



Institut für Schulqualität der Länder
Berlin und Brandenburg e.V.



Expertise MSA / P10 Mathematik

Bezug der Prüfungsaufgaben in Berlin und Brandenburg
2006/07 zu den kompetenzorientierten Anforderungen
der länderübergreifenden Bildungsstandards für das
Fach Mathematik



Institut für Schulqualität der Länder
Berlin und Brandenburg

Impressum

Alle Rechte vorbehalten

© Institut für Schulqualität der Länder Berlin und Brandenburg e. V.

Otto-von-Simson-Str. 15

14195 Berlin

030 84416680

www.isq-bb.de

info@isq-bb.de

Berlin 2008

Vorbemerkung

Das Brandenburger Ministerium für Bildung, Jugend und Sport (MBS) hat gemeinsam mit der Senatsverwaltung für Bildung, Wissenschaft und Forschung (SenBWF) des Landes Berlin im Oktober 2007 in einem Perspektivpapier beschlossen, Möglichkeiten der Angleichung der Prüfung 10 (P10) in Brandenburg und des Mittleren Schulabschlusses (MSA) in Berlin zu prüfen.

Die zu diesem Zweck eingerichtete Arbeitsgruppe auf Ebene der Bildungsverwaltungen umfasst neben Vertretern des MBS und der SenBWF auch Vertreter des ISQ und des Landesinstituts für Schule und Medien Berlin-Brandenburg (LISUM).

In diesem Zusammenhang wurde das ISQ gebeten, eine Expertise in Auftrag zu geben. Sie sollte auf der Basis der Prüfungsaufgaben in Mathematik und Deutsch des Schuljahres 2006/07 für beide Länder vergleichend darstellen:

- ➔ ob und ggf. wie die jeweiligen Prüfungsaufgaben den Bezug zu den kompetenzorientierten Anforderungen der länderübergreifenden Bildungsstandards inhaltlich und fachdidaktisch herstellen;
- ➔ wie gut die in den Prüfungsarbeiten verwendeten Aufgaben- und Antwortformate mit den Qualitätsstandards der Aufgabenentwicklung im Rahmen der Bildungsstandards übereinstimmen;
- ➔ welche Veränderungen empfohlen werden, um die Leistungsüberprüfungen am Ende der Jahrgangsstufe 10 im Sinne der KMK-Bildungsstandards aufgabenseitig zu optimieren.

Für die Fächer Mathematik und Deutsch wurden vom ISQ getrennte Expertisen beauftragt. Die Gutachtergruppen setzten sich für beide Fächer aus Vertretern des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin und ausgewiesenen Expertinnen und Experten der Fachdidaktik des jeweiligen Faches zusammen.

Die Expertise für das Fach Mathematik wird hiermit vorgelegt.

Berlin, im Januar 2008

Dr. Hans Anand Pant

Wissenschaftliche Leitung und Geschäftsführung des ISQ

U N I K A S S E L
V E R S I T Ä T



Institut zur Qualitätsentwicklung
im Bildungswesen



Kommentierung der
schriftlichen Prüfungsarbeit zum mittleren
Schulabschluss 2007 im Fach Mathematik (Berlin)
bzw. der Prüfungen am Ende der Jahrgangsstufe 10
Mathematik (Brandenburg)
hinsichtlich ihrer Orientierung an den
länderübergreifenden Bildungsstandards

Alexander Roppelt
(Humboldt-Universität zu Berlin)

Katrin Keller
Dominik Leiß
Christina Drüke-Noe
Werner Blum
(Universität Kassel)

Im Auftrag des
The logo features the letters 'ISQ' in a bold, blue, sans-serif font. A red checkmark is positioned below the 'Q'.
Institut für Schulqualität der Länder
Berlin und Brandenburg e.V.

Inhalt

THEORETISCHE GRUNDLAGEN	3
Bildungsstandards in Deutschland	3
Die Bildungsstandards Mathematik	5
Konstruktionsprinzipien eines Tests im Sinne der Bildungsstandards	8
Vergleichbarkeit von Leistungstests	10
BEFUNDE	12
Kontexte realitätsbezogener Aufgaben	12
Verteilung der Aufgaben auf Leitideen, Kompetenzen und Anforderungsbereiche	15
Vergleichbarkeit der Abschlussprüfungen mit den Tests der Bildungsstandards	19
GESAMTBEWERTUNG UND FAZIT	21
LITERATUR	24
ANHANG: KLASSIFIKATION DER AUFGABEN	25
Berlin	25
Brandenburg	26

Theoretische Grundlagen

Ziel der vorliegenden Expertise ist es zu überprüfen, inwieweit die Aufgaben der zentralen Prüfungsarbeiten am Ende der 10. Klasse in Berlin und Brandenburg jeweils die in den Bildungsstandards geforderte Kompetenzorientierung aufweisen und mit den Aufgaben vergleichbar sind, die der Normierung und Überprüfung der Bildungsstandards dienen. Als Grundlage für die Bewertung der Aufgaben wird im Folgenden erinnert an die konzeptionellen Hintergründe für die Bildungsstandards in Deutschland im Allgemeinen und im Fach Mathematik im Besonderen.¹ Wir nennen außerdem Kriterien, die ein Test oder eine Prüfung im Fach Mathematik erfüllen sollte, um nach unserer Meinung dem Geist der Bildungsstandards zu entsprechen. Dieser theoretische Teil wird abgeschlossen mit grundsätzlichen Überlegungen und einer Rahmenkonzeption zur Vergleichbarkeit von Leistungstests.

Bildungsstandards in Deutschland

Die deutsche Kultusministerkonferenz (KMK) hat bekanntlich im Jahre 2003 als Reaktion auf die Ergebnisse bei der PISA-Studie beschlossen, für einige zentrale Fächer abschlussbezogene Bildungsstandards einzuführen. Sie hat sich dabei auch an anderen (bei PISA erfolgreicher) Staaten orientiert, in denen es schon länger verbindlich vorgegebene Standards für das gibt, was das Bildungssystem erreichen soll, meist verbunden mit einer größeren Selbstständigkeit der Schulen in Bezug darauf, wie dies erreicht werden soll. Konzeptionelle Grundlage für den KMK-Beschluss war die so genannte „Klieme-Expertise“ (Klieme et al. 2003). Hiernach beschreiben Bildungsstandards die fachbezogenen Kompetenzen, die Schüler bis zum jeweiligen Abschluss erwerben sollen. Sie beruhen auf breit verstandenen, fachlich verankerten und die Fachstruktur widerspiegelnden Bildungszielen, die in der Schule erreicht werden sollen und die wiederum eingebettet sind in fachübergreifende Bildungsziele. Insofern sind Bildungsstandards in ihrer Substanz also *Leistungsstandards* (und nicht etwa Unterrichtsstandards). Dieser Charakter schafft

¹ Diese beiden Kapitel basieren in wesentlichen Teilen auf Kapitel 1 in Blum, Drüke-Noe, Hartung und Köller (2006).

eine natürliche Verbindung zwischen den Bildungsstandards und zentralen Abschlussprüfungen, wie sie auch in Berlin und Brandenburg nunmehr etabliert sind.

In inhaltlicher Hinsicht besteht der wesentliche Fortschritt der Bildungsstandards gegenüber herkömmlichen Lehrplänen in der Art und Weise, wie die zu erreichenden Ziele formuliert sind. Während in Lehrplänen meist die im Unterricht zu behandelnden Inhalte im Zentrum stehen, werden in Bildungsstandards die zu erreichenden *Kompetenzen* genannt. Allerdings lassen sich Kompetenzen nur anhand von konkreten Fachinhalten erwerben, d. h., es gibt keinen Gegensatz zwischen Inhalten und Kompetenzen. Dementsprechend legen Bildungsstandards *auch* verbindliche Inhalte fest, konzentriert auf Kerninhalte der jeweiligen Fächer. Genauer ausdifferenziert werden die Kompetenzanforderungen in so genannten Kompetenzmodellen. Für das Fach Mathematik wird das Kompetenzmodell der Bildungsstandards unten genauer beschrieben.

Konkretisiert werden Kompetenzen, ihre Facetten und ihre Stufen durch *Aufgaben*, zu deren Bearbeitung diese Kompetenzen erforderlich sind. Bildungsstandards als Ganzes werden demgemäß konkretisiert durch eine breit gefächerte Sammlung von Aufgaben. Gleichzeitig werden Standards durch solche Aufgaben auch einer empirischen Überprüfung zugänglich, wobei die Annahme zu Grunde liegt, dass sich aus der Bearbeitung von Aufgaben mit einem hohen Grad an Zuverlässigkeit auf das Vorhandensein oder Fehlen entsprechender Kompetenzen beim Aufgabenlöser schließen lässt.

