als verbindliche Prüfungsfächer erkennen, wenngleich länderspezifische Akzentuierungen hinsichtlich der genannten Reformbausteine offenkundig sind (vgl. Kühn 2010; Trautwein/Neumann 2008). Neben diesen strukturellen Veränderungen stehen auch fachinhaltliche und fachdidaktische Innovationen im Fokus des Interesses. Diese betreffen u.a. neuartige Unterrichtsinhalte, die Stärkung innovativer Unterrichtsmethoden sowie alternative Formen der Leistungsbewertung und innovative Aufgabenformate in der Abiturprüfung.

Die bundesweit gültigen *Einheitlichen Prüfungsanforderungen in der Abiturprüfung* (EPA) der KMK greifen zahlreiche dieser Aspekte auf und fokussieren in ihren aktuellen Fassungen auch auf Neuerungen hinsichtlich der Fachinhalte und der Strukturen der gymnasialen Oberstufe sowie der Abiturprüfung. Dabei beinhalten die Vorgaben im Sinne einer *Weiterentwicklung* zwar merkliche Veränderungen, knüpfen dabei jedoch an bestehende Praktiken an, z.B. durch die Ergänzung traditioneller Aufgabenarten um innovative Aufgabenformate. Sie schaffen zudem bewusst Handlungsspielräume, die länderspezifische Akzentuierungen ermöglichen, um dem Strukturprinzip des Föderalismus gerecht zu werden. Die EPA gelten damit als zentrales Steuerungsinstrument der KMK, um bundesweit gültige Standards zu formulieren, die zentrale Aspekte einer neuen Lehr-/Lern- und Aufgabenkultur aufgreifen; sie sollen die traditionelle Praxis ergänzen und weiterentwickeln, um so zur Qualitätsentwicklung in der gymnasialen Oberstufe und im Abitur beizutragen. Somit scheint ihnen ein deutliches Innovationspotenzial zugesprochen zu werden. Ob und inwieweit entsprechende Neuerungen jedoch tatsächlich in der Praxis aufgegriffen werden, ist für die Abiturprüfungsverfahren sowie die gymnasialen Oberstufen im Ländervergleich bislang nicht empirisch überprüft worden (vgl. Trautwein u.a. 2007, S. 24).

Vor dem Hintergrund neuer Steuerungstheorien (vgl. Altrichter/Maag Merki 2010) ist davon auszugehen, dass die Umsetzung von überregionalen Reformen nicht linear auf Grund formaler Weisungen seitens der KMK erfolgt, sondern dass sie durch die am Prüfungsverfahren beteiligten Akteure und ihre je eigenen Handlungsrouninen und organisationalen Rahmungen mit beeinflusst wird. Im Sinne des Governance-Konzeptes (vgl. Altrichter/Brüsemeister/Wissinger 2007) ist Schule als komplexes

Konstellationsgefüge zu verstehen, das nicht nur mittels formaler Weisungen hierarchisch höherer Ebenen, sondern vor allem auch durch von einzelnen Akteuren an lokalen Stellen getroffene, informelle Regelungen entsteht. Ergebnisse der Innovations- und Implementationsforschung bestätigen in diesem Zusammenhang eine nur geringe Übereinstimmung zwischen der konzeptionellen Planung bildungspolitischer Reformen und deren tatsächlicher Realisierung auf der Anwendungsebene (vgl. Gräsel/Parchmann 2004) bzw. machen deutlich, dass Innovationen im Bildungsbereich nur schwer Verbreitung finden und es häufig zu einem „Versanden“ kommt (vgl. Gräsel 2010, S. 7). Bezogen auf den Implementationserfolg der EPA, d.h. deren tatsächliche Umsetzung in den Abiturprüfungsverfahren sowie der gymnasialen Oberstufe im Ländervergleich, liegen jedoch bislang keine empirischen Befunde vor (s.o.). Die zunehmende Standardisierung des Abiturprüfungsverfahrens durch die Einführung zentraler Prüfungen in der Mehrheit der Länder ist Ausdruck der Annahme, dass die Einhaltung von Standards, aber auch die Umsetzung innovativer Ansätze über zentrale Prüfungsverfahren eher gesichert werden können als über dezentrale Prüfungsformen (vgl. van Ackeren 2007). Ein empirischer Nachweis dafür findet sich jedoch bis dato ebenfalls nicht.

Der vorliegende Beitrag betrachtet kategoriengeleitet und im retrospektiven Längsschnittdesign (1993-2008) Klausur- und Abiturprüfungsaufgaben auf Basis der EPA aus Ländern mit unterschiedlichen Prüfungstraditionen, und zwar exemplarisch in den Fächern Biologie, Chemie und Physik, um zu erfassen, ob und inwieweit fachinhaltliche und fachdidaktische Innovationen hinsichtlich der Abiturprüfungen und gymnasialen Oberstufen im Kontext der damit verbundenen Reformen implementiert worden sind. Vor der Darstellung zentraler Befunde dieser Studie wird einleitend zunächst knapp skizziert, was unter der *Implementation von Innovationen* im Bildungswesen verstanden wird, was eine erfolgreiche Implementation von Innovationen charakterisiert und welche Faktoren diesen Prozess beeinflussen können (Abschnitt 1). Es folgt eine Erläuterung des *methodischen Vorgehens* der o.g. Studie (Abschnitt 2), an die sich die Darstellung *zentraler Ergebnisse* anschließt (Abschnitt 3). Der vorliegende Beitrag schließt mit einer *Diskussion* zu potenziellen Einflussfaktoren der dargestellten Entwicklungen (Abschnitt 4).