Mit der Einführung von Bildungsstandards sind zwei Erwartungen verbunden: eine höhere Zielklarheit und eine bessere Chance zur Überprüfung des Erreichten, kurz: *Orientierung* und *Evaluation*. Durch eine möglichst klare Formulierung von Kompetenzerwartungen sollen Bildungsstandards Lehrkräften eine Hilfe dabei geben, die Unterrichtsarbeit geeignet zu fokussieren. Dabei spielt die oben beschriebene Manifestation der Standards in Aufgaben eine entscheidende Rolle.

Standardbasierte Tests (Bildungsmonitoring), standardorientierte Tests (Vergleichsarbeiten oder zentralen Prüfungen eines Bundeslandes), standardorientierte schulinterne Tests oder Klassenarbeiten sind Beispiele für Evaluationssituationen, wobei Ebene und Ziel der Evaluation ebenso wie die verwendeten Instrumente unterschiedlich sind (mehr zu den Unterschieden diverser Testsituationen im

Kapitel „Vergleichbarkeit von Leistungstests“). All diese standardorientierten Evaluationen sind kein Selbstzweck; sie dienen dazu, Ansatzpunkte für Verbesserungen auf allen Ebenen zu identifizieren. Deshalb ist es wichtig, nicht erst am Ende von Bildungsgängen nach der Standarderreicherung zu schauen, sondern rechtzeitig vorher, um konkrete *Fördermaßnahmen* einleiten zu können. Evaluations- und Fördermaßnahmen müssen stets Hand in Hand gehen.

Das hinter der Einführung von Bildungsstandards stehende zentrale Ziel ist eine Steigerung der *Unterrichtsqualität*. Alle beschriebenen Maßnahmen wären letztlich zwecklos, wenn sie keine Auswirkungen auf den Unterricht hätten. Nur ein fachlich anspruchsvoller, kognitiv aktivierender und zum Lernen motivierender Unterricht bietet Schülern die Möglichkeit, die durch Standards gesetzten Anforderungen auch zu erreichen. Die konkrete Umsetzung (*Implementation*) von Bildungsstandards im Schulsystem bedeutet im Kern also Qualitätsentwicklung im Unterricht.

Die Bildungsstandards Mathematik

Das Kompetenzmodell, das den Bildungsstandards Mathematik zu Grunde liegt, greift in Teilen zurück auf Modelle, die sich in anderen Zusammenhängen bereits bewährt haben. Es werden dabei bekanntlich drei Dimensionen unterschieden, die man kurz als Prozess-, Inhalts- bzw. Anspruchsdimension bezeichnen kann: 1. die allgemeinen mathematischen Kompetenzen, 2. die inhaltsbezogenen mathematischen Kompetenzen, geordnet nach Leitideen, 3. die Anforderungsbereiche.

Zu 1.: Die *allgemeinen mathematischen Kompetenzen* (kurz: *Kompetenzen*, K1-K6) sind: Mathematisch argumentieren, Probleme mathematisch lösen, mathematisch modellieren, mathematische Darstellungen verwenden, mit Mathematik symbolisch / formal / technisch umgehen und mathematisch kommunizieren.

Diese Kompetenzen bilden den Kern der Mathematik-Standards. Hiermit (siehe dazu Niss 2003) werden zentrale Aspekte des mathematischen Arbeitens in hinreichender Breite erfasst. Es ist dabei weder möglich noch beabsichtigt, diese Kompetenzen scharf voneinander abzugrenzen. Vielmehr gibt es natürliche Über-

lappungen, und beim mathematischen Arbeiten treten i. A. mehrere Kompetenzen im Verbund auf.

Zu 2. Die *inhaltlichen Leitideen* (kurz: *Leitideen*, L1-L5) sind: Zahl, Messen, Raum und Form, funktionaler Zusammenhang sowie Daten und Zufall.

Diese Leitideen versuchen, die Phänomene (Freudenthal, 1983) zu erfassen, die man sieht, wenn man die Welt mit mathematischen Augen betrachtet. Zur „Welt“ gehört auch die mentale Welt unserer Gedanken und Ideen. Aus diesen Leitideen heraus haben sich die mathematischen Stoffgebiete Arithmetik / Größen, Geometrie, Algebra und Stochastik entwickelt. Leitideen und Stoffgebiete sind aber nicht identisch.

Zu 3. Kompetenzen zeigen sich insbesondere in Form von Tätigkeiten beim Aufgabenlösen. Die *Anforderungsbereiche* sollen den kognitiven Anspruch, den solche kompetenzbezogenen Tätigkeiten beinhalten, auf theoretischer Ebene erfassen. Man unterscheidet in den Bildungsstandards Mathematik drei Anforderungsbereiche, die mit „I. Reproduzieren“, „II. Zusammenhänge herstellen“ und „III. Verallgemeinern und reflektieren“ überschrieben sind. Diese Überschriften geben Orientierungen, sie dürfen jedoch nicht wörtlich genommen werden. Wenn etwa ein Zusammenhang mit einem einzigen Schritt hergestellt werden kann, wird man die Tätigkeit dem Bereich I zuweisen. Sind hierfür dagegen komplexe, nicht aus anderen Kontexten abrufbare Tätigkeiten nötig, wird man dies in Bereich III einordnen. Natürlich hängt dieser kognitive Anspruch mit der Aufgabenschwierigkeit zusammen – tendenziell sind Aufgaben aus Bereich III deutlich schwieriger für Schüler als Aufgaben aus Bereich I.

Mit dieser Beschreibung der drei Dimensionen wird der „Geist“ der Bildungsstandards noch nicht zureichend erfasst. Die Bildungsstandards Mathematik beziehen sich explizit auf die von H. Winter (1995) formulierten „Grunderfahrungen“, die jedem Schüler im Mathematikunterricht zu ermöglichen sind:

- Erscheinungen der Welt um uns, aus Natur, Gesellschaft und Kultur, mit Hilfe von Mathematik in einer spezifischen Art wahrzunehmen und zu verstehen.
- Mathematische Gegenstände als geistige Schöpfungen und als eine Welt eigener Art kennen zu lernen und zu begreifen.

- In der Auseinandersetzung mit Mathematik heuristische Fähigkeiten, die über die Mathematik hinausgehen, zu erwerben.

Die Bildungsstandards Mathematik sind der Versuch, diese allgemeinen Bildungsziele in Form weit gefächerter Kompetenzanforderungen zu erfassen und zu konkretisieren. Hieraus legitimiert sich auch die Verwendung des Begriffs *Bildungsstandards*.

Wie im einleitenden Kapitel „Bildungsstandards“ bereits dargelegt, werden die Bildungsstandards durch ein breites Spektrum von Aufgaben illustriert und konkretisiert. Im Folgenden werden die Anforderungen präzisiert, die speziell Mathematikaufgaben erfüllen müssen, um konform mit den Bildungsstandards zu sein.

Im alltäglichen Mathematikunterricht – so berichten übereinstimmend alle Untersuchungen – dominieren oft kalkül- und verfahrensorientierte Aufgaben, bei denen der Anspruchsgrad durch die technische Komplexität der auszuführenden Operationen bestimmt ist. Solche Aufgaben sind auch unter Kompetenzgesichtspunkten sinnvoll (es geht hierbei ja vorwiegend um die Kompetenz des *symbolisch/technisch/ formalen Arbeitens* mit Mathematik); sie stellen aber nur einen kleinen Ausschnitt aus dem Spektrum sinnvoller Aufgabenstellungen dar. Wenn wir von *kompetenzorientierten* Aufgaben sprechen, meinen wir vor allem solche, die nicht ausschließlich technische Fertigkeiten erfordern. Kompetenzorientierung ist *das* zentrale Kriterium für gute Aufgaben im Sinne der Bildungsstandards.