## 1. Implementation von Innovationen im Bildungsbereich

Allgemein formuliert meint *Innovation* die Einführung einer neuen Idee oder Erfindung, von der sich eine Gesellschaft oder eine maßgebliche Gruppe der Gesellschaft wesentliche Verbesserungen in einem bestimmten Bereich (z.B. Technik, Sport) verspricht (vgl. Oerter 2010, S. 13; vertiefend Rürup in diesem Heft). Bezogen auf den Bildungsbereich kann Innovation demnach die Einführung system-, einzel-schul- oder unterrichtsbezogener Neuerungen meinen, die seitens einer bestimmten Gruppe von Akteuren im Bildungssystem (z.B. Bildungspolitik und -administra-

tion) als notwendig oder sinnvoll erachtet werden, um zu einer Weiterentwicklung in einem bestimmten Bereich des Bildungssystems beizutragen. Das Spektrum an Innovationen im Bildungsbereich geht nach Prenzel (vgl. 2010, S. 24) von *Tools* bzw. produktgestützten Hilfsmitteln (z.B. Diagnoseinstrumenten) über inhaltlich fokussierte *Maßnahmen* mit bestimmten Zielgruppen (z.B. Leseförderung von Jungen) bis hin zu komplexen Maßnahmenbündeln, die er als *Programme* bezeichnet. Letzteres umfasst beispielsweise länderspezifische Modellversuchsprogramme und bundesweite Programme, wie z.B. ‚Chemie im Kontext‘ (ChiK; vgl. Demuth u.a. 2008), aber auch umfassende Maßnahmenbündel, wie sie beispielsweise mit der Einführung der EPA in Deutschland verbunden sind.

In diesem Zusammenhang gilt es im Bildungsbereich zu berücksichtigen, dass allein die Einführung einer neuartigen Idee nicht ausreichend ist; diese muss auch umgesetzt bzw. implementiert werden. Berner, Oelkers und Reusser merken an, dass der Begriff der Implementation weit über den eindimensionalen Vorgang des ‚Einführens‘ hinausgeht und insbesondere den Prozess der ‚Umsetzung‘ einer Innovation meint (vgl. 2008, S. 224) – es handelt sich also um einen langfristig angelegten, mehrdimensionalen Prozess. Von einer erfolgreichen Implementation im Bildungsbereich spricht man folglich dann, wenn eine Innovation in der angezielten Praxis aufgenommen und standardmäßig umgesetzt wird.

Nach Gräsel, Jäger und Willke (vgl. 2006, S. 500ff.) kann die erfolgreiche Implementation einer Innovation an vier maßgeblichen Indikatoren festgemacht werden. Diese umfassen: 1. den Grad der *Verbreitung* der Innovation (z.B. innerhalb eines Bundeslandes), 2. die *Tiefe* der Umsetzung (oberflächliche Übernahme von Neuerungen vs. einschneidende Veränderungen der bisherigen Praxis), 3. das Ausmaß der gefühlten *Zugehörigkeit* seitens der betroffenen Akteure (insbesondere der Lehrkräfte) sowie 4. – vor dem Hintergrund angestrebter dauerhafter Veränderung in bestimmten Bereichen des Bildungssystems – die *Nachhaltigkeit* einer Innovation.

Die Implementation einer Innovation im Bildungsbereich ist ein hochgradig komplexer Prozess, der von zahlreichen Faktoren beeinflusst werden kann (vgl. vertiefend Fullan 1994; Thomas 1994). Diese Einflussfaktoren lassen sich nach Fullan (vgl. 1994, S. 2840ff.) in drei Kategorien einordnen: Die erste Kategorie umfasst die *Charakteristika der Innovation* (beispielsweise die Komplexität, Qualität und Praktikabilität der Innovation sowie die Klarheit über entsprechende Ziele und Wege), da deren Eigenschaften die tatsächliche Umsetzung in der Praxis beeinflussen. Zudem gilt die aus Sicht der betroffenen Akteure postulierte *Notwendigkeit* einer Innovation bzw. deren Bedürfnis nach einer Veränderung in einem bestimmten Bereich als entscheidender Einflussfaktor.

Die zweite Kategorie umfasst nach Fullan (vgl. ebd., S. 2842ff.) Aspekte des *lokalen Kontextes*, wobei diese lokale Prozesse und Charakteristika der Implementation betreffen. Dazu gehören sowohl die *regionale Verwaltung* (z.B. Bezirksregierung, lokale Schulverwaltung) als auch der einzelschulische Kontext (hier: Schulleitung, Kollegium), denen bei der Begleitung des Umsetzungsprozesses eine entscheidende Bedeutung zukommt. Letztlich stellen insbesondere die Kompetenzen und Einstellungen der einzelnen *Lehrpersonen* einen wichtigen Einflussfaktor dar.

Die dritte Kategorie umfasst nach Fullan (vgl. ebd., S. 2844f.) *externe Faktoren* im Implementationsprozess; dazu gehören insbesondere Akteure der *Bildungspolitik* und *-administration*, *externe Agenturen* (z.B. Landesinstitute für Schulqualität) sowie entsprechend involvierte *Interessengruppen* (z.B. Gewerkschaften, Eltern- und Schülervertretung), da diese über entsprechende Instrumente und Ressourcen zur Durchsetzung der Innovation verfügen und damit zur Veränderung relevanter Strukturen, Prozesse und Kulturen beitragen können.

Die nachfolgend dokumentierte Studie greift den Diskurs über das Innovationspotenzial der Veränderungstendenzen und Reformbemühungen, die sich in den letzten Jahren hinsichtlich der Abiturprüfungsverfahren und gymnasialen Oberstufen in Deutschland vollzogen haben, auf und diskutiert den Implementationserfolg administrativer Vorgaben seitens der KMK am Beispiel der *Einheitlichen Prüfungsanforderungen in der Abiturprüfung* (EPA).

## 2. Methodisches Vorgehen

Das Projekt zielt auf die Erfassung, Beschreibung und Analyse eindeutig definierter Merkmale der Aufgabenstellung *in Abitur- und Klausuraufgaben*, die sich aus den Vorgaben der EPA zur Aufgabenkonstruktion ergeben haben. Anhand der unterschiedlichen Aufgaben kann also gemessen werden, ob und inwieweit Innovationen in den Abiturprüfungen und gymnasialen Oberstufen der Länder mit unterschiedlichen Prüfungstraditionen tatsächlich implementiert worden sind. Somit können mittels Aufgabenanalyse der Implementationserfolg der EPA sowie darüber hinaus auch das Innovationspotenzial zentraler Prüfungen erfasst werden.

Dazu wurden exemplarisch für die drei naturwissenschaftlichen Prüfungsfächer Biologie, Chemie und Physik *zwei empirische Studien* durchgeführt: zum einen die Analyse von schriftlichen *Abiturprüfungsaufgaben* und zum anderen die Analyse von *Klausuraufgaben aus der Qualifikationsphase*.

Für die *erste Studie* wurden vier Länder mit unterschiedlicher Prüfungstradition im *Abitur* ausgewählt, und zwar Baden-Württemberg und das Saarland als traditi-

onell zentral prüfende Länder, Rheinland-Pfalz als traditionell dezentral prüfendes Bundesland sowie Nordrhein-Westfalen, das im Jahr 2007 vom dezentralen auf ein zentrales Prüfungsverfahren umgestellt hat.

Vor dem Hintergrund der skizzierten strukturellen und fachdidaktischen Veränderungen möchte die Studie mehr als eine Beschreibung des Ist-Zustandes leisten. Daher ist sie als retrospektive Längsschnittstudie angelegt und umfasst die Analyse von Aufgaben im Zeitraum von 1993 bis 2008, um mögliche Veränderungen der schriftlichen Abiturprüfungsaufgaben in den letzten fünfzehn Jahren aufzuzeigen. Hinsichtlich des Implementationserfolgs von Innovationen im Bildungsbereich ist dies einerseits aus *fachdidaktischer Perspektive* interessant, da die EPA im Kontext der Diskussion um das unbefriedigende Abschneiden deutscher Schüler/innen bei TIMSS und PISA modifiziert wurden. Im Zuge der Überarbeitung wurden zwar die grundsätzliche Struktur und die Begrifflichkeiten der ursprünglichen EPA aus den Jahren 1989 und 1975 beibehalten; allerdings wurden aus fachdidaktischer Sicht auch innovative Elemente aufgenommen (vgl. Kühn 2010, S. 177f.). In diesem Zusammenhang berücksichtigen die jüngsten Fassungen der EPA aus den Jahren 2002ff. insbesondere Erkenntnisse aus Schulleistungsstudien sowie die Expertise von Wissenschaftler/innen, die Ergebnisse neuerer fachdidaktischer Forschung miteingebracht haben. Andererseits gilt der *schulstrukturellen Perspektive* ein besonderes Interesse, da viele Länder erst kürzlich vom dezentralen zum zentralen Prüfungsverfahren umgestellt haben (vgl. Kühn u.a. 2009), wodurch möglicherweise Effekte auf die Aufgabengestaltung zu erwarten sind: Befürworter/innen zentraler Prüfungsverfahren schreiben dieser Organisationsform aus fachdidaktischer Sicht ein Innovationspotenzial zu, da angenommen wird, zeitnah und großflächig (fach-)didaktische und -methodische sowie curriculare Neuerungen implementieren zu können; Kritiker/innen derartiger Prüfungsverfahren merken hingegen an, dass einzelne Lehrkräfte zwar im Unterricht fachinhaltliche Neuerungen und alternative Formen von Aufgabenstellungen und Leistungsnachweisen testen, dies jedoch im Kontext zentraler Prüfungen kaum geschehe, da die Vorgaben von Inhalten und Aufgabenstellungen von externer Stelle erfolgen (vgl. van Ackeren 2007).

Insgesamt wurden für Kurse mit erhöhtem Anforderungsniveau 596 Abituraufgaben mit 7 399 Operatoraufgaben (also jede inhaltsbezogene Denk- und Handlungsaufforderung; vgl. Jatzwauk 2007) untersucht. In den beiden zentral prüfenden Ländern Baden-Württemberg und Saarland wurden drei Prüfungsvorschläge je Untersuchungszeitraum analysiert, in den dezentral prüfenden Ländern Nordrhein-Westfalen und Rheinland-Pfalz jeweils zwölf Prüfungsaufgaben pro Analysezeitraum. In Nordrhein-Westfalen wurden für den dritten Zeitraum nur die beiden Zentralabiturjahrgänge 2007 und 2008 berücksichtigt.

Im Kontext der Diskussion um den Implementationserfolg von Reformen und ihr innovatives Potenzial steht auch deren tatsächliche *Umsetzung im Unterricht*

der gymnasialen Oberstufe im Fokus des Interesses. Dabei bietet die Analyse von Klausuraufgaben eine adäquate Möglichkeit, Rückschlüsse auf den vorangehenden Unterricht zu ziehen, da anzunehmen ist, dass Klausuren diesbezüglich einen repräsentativen Querschnitt abbilden (vgl. Sacher 1999). So wurden im Rahmen einer *zweiten Studie exemplarisch Klausuren der Qualifikationsphase* aus jeweils fünf Schulen aus Rheinland-Pfalz und Baden-Württemberg analysiert – insgesamt ergab sich hier eine Stichprobe von 891 Klausuraufgaben mit 4 342 Operatoraufgaben.