Ein anderer Aspekt erscheint selbstverständlich, muss jedoch immer wieder betont werden: Aufgaben sollen interessant und anregend sein, die Schüler neugierig machen und dazu motivieren, sich mit ihnen zu beschäftigen. Für Testaufgaben mag diese Forderung unnötig streng erscheinen, doch wegen der Vorbildfunktion von Prüfungen (s. u.) ist sie ebenfalls sinnvoll und geboten. Insbesondere müssen die Kontexte von realitätsbezogenen Aufgaben ernst genommen werden und dürfen nicht als bloße Hülle für Fragen dienen, die sich in diesem Kontext nicht stellen würden und bei denen die Mathematik lediglich wieder aus dem – ansonsten irrelevanten – Kontext „auszukleiden“ ist. Eine realitätsbezogene Mathematikaufgabe im Sinne der Bildungsstandards nutzt einen Kontext, um ihn durch Fragen, die er auf natürliche Weise aufwirft, Schritt für Schritt auszuloten und dabei ein gewisses Spektrum unterschiedlicher mathematischer Kompetenzen einzufordern.

Es sei an dieser Stelle nochmals betont, dass es nicht genügt, kompetenzorientierte Aufgaben für den Unterricht vorzusehen, sondern dass sich die Kompetenzorientierung auch in Klassen- und Abschlussarbeiten adäquat widerspiegeln muss. Das Unterrichtsgeschehen orientiert sich daran, welche Kompetenzen am Ende von den Schülern erwartet werden. Insofern beinhaltet Unterricht stets sinnvollerweise ein Stück *“Teaching to the Test“*. Dies erfordert aber, dass Tests (und insbesondere zentrale Abschlussprüfungen) Vorbilder für gute Aufgaben liefern und damit ihre steuernd-orientierende Wirkung auf den Unterricht positiv entfalten können.

Über die inhaltliche Qualität hinaus müssen Testaufgaben im Vergleich zu Aufgaben, die im Unterricht eingesetzt werden, noch zusätzlichen Bedingungen genügen, vor allem prinzipielle Verständlichkeit ohne externe Unterstützung, individuelle Bearbeitbarkeit in überschaubarer Zeit und verlässliche Korrigierbarkeit.

Konstruktionsprinzipien eines Tests im Sinne der Bildungsstandards

Eine Sammlung von Aufgaben, von denen jede für sich gut (im Sinne der Kompetenzorientierung) ist, ist zwar notwendige Voraussetzung, jedoch keine Garantie für eine gute Prüfung. Um den Geist der Bildungsstandards zu repräsentieren, gelten für einen *Test als Ganzes* zusätzliche Prinzipien, die im folgenden Kapitel umrissen werden sollen.

Über die Kompetenzorientierung der Aufgaben hinaus muss ein bildungsstandardorientierter Test selbstverständlich auch die hinlänglich bekannten Testgütekriterien aus der Testtheorie erfüllen: Er muss *valide* das messen, was man messen möchte; nämlich Mathematikkompetenz im Sinne der Standards. Dafür muss die Messung notwendigerweise hinreichend zuverlässig und genau, also *reliabel* sein. Und schließlich darf die Leistung jedes Getesteten nur von dessen Fähigkeit abhängen und beispielsweise nicht von den Umständen der Testdurchführung oder -korrektur beeinflusst werden (*Objektivität*).

Zusätzlich zu diesen eher allgemeinen Forderungen an einen guten Test gelten im Rahmen der Bildungsstandards Mathematik noch speziellere Spezifikationen,

um sicherzustellen, dass der Test als Ganzes alle Facetten mathematischer Kompetenz ausgewogen erfasst.

Wie oben beschrieben, bildet die Prozessdimension (allgemeine mathematische Kompetenzen) den Kern der Mathematik-Standards und muss deshalb stets in ihrer ganzen Breite in einem standardkonformen Test repräsentiert sein. D. h. jede allgemeine mathematische Kompetenz muss in substanzieller Menge vertreten sein und dies idealerweise auf jedem Anspruchsniveau (was aber nicht immer zu realisieren ist). Aufgrund der Natur von Mathematikaufgaben wird die allgemeine Kompetenz *mit Mathematik symbolisch / formal / technisch umgehen* (K5) überproportional häufig vertreten sein. Jedoch ist anzustreben, sie nicht bloß als Selbstzweck isoliert abzuprüfen – wie bisher häufig zu finden. Vielmehr sollte K5 vorrangig als Hilfsmittel in Aufgaben auftreten, die primär andere Kompetenzen ansprechen. Eine besondere Herausforderung der Aufgabenentwicklung ist es dabei, nichttriviale Aufgaben zum Argumentieren, Problemlösen und Modellieren zu finden, die auch außerhalb der Gymnasien noch eine vertretbare Lösungshäufigkeit haben.

Auch die Inhaltsdimension (Leitideen) sollte in einem solchen Test vollständig vertreten sein, wobei eine ungefähre Gleichverteilung der Leitideen sinnvoll ist. Dabei ist jedoch zu bedenken, dass es auch bei den inhaltlichen Kompetenzen teilweise erhebliche Überlappungsbereiche gibt, die in Statistiken häufig nicht in vollem Umfang abzubilden sind. Insbesondere gilt das innerhalb des üblichen Verfahrens, nur die primäre Leitidee einer Aufgabe zu berücksichtigen. Um bei begrenzter Testzeit diagnostische Aussagen über Fähigkeiten machen zu können, die nach Inhaltsbereichen differenziert sind, wird es gelegentlich nötig sein, inhaltliche Schwerpunkte zu setzen. Dies sollte dann aber explizit gesagt werden, damit nicht das falsche Signal gesendet wird, bestimmte Leitideen seien bevorzugt zu unterrichten.

Bei der dritten Dimension des Kompetenzmodells der Mathematik-Standards, den Anforderungsbereichen, ist keine Gleichverteilung, sondern eher ein Verhältnis von 1:2:1 anzustreben. Entscheidender als das genaue Verhältnis ist jedoch, dass in einem standardorientierten Test alle kognitiven Anspruchsniveaus gefordert werden. Dies gilt auch für Tests, die auf einzelne Bildungsgänge zugeschnitten sind. Insbesondere muss sich also auch in Tests für Hauptschulzweige Anforderungsbe-

reich III in substanzieller Menge finden lassen. Analog zu den allgemeinen mathematischen Kompetenzen stellt dies eine besondere Herausforderung an die Aufgabenentwicklung dar.

Die Formate der Aufgaben (Multiple Choice, Kurzantwort, offene Antwort) spielen im Sinne der Bildungsstandards eine eher untergeordnete Rolle. Allerdings ist zu beachten, dass das Aufgabenformat häufig konfundiert ist mit den Kompetenzen und dem Anspruchsniveau. Zum Beispiel wird die Fähigkeit zum eigenständigen mathematischen Argumentieren kaum anders als in einer offenen Frage zu demonstrieren sein.

Vergleichbarkeit von Leistungstests

Im vorangegangenen Kapitel wurden inhaltliche Spezifikationen beschrieben, welche die Menge aller Aufgaben eines Tests als Ganzes erfüllen muss, um dem Geist der Bildungsstandards zu entsprechen. Dies allein ist jedoch nicht hinreichend, um Testergebnisse auf den nationalen Skalen und den darauf beruhenden Kompetenzstufen verorten zu können. Das folgende Kapitel zeigt Kriterien auf, die nach internationaler Erfahrung darüber hinaus beim Vergleich von Leistungstests zu berücksichtigen sind. Da diese Thematik in Deutschland noch vergleichsweise unbekannt ist, führen wir sie hier etwas weiter als unbedingt notwendig aus.

In der internationalen testmethodischen Literatur wird die Vergleichbarkeit verschiedener Tests üblicherweise innerhalb der Thematik des so genannten *Linkings* diskutiert.² Dabei gilt der Grad der Ähnlichkeit der jeweils gemessenen Konstrukte als entscheidend für jeden Vergleich „verlinkter“ Tests: Messen die Tests das Gleiche? Hierbei sind inhaltliche Aspekte im engeren Sinne (wie sie im vorangehenden Kapitel beschrieben wurden) und die äußeren Rahmenbedingungen der Tests nicht voneinander zu trennen. Beispielsweise beeinflusst die Testlänge, inwiefern Ausdauer und Bearbeitungsgeschwindigkeit (*speededness*) für das Testergebnis von Bedeutung sind – und damit das gemessene Konstrukt: Wie lange muss man sich konzentrieren können? Kommt es darauf an, Aufgaben überhaupt lösen zu können, oder sollen sie innerhalb einer gewissen Zeit gelöst werden?

² Allgemein versteht man unter Linking das Herstellen einer Beziehung zwischen den Punktwerten zweier Tests, um diese miteinander vergleichen zu können (Kolen, 2004).