Zur Aufgabenanalyse wurde in Kooperation mit Fachdidaktiker/inne/n sowie erfahrenen Lehrkräften ein formal einheitliches, jedoch fachinhaltlich spezifisch ausgestaltetes Kategoriensystem entwickelt, das insgesamt 16 Analysekategorien vorsieht, die objektive, klar definierte Merkmalsbeschreibungen sowie Handlungsanleitungen für die Beurteiler/innen enthalten. Dabei werden drei Bereiche erfasst: Im Bereich der *deskriptiven Angaben* werden die Anzahl der Aufgaben pro Prüfungsaufgabe, die Anzahl der Teilaufgaben pro Aufgabe sowie die Anzahl der Operatoraufgaben pro Aufgabe aufgezeigt. Die *Sichtstruktur der Aufgaben* schließt alle formalen, sichtbaren Aufgabenmerkmale ein. Dazu gehört zunächst die Kategorie *Antwortformat*, welche die Art und Weise bezeichnet, in der die Bearbeitung einer Aufgabe erfolgt – gebundene oder freie Aufgabenbeantwortung –, und damit auf die äußere Form der Aufgabe fokussiert. Diese darf jedoch nicht mit der Kategorie *Offenheit* verwechselt werden, die sich auf verschiedene Lösungswege und mehrere denkbare richtige Ergebnisse bezieht. Darüber hinaus wird auch die *Art der Aufgabe* erfasst, wobei ein spezifischer Fokus auf die *Bedeutung von Experimenten* sowie – falls experimentelle Aufgaben durch die Prüflinge selbst durchzuführen sind – auf das *geforderte Experimentierverhalten* gelegt wird. Zuletzt wird im Rahmen der Sichtstruktur der Anwendungs- bzw. Alltagsbezug – der *Kontext* – der Aufgaben in die Betrachtungen mitaufgenommen. Darüber hinaus wird die *Sachstruktur der Aufgaben*, verstanden als inhaltsbezogene Aufgabenmerkmale, analysiert. Die Aufgaben werden in diesem Bereich zunächst bestimmten *Fachinhalten* zugeordnet, wobei hier zusätzlich die Zuordnung zum jeweiligen *Lehrplan* des entsprechenden Bundeslandes und damit auch die curriculare Validität der Aufgaben erfasst werden. Zudem werden die Aufgaben den Kategorien *Kompetenzbereiche* und *Anforderungsbereiche*, verstanden als kognitives Anspruchsniveau einer Aufgabe, zugeordnet. Unter fachdidaktischer Perspektive werden damit auch innovative Aufgabenmerkmale erfasst, die Bestandteil einer ‚neuen Aufgabenkultur‘ in den naturwissenschaftlichen Prüfungsfächern sein sollen (zur Diskussion um die Weiterentwicklung der Aufgabenkultur im naturwissenschaftlichen Unterricht vgl. BLK 1998).

Zur Sicherung der Qualität des Kategoriensystems wurden verschiedene Maßnahmen durchgeführt; dazu gehörte insbesondere ein mehrstufiges Training der Beurteiler/innen, welches nicht nur die generelle Anwendung des Kategoriensystems, sondern insbesondere den Umgang mit bestimmten Sonderfällen sowie die Festlegung von Grenzfallentscheidungen beinhaltete. Um festzustellen, inwieweit die Beurteile-

r/innen hinreichend trainiert waren, wurde die Beurteilerübereinstimmung berechnet, was als gebräuchlichste Methode zur Überprüfung der Objektivität und Reliabilität gilt. Nach Abschluss des Trainings konnte für alle Kategorien in allen Fächern eine Beurteilerübereinstimmung von  $\kappa \geq .75$  erzielt werden. Damit ist das entwickelte Kategoriensystem ein verlässliches Erhebungsinstrument zur Analyse von Abiturprüfungsaufgaben und Klausuraufgaben der Qualifikationsphase aus Kursen mit erhöhtem Anforderungsniveau in den drei naturwissenschaftlichen Unterrichtsfächern.

### 3. Zentrale Ergebnisse

Im Folgenden werden zentrale Befunde der Analyse von Abiturprüfungs- und Klausuraufgaben der gymnasialen Oberstufe vorgestellt (vgl. auch Kühn 2010) sowie zentrale Ergebnisse (über alle Kategorien hinweg) hinsichtlich des Implementationserfolgs der EPA sowie des angenommenen Innovationspotenzials zentraler Abiturprüfungen zusammengefasst.

#### 3.1 Zum Implementationserfolg der EPA – die fachdidaktische Perspektive

Wie bereits dargestellt, berücksichtigen die aktuellen EPA aus fachdidaktischer Sicht zahlreiche innovative Elemente im Hinblick auf die Aufgabenkonstruktion. Sie spiegeln zahlreiche Aspekte einer neuen, kompetenzorientierten Aufgabenkultur der naturwissenschaftlichen Fächer wider, die seitens der Fachdidaktik proklamiert werden (vgl. Kühn 2010).

Die Ergebnisse der Aufgabenanalysen können wie folgt zusammengefasst werden: Die Mehrheit der vorliegenden Abiturprüfungs- und Klausuraufgaben weist eine kleinschrittige Gliederungsstruktur – gekennzeichnet durch eine Vielzahl von Handlungsanweisungen – auf, ermöglicht kaum variable Lösungswege und -ergebnisse, enthält nur wenige außerfachliche, d.h. anwendungs- und lebensweltorientierte Aufgaben, ist durch eine Monostruktur an Aufgabenarten und Materialien ohne experimentelle Bezüge gekennzeichnet, fokussiert eher eng auf bestimmte Themenbereiche, weist nur wenige Transferaufgaben auf und dient primär der Reproduktion von Wissen, hingegen kaum der Anwendung von Kompetenzen. Innovative Elemente finden – mit wenigen Ausnahmen (s.u.) – kaum Berücksichtigung.

Die hier knapp skizzierten Merkmale von Abitur- und Klausuraufgaben gelten für die überwiegende Mehrheit aller im Rahmen der Studie untersuchten Prüfungsaufgaben. Allerdings zeigen sich länderspezifisch teilweise beträchtliche Unterschiede. Auch wenn die Vergleiche zwischen den Ländern oftmals nur auf einen kleinen bis mitt-

leren Effekt hindeuten, zeigen paarweise Ländervergleiche zum Teil erhebliche Divergenzen und das in nahezu allen Kategorien. Insgesamt lassen sich jedoch keine eindeutigen Ländertypen oder Vergleichsländer identifizieren, die in der Mehrheit der Kategorien Kongruenzen aufweisen oder sich grundsätzlich, d.h. kategorienübergreifend, unterscheiden. Unter den erfassten Bundesländern fällt keines durch eine verstärkte Umsetzung innovativer Aufgabenmerkmale besonders auf; vielmehr setzen die Länder, mit wenigen Ausnahmen (s.u.), die traditionelle Aufgabenkultur kontinuierlich fort. Vor dem Hintergrund der skizzierten Veränderungen im Hinblick auf die gymnasiale Oberstufe und die Abiturprüfung in den letzten Jahren wäre eigentlich anzunehmen gewesen, dass eine Weiterentwicklung der Aufgabenkultur im Abitur und damit eine Qualitätssteigerung stattgefunden hätten. Auf Grundlage der retrospektiven Längsschnittstudie lässt sich jedoch länderübergreifend feststellen, dass dies *kaum* geschehen ist – vielmehr werden mehrheitlich seit (mindestens) 15 Jahren vergleichbare Aufgaben gestellt, die prototypisches Abbild der skizzierten traditionellen Aufgabenkultur sind.

Zudem scheinen sich die beteiligten Akteure die Innovationen überwiegend eher nicht zu eigen gemacht zu haben: Insgesamt scheint das (mehrheitlich traditionelle) Verständnis der Aufgabenentwickler/innen und -steller/innen (Lehrkräfte, Aufgabenkommissionen, Mitglieder der Bildungsadministration) bezüglich der Gestaltung von Abituraufgaben, also die ‚Abituraufgabenkultur‘, den Steuerungsanspruch und das intendierte innovative Potenzial der EPA zu überlagern. Dies gilt auch für die Klausuraufgaben aus der Qualifikationsphase. Das innovative Potenzial, das den EPA zugeschrieben wird, kommt in der tatsächlichen Praxis *nicht* an – die Implementation der modifizierten EPA im Sinne einer *Weiterentwicklung* der gymnasialen Oberstufe und des Abiturs muss auf Grundlage der vorliegenden Stichprobe als *nicht erfolgreich* bezeichnet werden.

### 3.2 Zum Implementationserfolg der EPA – die strukturelle Perspektive

Hinsichtlich der Bewertung der Implementationswirkung der EPA gilt es zudem, schulstrukturelle Aspekte in die Betrachtungen zu integrieren. Wie bereits dargestellt, hat die Mehrheit der Länder in jüngster Zeit vom dezentralen auf ein zentrales Abiturprüfungsverfahren umgestellt, wenngleich die KMK beide Prüfungsverfahren als funktionale Äquivalente betrachtet, da die vereinbarten Prüfungsstandards bundesweit – unabhängig von der generellen Prüfungsorganisation – gelten. Trotz einer defizitären Forschungslage unterliegen zentrale Abiturprüfungen allerdings einer Art von öffentlichem Glauben daran, dass die vereinbarten Prüfungsstandards und innovative Elemente mit ihnen eher umzusetzen seien als mit dezentralen Prüfungssystemen.

Diese Annahme wird durch die vorliegende Stichprobe *nicht* bestätigt: Hinsichtlich des vermuteten Innovationspotenzials zentraler Prüfungsverfahren zeigt die re-

trospektive Längsschnittstudie, dass die beiden traditionell zentral prüfenden Bundesländer Baden-Württemberg und Saarland eine erhebliche Konstanz bezüglich der Gestaltung der Prüfungsaufgaben aufweisen und mit wenigen Ausnahmen kaum Veränderungen deutlich werden. Innovative Aspekte zur Aufgabengestaltung aus den modifizierten EPA finden nahezu keine Berücksichtigung. Auch im Zuge der Umstellung vom dezentralen zum zentralen Prüfsystem in Nordrhein-Westfalen zeigen sich weit weniger Unterschiede hinsichtlich der Umsetzung innovativer Elemente, als es die zuvor geführten Diskurse vermuten lassen, was sich mit den Befunden anderer Studien deckt (vgl. Maag Merki 2008; Maag Merki/Klieme/Holmeier 2008).