Den Grad der Ähnlichkeit der gemessenen Konstrukte kann man entweder quantitativ in Form der Stufen einer Taxonomie ausdrücken, wie sie etwa von Mislevy (1992) und Linn (1993) vorgeschlagen wird. Alternativ kann man die Gleichheit auch eher qualitativ an Übereinstimmungen hinsichtlich bestimmter Kriterien festmachen. Aufbauend auf einem umfassenden Report des amerikanischen National Research Council (Feuer, Holland, Green, Bertenthal und Hemphill, 1999) verbinden Kolen und Brennan (2004, vgl. auch Kolen, 2004) die beiden Ansätze, indem sie zunächst einen Kriterienkatalog aufstellen und die gewonnenen Befunde auf die Mislevy-Linn-Taxonomie anwenden. Auf diese Weise lässt sich die Stärke des Linking kriterienbasiert quantifizieren. Das Schema zur Charakterisierung von Linkingsituationen, das Kolen und Brennan (2004) vorschlagen, umfasst folgende Aspekte:

- *Schlussfolgerungen*: In welchem Umfang werden die Ergebnisse der Tests genutzt, um ähnliche Schlussfolgerungen zu ziehen? Dies ist im Wesentlichen die Frage, ob die Tests die gleichen Zwecke verfolgen (Bildungsmonitoring, Selbstevaluation, Vergabe von Abschlüssen, ...).
- *Konstrukte*: In welchem Umfang messen die Tests dieselben Konstrukte? Sind wirklich dieselben kognitiven Fähigkeiten zum Lösen der Tests erforderlich, oder gibt es doch unterschiedliche Facetten?
- *Populationen*: In welchem Umfang sind die Tests konstruiert, um in denselben Populationen eingesetzt zu werden? Zwei Tests mögen im Großen und Ganzen dasselbe Konstrukt messen, sind jedoch unter Umständen nicht für die gleiche Population geeignet; zum Beispiel, weil sie eine andere Sprache verwenden oder die Schwierigkeitsverteilung der Aufgaben eine andere ist.
- *Messbedingungen*: In welchem Umfang werden die Tests unter den gleichen Messbedingungen durchgeführt? Gemeint sind zum Beispiel Testlänge, Testformat, Durchführungsbedingungen, Art der Kodierung und ähnliche Dinge, die Auswirkungen auf die Ergebnisse haben könnten.

Eins-zu-eins vergleichbar sind Tests nur dann, wenn sie in *allen* genannten Kriterien übereinstimmen. Nur dann lassen sich die Punktwerte in einem Test in Punktwerte in einem anderen Test ohne weiteres übersetzen. In je mehr Kriterien sich die Tests jedoch mehr oder weniger deutlich unterscheiden, desto schwächer

wird das Linking und desto vorsichtiger müssen Interpretationen ausfallen, die auf Vergleichen beruhen. Insbesondere ist jeder Versuch, zentrale Abschlussarbeiten mit nationalen Kompetenzstufen zu vergleichen, mit erheblichen Unsicherheiten behaftet, so lange die Tests nicht in den genannten Punkten hinreichend übereinstimmen.

Befunde

Im Kapitel „Theoretische Grundlagen“ wurde ausgeführt, dass sich die Kompetenzorientierung der Bildungsstandards sowohl in den einzelnen Aufgaben als auch in der Zusammenstellung des Tests als Ganzes wiederfinden muss. Diesem Punkt widmen sich die nächsten beiden eher deskriptiven Kapitel. Das erste thematisiert speziell die Kontexte der realitätsbezogenen Aufgaben, da wir in diesem Punkt insgesamt die größten Optimierungsmöglichkeiten auf der Ebene der einzelnen Aufgaben sehen. Die Tests als Ganzes rücken im darauf folgenden Kapitel in den Blick. Darin werden Beobachtungen und Statistiken zur Verteilung der Aufgaben im Kompetenzmodell der Bildungsstandards referiert. Schließlich berichten wir unter Rückgriff auf die oben vorgestellte Rahmenkonzeption über die Vergleichbarkeit der vorliegenden Prüfungen mit den Tests, die zur Normierung der länderübergreifenden Bildungsstandards entwickelt worden sind.

Kontexte realitätsbezogener Aufgaben

Mit der Kompetenzorientierung gewinnen realitätsbezogene Aufgaben stärker an Bedeutung, und der Prozess des mathematischen Modellierens rückt in den Mittelpunkt. Der Kontext solcher Aufgaben und deren am Prozess des Modellierens orientierte Lösung machen den wesentlichen Unterschied zu traditionellen eingekleideten Textaufgaben, die letztlich neben einer „Auskleidekompetenz“ doch nur technische Fertigkeiten erfordern (vgl. auch Leuders und Leiß, 2006). Im folgenden Kapitel werden zunächst eher allgemeine Beobachtungen zum Thema Kontexte zusammengestellt, wobei die Prüfungsaufgaben aus Berlin und die aus Brandenburg zu unterscheiden sein werden. Schließlich wird beispielhaft an der Aufgabe „Bermudadreieck“ im Detail erläutert, wo unseres Erachtens Verbesserungsbedarf besteht.

Die Prüfungsaufgaben aus Berlin unterscheiden sich von denen aus Brandenburg spürbar. In Berlin ist insgesamt eine ausgewogenere Mischung aus traditionellen Aufgaben zum Grundwissen und aus Aufgaben zu beobachten, die ein breiteres Kompetenzspektrum abrufen. Die Kontexte der Berliner Aufgaben wirken realistischer und kommen weitgehend ohne Einkleidungen aus. Gleichzeitig stehen die einzelnen Teilaufgaben nicht losgelöst vom Kontext für sich allein, sondern besitzen erkennbare Bezüge untereinander. In Hinblick darauf, dass im Gesamttest anspruchsvolle, eigenständige Modellierungen fehlen (vgl. nächstes Kapitel), wäre es allerdings überlegenswert, ob auf die Hinweise zum Lösungsweg bei den Aufgaben 3 („Laser“) und 8 („Hubschrauber“) nicht besser verzichtet worden wäre. Darüber hinaus ist der Graph in Aufgabe 5 („Auto“) gegenüber realistischen Beschleunigungsvorgängen so stark vereinfacht, dass es wünschenswert wäre, dies deutlicher zu thematisieren. Dies könnte z. B. zum Inhalt einer weiteren Teilaufgabe gemacht werden. Fraglich erscheint außerdem, ob es ein gutes Bild von praktischer Genauigkeit liefert, wenn die Fahrzeit von Berlin nach Moskau auf eine Minute genau angegeben wird (selbst wenn dies mit Fahrplänen theoretisch möglich ist).

Im Vergleich hierzu müssen bei den Brandenburger Versionen deutlich mehr problematische Punkte konstatiert werden.³ Dies beginnt mit der Formulierung der Aufgaben, die z. T. unklar und unnötig kompliziert gerät und mitunter Hinweise auf anzuwendende Verfahren enthält, die die Lösungsvielfalt einschränken (auf die Anweisung „rechnerisch“ in Aufgabe 4.2 c) ließe sich etwa verzichten). Vor allem aber ist die Vielzahl künstlich konstruierter Kontexte zu bemängeln. Zum Beispiel ist bei der Aufgabe 4.2 („Camcorder“) der Abstand zwischen Kaufentscheidung und tatsächlichem Kauf unwahrscheinlich groß (was aber noch vertretbar sein mag). Entscheidender ist, dass die Angebote, die der Vater seinem Sohn Felix macht, sehr konstruiert erscheinen und offenkundig nur dazu dienen, lineares und exponentielles Wachstum miteinander zu vergleichen. Die einzelnen Teilaufgaben zu den gewählten Kontexten werden zudem oft ohne inhaltlichen Bezug aneinandergereiht. Für die Teilaufgaben d) und e) der Wahlaufgabe 4.3 etwa ist der ur-

³ Der Einfachheit halber beziehen wir uns bei konkreten Aufgabenformulierungen und -numerierungen stets auf die Version für die Realschulen (wenn nichts anderes angegeben).

sprüngliche Kontext des Postkastens vollkommen irrelevant. Teilweise leiden die Aufgaben unter dem Bemühen, sie an verschiedene Schultypen und -niveaus anzupassen. So erscheint die Version der Aufgabe 3 („Bermudadreieck“) aus der Abschlussarbeit der Gesamtschule für den Grundkurs gelungener als ihr Pendant für die Realschule (vgl. die Anmerkungen zu den einzelnen Teilaufgaben unten).

Um die beschriebenen Probleme zu verdeutlichen, betrachten wir als Beispiel die Aufgabe 3 „Bermudadreieck“ im Detail. Ein grundsätzliches Problem der Aufgabe ist, dass der gesamte Kontext per se unklar ist, da in der Literatur z. B. die Angaben zur Zahl der verschollenen Objekte, zur Größe des Bermudadreiecks und dessen Begrenzung je nach Informationsquelle voneinander abweichen. Vor diesem Hintergrund lässt sich die Genauigkeit, die in den einzelnen Teilaufgaben gegeben ist, nicht rechtfertigen. Andererseits würde gerade eine gewisse „Unbestimmtheit“ des Kontextes Gelegenheiten für kognitiv anspruchsvollere Aufgaben bieten. Statt die Schüler Vereinfachungen und Annahmen selbst treffen zu lassen, sind diese jedoch weitgehend schon in der Aufgabenstellung vorweggenommen. Daher liegen alle Teile dieser langen Aufgabe in Anforderungsbereich I oder II.

Zum Abschluss dieses Kapitels seien exemplarisch noch die Teilaufgaben des „Bermudadreiecks“ im Einzelnen kommentiert:

- a) Vermutlich um die Aufgabe anspruchsvoller zu machen, wird im Gegensatz zur Version des Gesamtschul-Grundkurses nicht nach dem Prozentsatz, sondern nach dem Grundwert gefragt – eine Aufgabenstellung, die in der „wirklichen Welt“ so wohl kaum auftreten wird. In Verbindung mit der ggfs. verwirrenden Folge von Begriffen („auch 54 Segelschiffe“, „36,73 % aller Schiffe und Flugzeuge“, „wie viel [*sic!*] Objekte“) und der nicht sinnvollen Genauigkeit des Prozentsatzes wirkt die Aufgabe in dieser Version sehr konstruiert.
- b) Auch wenn sie nur ein relativ schmales Kompetenzspektrum ansprechen, ist gegen diese Teilaufgaben an sich nichts einzuwenden. Sie stehen allerdings weder in Bezug zum Bermudadreieck noch zu einer der folgenden Teilaufgaben.
- c) Bei dieser Teilaufgabe stellt sich die Frage, wozu jemand die Fläche des Bermudadreiecks wissen wollen könnte. Sollte sie (evtl. in der Schiff-

fahrt?) in irgendeiner Weise von Bedeutung sein, wäre ein Hinweis wünschenswert. Etwas besser ist es da schon, nicht von der Fläche des Bermudadreiecks, sondern von der des Dreieck MSB zu sprechen (wie in der Grundkursversion), da die Aufgabe damit einen eher innermathematischen Charakter erhält. Die Genauigkeit der Angaben (Winkel α !) ist auch in dieser Aufgabe wenig sinnvoll. Davon unabhängig ließen sich die Anforderungen an die Modellierungskompetenz erhöhen, wenn auf den Hinweis zur Erdkrümmung und auf die Skizze verzichtet würde.

- d) Der Kontext dieser Aufgabe wirkt konstruiert: Wenn es möglich ist, den (rechten) Winkel MPS zu bestimmen, dann ist sicherlich auch die zurückgelegte Fahrstrecke bekannt (auch hier wirkt die Grundkursversion weniger künstlich). Sollte es jedoch eine Möglichkeit zur Winkelbestimmung geben, bei der dies nicht notwendigerweise der Fall ist, so wäre auch hier ein Hinweis darauf wünschenswert.
- e) Diese letzte Teilaufgabe soll positiv herausgehoben werden: Auch wenn die Modellierungskompetenz keine vordergründige Rolle spielt, enthält die Aufgabe einen realen Kern und verlangt eine gewisse Flexibilität im Umgang mit Vergleichsgrößen. Fragestellungen dieser Art dürften von größerer Relevanz sein als etwa die Fläche des Bermudadreiecks. Allerdings fehlt auch hier der Bezug zum Rahmen-Kontext.

Verteilung der Aufgaben auf Leitideen, Kompetenzen und Anforderungsbereiche

In diesem Kapitel stehen die Tests als Ganzes und vor allem die Abdeckung des gesamten Kompetenzspektrums im Fokus, wie sie im Abschnitt zu den „Konstruktionsprinzipien eines Test im Sinne der Bildungsstandards“ beschrieben ist. Da sich die Verteilung der Aufgaben auf Leitideen, Kompetenzen und Anforderungsbereiche in Berlin und in Brandenburg nur in Nuancen unterscheidet, behandeln wir die Prüfungen weitgehend gemeinsam. In Brandenburg konzentrieren wir uns außerdem – wie im letzten Kapitel – auf die Version für die Realschulen. Die Wahlaufgaben in Brandenburg gehen in die berichteten Statistiken so ein, als ob sie alle zu bearbeiten wären, da sie am ehesten alle zusammen für das erwartete Kompetenz-

spektrum stehen. Eine Tabelle mit der Klassifikation jeder einzelnen Teilaufgabe gemäß Leitidee, Kompetenzen und Anforderungsbereich findet sich im Anhang.

Die Verteilungen der Aufgaben (genauer: Teilaufgaben) auf die Leitideen, Kompetenzen und Anforderungsbereiche ist für die Berliner Prüfung in Abbildung 1 und für die Brandenburger Version für Realschulen in Abbildung 2 dargestellt. Der augenfälligste Befund ist, dass die Kategorien in allen Dimensionen deutlich verschieden stark besetzt sind. Im Folgenden wird berichtet, was dies im Einzelnen bedeutet.

Bei den Leitideen fällt als erstes auf, dass *Raum und Form* (L3) stark unterrepräsentiert ist: Nur zwei Teilaufgaben in Brandenburg und gar keine in Berlin sind dieser Leitidee zuzuordnen. Ein ähnliches Bild bietet *Daten und Zufall* (L5): Während es in Brandenburg zumindest einige Aufgaben aus diesem Inhaltsbereich gibt, fehlt er in Berlin wiederum vollständig.⁴ Diese Ungleichverteilung ist zwar ungewöhnlich deutlich ausgeprägt, spiegelt jedoch letztlich wohl die schulische Praxis wider.

Ein mindestens ebenso starkes Ungleichgewicht tritt bei den allgemeinen mathematischen Kompetenzen auf. Selbst wenn man ein Übergewicht des *symbolisch/technisch/ formalen Arbeitens* (K5) als „natürlich“ für die Mathematik (insbesondere für die Schulmathematik) ansieht, dürfen andere Kompetenzen nicht ausgespart werden. Tatsächlich werden in den Prüfungen beider Länder die Kompetenzen *Probleme mathematisch lösen* (K2) und vor allem *mathematisch Argumentieren* (K1) und *mathematisch Kommunizieren* (K6) nur sehr selten gefordert. Darüber hinaus ist zu der relativ großen Zahl an Aufgaben zum *mathematischen Modellieren* (K3) zu sagen, dass diese Aufgaben offenbar primär dazu dienen, innermathematische Themen zu behandeln und – wie oben ausführlich dargelegt – der Kontext beim Lösen eher vernachlässigt werden kann. Wünschenswerte Modellierungsaktivitäten wie zum Beispiel das Treffen von Annahmen beim Strukturieren einer Realsituation sind höchstens implizit nötig.

⁴ Das Fehlen von *Daten und Zufall* in Berlin ist möglicherweise einer Übergangsregel geschuldet.