Hingegen weisen Abituraufgaben aus dezentralen Verfahren vergleichsweise mehr Veränderungen auf, wenngleich sich auch hier kein eindeutiger Trend feststellen lässt. Unter den dezentral gestellten Abituraufgaben aus jüngster Zeit finden sich auch einige (wenige) Aufgaben, die – aus fachdidaktischer Sicht – im Sinne von Good Practice-Beispielen konstruiert sind: Diese ermöglichen die Anwendung vielfältiger Kompetenzen in unterschiedlichen Kontexten zur Lösung fachspezifischer Fragen und Probleme, teilweise auch unter Einbezug real-experimenteller Aufgabenstellungen. Dabei kommen abwechslungsreiche Anwendungsaufgaben in variiierenden Kontexten zum Einsatz, die mehrere Vorgehensweisen und verschiedene Lösungsmöglichkeiten auf unterschiedlichen kognitiven Anforderungsniveaus zulassen. Jedoch handelt es sich lediglich um wenige Einzelfälle, die im Gros der traditionell gestalteten Aufgaben kaum auffallen. Dies deckt sich mit Befunden anderer Studien im Kontext der Implementationsforschung im angloamerikanischen Raum, wonach breit angelegte Innovationen nur in einer kleinen Zahl von Schulen realisiert und institutionalisiert worden sind, während sie für die Breite des Schulwesens weitgehend folgenlos blieben (vgl. Wiechmann 2002).

### 3.3 Auswirkungen auf den Unterricht

Im Hinblick auf den angenommenen Backwash-Effekt auf die Unterrichtsebene zeigt der direkte Vergleich zwischen den jeweiligen Klausur- und Abituraufgaben des Abiturjahrgangs 2008 aus Baden-Württemberg und Rheinland-Pfalz – bis auf vereinzelte Ausnahmen in den Fächern – bezogen auf einzelne Kategorien keine bedeutsamen Unterschiede, so dass eine hohe Kongruenz zwischen beiden Aufgabengruppen besteht. Es kann also festgestellt werden, dass die Schüler/innen anhand der Klausuren in der vorgelagerten Qualifikationsphase formal ‚optimal‘ auf die entsprechende Abiturprüfung vorbereitet werden, wobei diese Vorbereitung nicht mit dem übereinstimmt, was aus fachdidaktischer Sicht als optimal bezeichnet wird. Innovative Elemente lassen sich also auch in den Klausuraufgaben der gymnasialen Oberstufe kaum finden. Vor dem Hintergrund einer proportionalen Abbildung des Unterrichts durch die Klausuraufgaben (vgl. Sacher 1999) muss damit angenommen werden, dass ein Großteil der Unterrichtszeit für die abzuprüfenden Inhalte verwen-

det wird und dass dies im Sinne einer etablierten Unterrichtskultur über traditionelle Lehr- und Lernstrategien geschieht. Darüber hinaus ist auch zu vermuten, dass fachdidaktische Neuerungen, wie beispielsweise innovative Aufgabenformate, nicht ausreichend im Unterricht der gymnasialen Oberstufe verankert sind.

Insgesamt ist die Annahme, dass zentrale Abiturprüfungen besser als dezentrale geeignet seien, Innovationen zeitnah und flächendeckend umzusetzen, *nicht plausibel*. Allerdings, auch das zeigen die Ergebnisse der hier dokumentierten Studie, kann die von der KMK angenommene Äquivalenz zentraler und dezentraler Prüfungsverfahren nachgewiesen werden, allerdings in einem anderen Sinne als ursprünglich von dieser Instanz beabsichtigt: Die EPA werden in beiden Prüfungssystemen nicht wie vereinbart umgesetzt, sondern jedes Land – unabhängig von der generellen Prüfungsorganisation – legt diese so aus, wie es ihm angemessen erscheint. Dabei scheinen das Innovationspotenzial einzelner Lehrkräfte bzw. wechselseitige Adaptionprozesse größer zu sein als das Potenzial einer zentral verantwortlichen Aufgabenkommission.

#### 4. Fazit

Vor dem Hintergrund der dargelegten Befunde muss die Implementation der EPA als *wenig erfolgreich* bezeichnet werden: Innovative Elemente werden *kaum* in der tatsächlichen Aufgabenpraxis realisiert, sondern sie sind anscheinend im Rahmen der fachdidaktischen Diskurse im Kontext der Schulleistungsstudien formal in die modifizierte EPA aufgenommen worden, ohne dass sich an der vorherrschenden Aufgabenkultur grundsätzlich etwas geändert hat. Elmore (1996) sieht es sogar als zentralen Grund für das Misslingen zahlreicher Bildungsreformen an, dass zwar Merkmale auf der Makroebene erneuert werden, ohne dass sich allerdings tief sitzende Überzeugungen und Handlungsmuster von Lehrkräften, die zum Teil über Generationen tradiert werden, verändern (zit. nach Gräsel/Jäger/Willke 2006, S. 501). Die beteiligten Akteure scheinen sich die diskutierten Innovationsansätze zudem nicht zu eigen gemacht zu haben, sondern übernehmen offensichtlich wenige und nur jene Elemente, die mit ihren Überzeugungen, ihrem bisherigen Wissen und der bisher ausgeübten Praxis in Einklang stehen (vgl. ebd.). Insgesamt scheint das (mehrheitlich traditionelle) Verständnis der Aufgabenentwickler/innen und -steller/innen bezüglich der Gestaltung von Abituraufgaben, also die Abituraufgabenkultur, den Steuerungsanspruch und das innovative Potenzial der EPA zu überlagern. Dies trifft auch für die Klausuraufgaben aus der Qualifikationsphase zu und gilt für zentrale und dezentrale Prüfungsverfahren gleichermaßen.