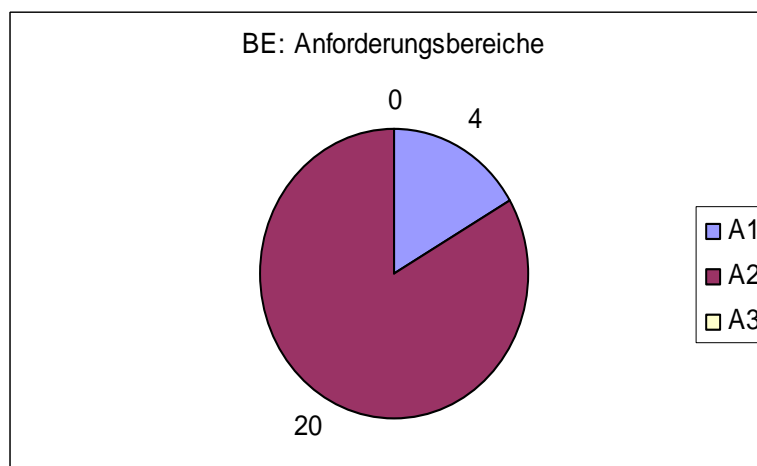
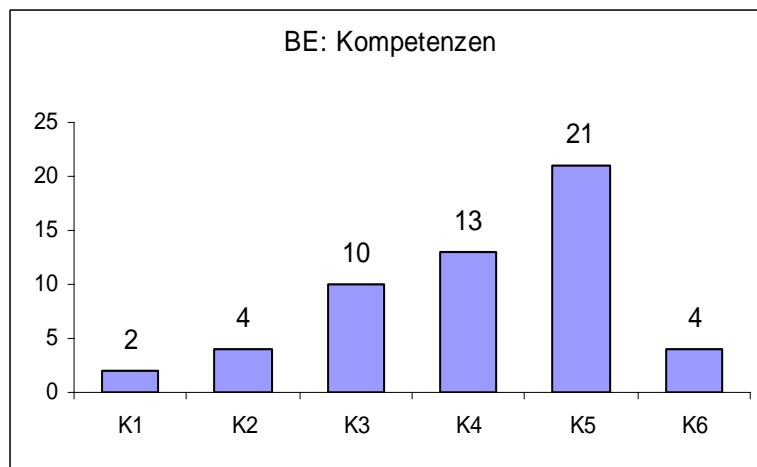
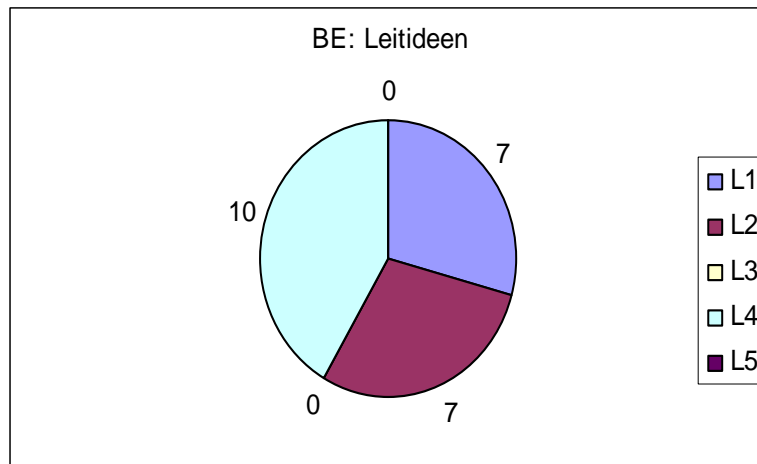


Abbildung 1: Verteilung der Leitideen, Kompetenzen und Anforderungsbereiche in der Berliner Prüfung. Die Zahlen geben die Anzahl der Teilaufgaben in der jeweiligen Kategorie an.

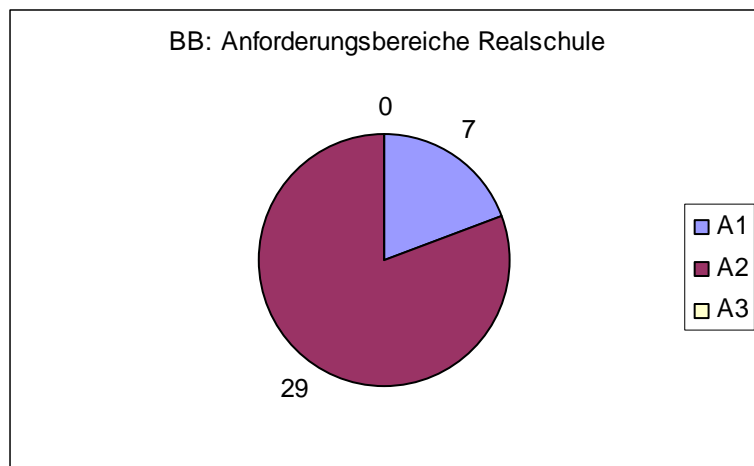
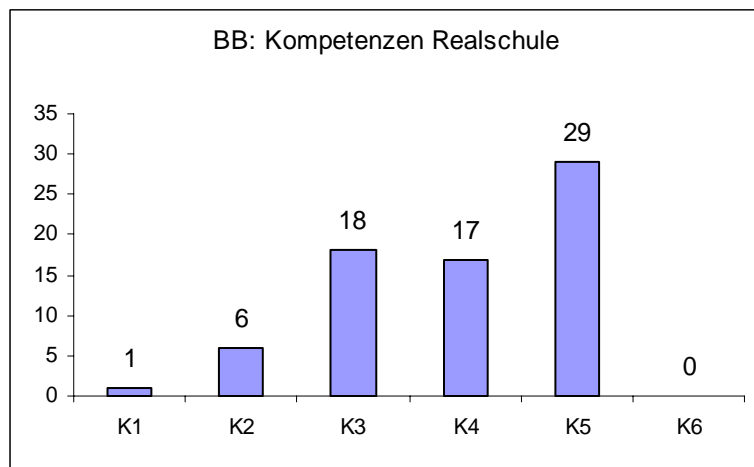
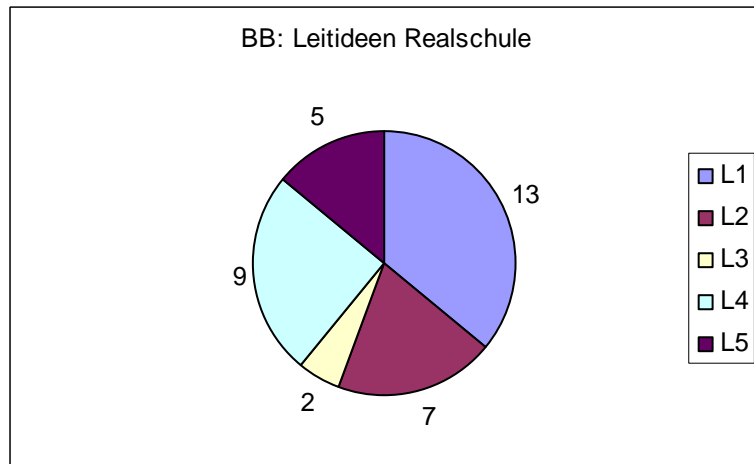


Abbildung 2: Verteilung der Leitideen, Kompetenzen und Anforderungsbereiche in der Brandenburger Prüfung für Realschulen. Die Zahlen geben die Anzahl der Teilaufgaben in der jeweiligen Kategorie an. Für die Statistik wurden alle Wahlaufgaben gleichermaßen berücksichtigt.

Der letzte Punkt berührt die dritte Dimension des Kompetenzmodells der Mathematik-Standards, die Anforderungsbereiche. Analysiert man die Verteilung der Aufgaben auf die Anforderungsbereiche (AB), so fällt neben der starken Dominanz von AB II (was z. T. vielleicht auch in dessen Konzeption begründet ist) insbesondere auf, dass kein Item zu AB III vorliegt. Müsste der Anspruch eigentlich lauten, in allen Bildungsgängen Aufgaben aus diesem Bereich zu stellen, so ist besonders verwunderlich, dass nicht einmal im Gymnasium eine höhere kognitive Komplexität angestrebt wird. Aber auch die geringe Anzahl an Aufgaben aus AB I erscheint für einen Leistungstest im „Realschulbereich“ sehr problematisch. Das Problem verschärft sich noch dadurch, dass es sich bei den vorliegenden – keinesfalls immer leichten – Testaufgaben häufig um solche handelt, die nur aufgrund der Mehrschrittigkeit der (oft technischen und im Einzelnen eher wenig komplexen) Anforderungen dem AB II zuzuordnen sind. Eine globale Sichtweise auf das in einer Aufgabe gestellte Problem und ein vertieftes begriffliches Verständnis werden kaum einmal verlangt.

Vergleichbarkeit der Abschlussprüfungen mit den Tests der Bildungsstandards

Mit der Synopse von Neumann (2006) liegt bereits eine umfassende Gegenüberstellung der Prüfungen zum Mittleren Schulabschluss in Berlin und Brandenburg vor. Auf diese aufbauend wird im Folgenden die Vergleichbarkeit der vorliegenden Abschlussprüfungen mit den länderübergreifenden Tests zur Normierung der Bildungsstandards diskutiert, die unter der Ägide des IQB entwickelt worden sind. Zu diesem Zweck betrachten wir die von Kolen und Brennan (2004) genannten vier Kriterien, wie sie im theoretischen Hintergrund vorgestellt wurde.