Über mögliche Faktoren, die diese Entwicklung beeinflusst haben, kann an dieser Stelle nur spekuliert werden: Hinsichtlich der *Charakteristika der EPA* ist weitgehend unklar, ob und inwiefern die unmittelbar betroffenen Akteure eine Notwendigkeit

zur Weiterentwicklung der Aufgabenkultur sehen. Es wird hingegen deutlich, dass tradierte Praktiken im Sinne etablierter Aufgaben- und Prüfungskulturen, auch vor dem Hintergrund einer spezifischen (traditionellen) Unterrichtskultur, die Gestaltung der Aufgaben deutlich stärker beeinflussen als die administrativen Vorgaben der EPA. Eine weitere Ursache könnte auch in mangelnder Eindeutigkeit der Konstruktionsvorgaben liegen; so sehen die EPA bislang von Detailregelungen ab und haben bewusst Handlungsspielräume geschaffen, die länderspezifische Akzentuierungen ermöglichen. Diese Vorgehensweise suggeriert jedoch eine gewisse Unverbindlichkeit bzw. Uneindeutigkeit der Anforderungen.

Auch im Hinblick auf *lokale Prozesse und Charakteristika der Implementation* lassen sich nur Vermutungen anstellen: Sowohl in der Einzelschule als auch in den Aufgabenkommissionen kann der Umsetzungsprozess durch zahlreiche Faktoren beeinflusst werden. Dies betrifft zum einen die Schulleitung bzw. den oder die Kommissionsvorsitzende/n, die entscheidend darauf Einfluss nehmen können, ob eine Innovation in der gymnasialen Oberstufe der Einzelschule bzw. in den zentral gestellten Aufgaben übernommen und mit welcher Wirkung sie nachhaltig umgesetzt wird. Auch der Austausch im Kollegium bzw. der Fachkommission stellt einen weiteren potenziellen Einflussfaktor für die Verbreitung innovativer Elemente in Klausur- und Abiturprüfungsaufgaben dar. Letztlich kommt aber auch in diesem Kontext den Einstellungen und Verhaltensweisen der einzelnen Personen eine wichtige Bedeutung hinsichtlich des Implementationsprozesses zu. Maßgeblichen Einfluss auf die Gestaltung der Aufgaben hat zudem das Professionswissen der Aufgabenentwickler/innen, welches die konkrete Ausgestaltung der Aufgaben und damit auch die Umsetzung fachdidaktischer Innovationen mitbestimmt.

Insgesamt zeigt sich eine Fülle möglicher Einflussfaktoren. Diese reichen von der bildungspolitischen Ebene über die spezifischen lokalen Rahmenbedingungen der Implementation bis hin zu den Einstellungen und Dispositionen der einzelnen Akteure und umfassen auch die spezifischen Charakteristika von Innovation. Die dokumentierten Befunde bestätigen damit die Annahme schulischer Governance-Ansätze, dass die Umsetzung von Innovationen in den komplexen Akteurkonstellationen im Mehrebenensystem Schule *nicht* linear auf Grund formaler Weisungen der KMK erfolgt, sondern offensichtlich durch spezifische Übersetzungs- und wechselseitige Adaptionprozesse der beteiligten Akteure auf bildungsadministrativer, fach-, kurs- und schulkultureller Ebene – auch unter dem Einfluss informeller Regelungen und tradierter Praktiken.

Letztlich ist jedoch bislang noch unklar, ob und inwiefern die angeführten Faktoren die erfolgreiche Implementation der EPA konkret verhindert haben und in welcher Weise sie sich wechsel- und gegenseitig beeinflussen. Sie gelten damit als „black box“ im Kontext der Implementationsforschung (vgl. Fullan 1994, S. 2839) und sollten in weiterführenden empirischen Untersuchungen differenziert betrachtet werden.