Schlussfolgerungen. Während die Zielsetzungen der Abschlussprüfungen der beiden Bundesländer naturgemäß sehr ähnlich sind (der Leistungsstand jedes Schülers soll unter einheitlichen Bedingungen festgestellt werden), verfolgen die Tests des IQB einen anderen Zweck, nämlich die Normierung und Illustrierung der Bildungsstandards. Sie sind ein Instrument des Monitorings auf der Ebene des Bildungssystems. Folglich sind die Schlussfolgerungen, die aus den Tests jeweils gezogen werden, gänzlich anderer Art. Werden aufgrund der Ergebnisse in den Abschluss-

prüfungen Schlussfolgerungen über *Individuen* gezogen, die für diese persönliche, unter Umständen andauernde Folgen haben, stehen bei den länderübergreifenden Tests des IQB *Systeme* auf dem Prüfstand. Die Unterschiede im Zweck der Tests haben weit reichende Konsequenzen auf verschiedensten Gebieten. Sie beginnen beim Testdesign, gehen über die Vorbereitung auf die Tests in den Schulen bis hin zur Motivation der Schüler während der Testsitzung.

Konstrukte. In den vorherigen Kapiteln wurde gezeigt, dass sich die vorliegenden Prüfungen sowohl in den einzelnen Aufgaben als auch in der Zusammensetzung des Tests als Ganzes doch spürbar von den Bildungsstandardtests unterscheiden, mithin die gemessenen Konstrukte verschieden sind. Darüber hinaus darf nicht vergessen werden, dass letztlich auch alle Unterschiede, die unter den anderen Stichworten dieses Kapitels aufgeführt sind, Nebenwirkungen auf das gemessenen Konstrukt besitzen (vgl. die Ausführungen im theoretischen Hintergrund).

Andererseits bleibt der Testgegenstand der Abschlussprüfungen der Länder ebenso wie der Normierungstests des IQB erklärtermaßen die mathematische Kompetenz von Schülern am Ende der Sekundarstufe I. Deshalb dürften die Effekte auf interindividuelle Unterschiede in den Testergebnissen wahrscheinlich nicht allzu groß sein. So weiß man zum Beispiel aus dem PISA-Längsschnitt, dass Tests, die zwar auf unterschiedliche theoretische Konzeptionen ausgerichtet sind (Literacy-Konzept versus Curriculum-Orientierung), jeweils aber Mathematik zum Inhalt haben, sehr ähnliche Ergebnisse produzieren können (Ehmke, Blum, Neubrand, Jordan und Ulfig, 2006).

Populationen. Die größte Ähnlichkeit zwischen den vorliegenden Prüfungen und den Normierungstests des IQB liegt naturgemäß in der Zielpopulation: In beiden Fällen werden Schüler kurz vor dem Mittleren Schulabschluss getestet, wenn auch die IQB-Tests deutschlandweit eingesetzt werden.

Messbedingungen. In den Bedingungen der Messung liegen die wohl auffälligsten Unterschiede; und zwar sowohl zwischen den Prüfungen der beiden Länder als auch zwischen diesen und den länderübergreifenden Tests des IQB. Die Testlänge schwankt zwischen 120 und 160 Minuten. Das Testformat unterscheidet sich sowohl hinsichtlich des Anteils von offenen und geschlossenen Aufgabenformaten als auch darin, in welcher Form die Antworten gegeben werden müssen (ins Test-

heft oder auf ein separates Blatt). Während für die Bildungsstandards voll standardisierte Durchführungsbedingungen vorgegeben sind (externe Testleiter mit strikt vorgegebenem Skript), genügen für die Abschlussprüfungen vergleichsweise weiche Vorgaben. Der vielleicht folgenreichste Unterschied liegt jedoch in der Organisation der Korrektur, die bei den länderübergreifenden Tests grundsätzlich extern erfolgt, während sie bei den Abschlussprüfungen vollständig in der Hand der Schulen liegt.

Zusammenfassend muss also festgehalten werden, dass sich die Abschlussprüfungen aus Berlin und Brandenburg und die Tests zur Normierung der Bildungsstandards für den Mittleren Abschluss – z. T. in natürlicher Weise – bezüglich fast aller Kriterien von Kolen und Brennan (2004) unterscheiden. Das Ausmaß einer möglichen Verknüpfung der Ergebnisse der Tests wäre demnach eher gering. In der Taxonomie von Linn und Mislevy wäre sie der Stufe „Projection“ zuzuordnen, die Linn (1993) als „weakest form of statistical linking“ bezeichnet.

Gesamtbewertung und Fazit

Im Kern dieser Expertise steht die mithilfe der Bildungsstandards angestrebte Kompetenzorientierung. Nachdem wir im „Theoretischen Hintergrund“ deren zentrale Rolle für die Bildungsstandards erläutert haben, legt der Abschnitt „Befunde“ im Detail dar, inwieweit die vorliegenden Prüfungen aus Berlin und Brandenburg diesem „Geist“ der Bildungsstandards entsprechen. Dies fassen wir im Folgenden nochmals in einer Gesamtbewertung zusammen.

Betrachtet man die Tests beider Länder jeweils als Ganzes, zeigt sich, dass es bezüglich des Ausschöpfens der Breite des Kompetenzspektrums noch Entwicklungsbedarf gibt. So sollten die Leitideen *Raum und Form* (L3) und *Daten und Zufall* (L5) eine stärkere Berücksichtigung finden, auch um damit einer entsprechenden Benachteiligung dieser Inhaltsbereiche im Unterricht entgegenwirken zu können. Noch wichtiger im Sinne der Kompetenzorientierung wäre es, die allgemeinen mathematischen Kompetenzen *Probleme mathematisch lösen* (K2) und vor allem *mathematisch Argumentieren* (K1) und *mathematisch Kommunizieren* (K6) stärker in den Prüfungen zu berücksichtigen. Mehrschrittig-technische Aufgaben dominieren insgesamt noch zu sehr das Bild der Prüfungen. Auch hier ist aufgrund der Vorbereitung der

Schüler auf den Test von einer Rückkopplung mit dem Unterricht auszugehen. Des Weiteren sollte ein stärkeres Gewicht auf ein begriffliches Verständnis gelegt und sollten unbedingt Aufgaben aus dem Anforderungsbereich III gestellt werden (in allen Schulformen). Ein Schritt in diese Richtung könnte darin bestehen, bei einigen Anwendungsaufgaben Vereinfachungen und Annahmen nicht vorwegzunehmen und dadurch die Modellierungsanforderungen in vollem Umfang zu erhalten.

Trotz dieser Unausgewogenheit bietet die vorliegende Prüfung aus Berlin ein insgesamt befriedigendes Bild. Sie zeichnet sich durch eine ausgewogene Mischung traditioneller Aufgaben zum Grundwissen und solcher Aufgaben aus, die ein breiteres Kompetenzspektrum abrufen. Die einzelnen Teilaufgaben loten die der gesamten Aufgabe zugrunde liegenden Kontexte aus, welche wiederum weitgehend realistisch wirken. Allerdings ist darauf zu achten, dass die Genauigkeiten in den Angaben stets sinnvoll gestaltet werden.

Obwohl auch bei den Brandenburger Abschlussarbeiten eine positive Tendenz zu Modellierungsaufgaben zu bemerken ist, können diese insgesamt noch nicht als zufriedenstellend bezeichnet werden. Neben dem oben für beide Länder gemeinsam beschriebenen Ungleichgewicht bezüglich der Leitideen, Kompetenzen und Anforderungsbereiche ist vor allem der Umgang mit den Kontexten realitätsbezogener Aufgaben zu kritisieren. Insbesondere sollten eingekleidete Aufgaben, die von den Schülern nur entkleidet und anschließend mittels Routineverfahren bearbeitet werden müssen, im Umfang deutlich reduziert werden. Des Weiteren sollten die einzelnen Teile einer Aufgabe im direkten Bezug zum ausgewählten Kontext stehen und diesen möglichst umfassend ausloten. Dabei muss die Idee, ausgehend von einer Stammaufgabe schulformdifferenzierende Aufgabenvariationen zu entwickeln, keineswegs aufgegeben werden. Allerdings ist zu berücksichtigen, dass die Bildungsstandards nicht nach Schulformen, sondern lediglich nach Abschlüssen differenzieren.

Abschließend sei noch bemerkt, dass es sehr wünschenswert ist, die inhaltlichen Spezifikationen der zentralen Abschlussprüfungen möglichst ähnlich zu denen der Tests zur Normierung der länderübergreifenden Bildungsstandards zu gestalten. Es werden dennoch stets zahl- und folgenreiche Unterschiede zwischen

diesen Tests bestehen bleiben, die sich vor allem aus den unterschiedlichen Zielsetzungen ergeben. Dies ist in gewisser Weise natürlich und an sich gänzlich unproblematisch, muss jedoch bei jeder Interpretation, die die Tests in Bezug zueinander setzt, und insbesondere bei jedem quantitativen Vergleich unbedingt berücksichtigt werden.