## Literatur

- Ackeren, I. van (2007): Zentrale Abschlussprüfungen. Entstehung, Struktur und Steuerungsperspektiven. In: Pädagogik 59, H. 3, S. 12-15.
- Altrichter, H./Brüsemeister, T./Wissinger, J. (Hrsg.) (2007): Educational Governance. Handlungskoordination und Steuerung im Bildungssystem. Wiesbaden: VS.
- Altrichter, H./Maag Merki, K. (2010): Steuerung der Entwicklung des Schulwesens. In: Altrichter, H./Maag Merki, K. (Hrsg.): Handbuch neue Steuerung im Schulsystem. Wiesbaden: VS, S. 15-40.
- Berner, E./Oelkers, J./Reusser, K. (2008): Implementation von Bildungsstandards. Bedingungen des Gelingens (und Scheiterns) aus internationaler Sicht. 53. Beiheft der Zeitschrift für Pädagogik, S. 210-226.
- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK) (1998): Gutachten zur Vorbereitung des Programms Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts. Heft 60. Bonn.
- Demuth, R./Gräsel, C./Parchmann, I./Ralle, B. (2008): Chemie im Kontext: Von der Innovation zur nachhaltigen Verbreitung eines Unterrichtskonzepts. Münster: Waxmann.
- Fullan, M. (1994): Implementation of Innovations. In: Husen, T./Postlethwaite, T. (Hrsg.): The International Encyclopedia of Education. 2. Aufl. Oxford: Pergamon, S. 2839-2847.
- Gräsel, C. (2010): Stichwort: Transfer und Transferforschung im Bildungsbereich. In: Zeitschrift für Erziehungswissenschaft 13, H. 1, S. 7-20.
- Gräsel, C./Jäger, M./Willeke, H. (2006): Konzeption einer übergreifenden Transferforschung und Einbeziehung des internationalen Forschungsstandes. In: Nickolaus, R./Abel, M. (Hrsg.): Innovation und Transfer. Expertisen zur Transferforschung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 447-553.
- Gräsel, C./Parchmann, I. (2004): Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. In: Unterrichtswissenschaft 32, H. 3, S. 196-214.
- Jatzwauk, P. (2007): Aufgaben im Biologieunterricht. Eine Analyse der Merkmale und des didaktisch-methodischen Einsatzes von Aufgaben im Biologieunterricht. Berlin: Logos.
- Kühn, S. (2010): Steuerung und Innovation durch Abschlussprüfungen? Wiesbaden: VS.
- Kühn, S./Ackeren, I. van/Block, R./Klein, D. (2009): Zentrale Abiturprüfungen. Zur Heterogenität der Prüfungsverfahren in Deutschland. In: Schulverwaltung, Ausgabe Hessen und Rheinland-Pfalz 14, H. 10, S. 281-285.
- Maag Merki, K. (2008): Die Einführung des Zentralabiturs in Bremen. Eine Fallanalyse. In: Die Deutsche Schule 100, H. 3, S. 357-368.
- Maag Merki, K./Klieme, E./Holmeier, M. (2008): Unterrichtsgestaltung unter den Bedingungen zentraler Abiturprüfungen. Differenzielle Analysen auf Schulebene mittels Latent Class Analysen. In: Zeitschrift für Pädagogik 54, H. 6, S. 791-808.
- Oerter, R. (2010): Innovation. In: Oerter, R./Frey, D./Mandl, H./Rosenstiel, L. von/Schneewind, K.A. (Hrsg.): Neue Wege wagen. Innovation in Bildung, Wirtschaft und Gesellschaft. Stuttgart: Lucius & Lucius, S. 13-25.
- Prenzel, M. (2010): Geheimnisvoller Transfer? Wie Forschung der Bildungspraxis nützen kann. In: Zeitschrift für Erziehungswissenschaft 13, H. 1, S. 21-37.
- Sacher, W. (1999): Tests und Klausuren in der Schule. Wie mache ich das? In: Pädagogik 51, H. 4, S. 43-47.
- Thomas, R. (1994): Implementation of Educational Reform. In: Husen, T./Postlethwaite, T. (Hrsg.): The International Encyclopedia of Education. 2. Aufl. Oxford: Pergamon, S. 1852-1857.
- Trautwein, U./Köller, O./Lehmann, R./Lüdtke, O. (2007): Öffnung von Bildungswegen, erreichtes Leistungsniveau und Vergleichbarkeit von Abschlüssen. In: Trautwein, U./

- Köller, O./Lehmann, R./Lüdtke, O. (Hrsg.): Schulleistungen von Abiturienten. Regionale, schulformbezogene und soziale Disparitäten. Münster: Waxmann, S. 11-30.
- Trautwein, U./Neumann, M. (2008): Das Gymnasium. In: Cortina, K./Baumert, J./Leschinsky, A./Mayer, K.U./Trommer, L. (Hrsg.): Das Bildungswesen in der Bundesrepublik Deutschland. Reinbek: Rowohlt, S. 467-502.
- Wiechmann, J. (2002): Der Innovationstransfer in der Breite des Schulwesens. In: Zeitschrift für Erziehungswissenschaft 5, H. 1, S. 95-117.

*Svenja Mareike Kühn*, Dr., geb. 1983, wissenschaftliche Mitarbeiterin in der Arbeitsgruppe Bildungsforschung, Arbeitseinheit Bildungssystem- und Schulentwicklungsforschung, an der Universität Duisburg-Essen.

Anschrift: Universität Duisburg-Essen, Fakultät für Bildungswissenschaften, Institut für Pädagogik – AG Bildungsforschung, Campus Essen/Weststadttürme, 45117 Essen  
E-Mail: svenja.kuehn@uni-due.de

**WAXMANN**

Georg Stöckli, Rita Stebler

### Auf dem Weg zu einer neuen Schulform

**D**ie Grundstufe vereinigt den zweijährigen Kindergarten und die erste Grundschulklasse zu einer neuen Schulform und geht nicht mehr von einem zeitlich genau festgelegten Schuleintritt aus, sondern von den individuellen Voraussetzungen und Lernbereitschaften der Kinder. Daraus erwächst eine neue pädagogische Herausforderung: Individuelle Lernwege und lernstandsbezogene Zugänge zu den Kulturtechniken sind zu ermöglichen und zu unterstützen. Dieses Buch enthält Ergebnisse einer dreijährigen Begleitstudie, die den Unterricht und die individuelle Entwicklung der Kinder in fünf Grundstufenklassen des Kantons Zürich beleuchtet hat. Welche didaktischen Arrangements werden in der Grundstufe gestaltet? Wie werden Spielen und Lernen in den Unterrichtsalltag integriert? Wie wird die Altersdurchmischung genutzt? Wie entwickeln sich die Kinder im Verlauf der Grundstufenzeit? Es werden differenzierte Fallbeispiele geliefert, die mit Anregungen für die praktische Gestaltung des Grundstufenunterrichts verknüpft werden.



Georg Stöckli,  
Rita Stebler

### Auf dem Weg zu einer neuen Schulform

Unterricht und Entwicklung in der Grundstufe

2011, 292 Seiten  
Broschur: 39,90 €

ISBN 978-3-8309-2432-6  
E-Book: 35,90 €  
ISBN 978-3-8309-7432-1

Waxmann Verlag GmbH  
Steinfurter Straße 555  
48159 Münster  
Fon: 02 51 / 2 65 04-0  
Fax: 02 51 / 2 65 04-26  
E-Mail: order@waxmann.com  
www.waxmann.com