Literatur

- Blum, W., Drücke-Noe, C., Hartung, R. & Köller, O. (Hrsg.). (2006). *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. Berlin: Cornelsen Scriptor.
- Blum, W. & Leiß, D. (2005). Modellieren im Unterricht mit der "Tanken"-Aufgabe. *mathematik lehren*, 128, 18-21.
- Ehmke, T., Blum, W., Neubrand, M., Jordan, A. & Ulfig, F. (2006). Wie verändert sich die mathematische Kompetenz von der neunten zur zehnten Klassenstufe? In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. G. d. Schiefele (Hrsg.), *PISA 2003*. (S. 63-85). Münster, Westfalen u.a.: Waxmann.
- Feuer, M. J., Holland, P. W., Green, B. F., Bertenthal, M. W. & Hemphill, F. C. (Hrsg.). (1999). *Uncommon Measures: Equivalence and Linkage among Educational Tests*. Washington, DC: National Academy Press.
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: Reidel.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: BMBF.
- Kolen, M. J. (2004). Linking Assessments: Concept and History. *Applied Psychological Measurement*, 28(4), 219-226.
- Kolen, M. J. & Brennan, R. L. (2004). *Test Equating, Scaling, and Linking: Methods and Practices (2nd ed.)*. New York: Springer.
- Leuders, T. & Leiß, D. (2006). Realitätsbezüge. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. Berlin: Cornelsen Scriptor.
- Linn, R. L. (1993). Linking results of distinct assessments. *Applied Measurement in Education*, 6(1), 83-102.
- Mislevy, R. J. (1992). *Linking Educational Assessments: Concepts, Issues, Methods, and Prospects*. Princeton, NY: ETS Policy Information Center.
- Neumann, A. (2006). *Vergleich der Konzeptionen und Inhalte der Prüfungen zum Mittleren Schulabschluss in Berlin und der Prüfungen Jahrgangsstufe 10 in Brandenburg Schuljahr 2005/06*: Institut für Schulqualität der Länder Berlin und Brandenburg e. V.
- Niss, M. (2003). *Mathematical Competencies and the Learning of Mathematics: The Danish KOM Project*, Verfügbar unter:
http://www7.nationalacademies.org/mseb/mathematical_competencies_and_the_learning_of_mathematics.pdf [2007-08-14]
- Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61, 37-46.

Anhang: Klassifikation der Aufgaben

Die folgenden Tabellen enthalten die Klassifikationen der einzelnen Aufgaben der Abschlussarbeiten aus Berlin und Brandenburg aus dem Jahr 2007 nach primärer Leitidee, allgemeinen mathematischen Kompetenzen und Anforderungsbereich. Außerdem wurde eine Bewertung vorgenommen, der die Maßstäbe zugrunde liegen, wie sie für Aufgaben gelten, die in den Tests zur länderübergreifenden Normierung der Bildungsstandards eingesetzt werden.

Die Skala reicht von 0 bis 3, wobei die Kategorien ungefähr folgende Bedeutung haben:

0 Aufgabe sollte so auf keinen Fall verwendet werden

1 Aufgabe mit deutlichen Mängeln

2 Akzeptable bis gute Aufgabe

3 Sehr gute Aufgabe.

Berlin

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3		
1a	x									x		x			2	
1b	x									x		x			0	Mischung aus Dezimalbruch und Bruch fragwürdig
1c				x						x			x		2	
1d	x									x			x		2	
2a	x							x	x	x			x		2	
2b	x							x	x	x			x		2	
2c	x					x		x	x	x			x		1	
3		x						x	x	x	x		x		2	
4				x				x		x	x		x		1	
5a				x				x	x				x		0	Der Graph stellt die Sachsituation stark vereinfacht da, ohne dass dies thematisiert wird.
5b				x		x		x	x				x		0	
6a	x							x	x	x			x		1	Die Sinnhaftigkeit der Genauigkeit ist fragwürdig.
6b	x							x	x	x			x		2	

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung	
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3			
7a				x						x	x		x			2	
7b				x						x			x			2	
8		x						x	x	x	x		x			2	
9a				x					x			x				2	
9b				x					x	x			x			2	
9c		x							x	x			x			2	
9d				x			x		x	x			x			2	
10a		x								x		x				2	
10b		x								x		x				2	
10c		x					x			x			x			2	
10d				x		x				x			x			2	

Brandenburg

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung	
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3			
Brandenburg Realschule																	
1a	x									x			x			2	
1b				x						x		x				1	
1c	x						x			x			x			2	
1d			x						x	x		x				1	
1e	x									x			x			2	
1f		x					x			x		x				2	
2a1				x					x	x			x			2	
2a2				x						x			x			2	
2b				x			x			x			x			2	
2c				x			x		x	x			x			2	
2d1				x					x	x			x			2	
2d2		x					x		x	x			x			2	
2e				x						x		x				2	

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3		
2e				x						x		x			2	Der Kontext dieser Aufgabe ist vor allem aufgrund abweichender Angaben verschiedener Informationsquellen zur Begrenzung des Bermudadreiecks oder der Zahl der verschwundenen Objekte nicht eindeutig umrissen. Auf diesem Hintergrund ist eine Genauigkeit wie "etwa 36,73%" nicht akzeptabel.
3a	x							x		x			x		0	
3b1	x							x		x		x			2	
3b2	x							x		x			x		1	Durch die Vorwegnahme wichtiger Vereinfachungen und Annahmen (wie beispielsweise die Vernachlässigung der Erdkrümmung) werden den Schülern wichtige Modellierungsschritte abgenommen. Das Anforderungsniveau der Aufgabe verlagert sich somit auf ein maximal mittleres Niveau.
3c1		x						x	x	x			x			
3c2		x						x	x	x			x		1	Die Aufgabenstellung erscheint künstlich konstruiert, da wichtige Eckdaten der Sachsituation (z.B. die Orthogonalität der Fahrtrichtung zur Blickrichtung) ermittelt, einfacherer festzustellende Daten (z.B. die zurückgelegte Fahrtstrecke) jedoch nicht angegeben sind.
3d		x					x		x	x			x		0	Diese Teilaufgabe steht nicht im Bezug zu dem der Aufgabe zugrundeliegenden Kontext.
3e	x							x		x			x		2	
41a					x				x				x		1	Um die Wahrscheinlichkeit berechnen zu können, mit der der erste Besucher eine Blume bekommt, muss für das "zufälligem Auswählen" die gleiche Wahrscheinlichkeit für alle Überraschungsgeschenke angenommen werden. Aufgrund der unterschiedlichen Form und Größe der einzelnen Geschenke ist dies jedoch eher nicht anzunehmen. Die Teilaufgabe
41b					x			x		x			x		1	kann somit kaum beantwortet werden.

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3		
41c	x							x		x			x			1
41d	x							x		x			x			2
41e					x			x	x			x				2
42a					x			x	x				x			0
42b1				x				x	x				x			0
42b2	x					x		x		x			x			0
42c				x				x		x			x			0
42d	x							x		x			x			0
43a1			x						x				x			1
43a2		x							x	x		x				1
43b		x							x	x			x			1
43c	x							x		x			x			2
43d					x			x	x				x			1

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung	
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3			
43e	x							x	x				x			2	
Brandenburg Gesamtschule Grundkurs																	
1a	x										x	x				2	
1b				x							x		x			2	
2a				x					x	x			x			2	
2b				x						x			x			2	
2c				x						x		x				2	
2d				x					x				x			2	
2e1				x					x	x			x			2	
2e2		x							x	x			x			2	
Brandenburg Gesamtschule Erweiterungskurs																	
4.1a	x							x		x	x		x			1	Die Bedingungen der Rabattkarte (insb. der Kaufpreis) wirken für ein Kaufhaus künstlich.
4.1b					x			x		x			x			1	Um die Wahrscheinlichkeit berechnen zu können, mit der der erste Besucher eine Blume bekommt, muss für das "zufälligen Auswählen" die gleiche Wahrscheinlichkeit für alle Überraschungsgeschenke angenommen werden. Aufgrund der unterschiedlichen Form und Größe der einzelnen Geschenke ist dies jedoch eher nicht anzunehmen. Die Teilaufgabe kann somit kaum beantwortet werden.
4.1c	x							x		x			x			2	
4.1d					x			x	x	x	x		x			2	
4.2a1				x				x	x	x			x			0	
4.2a2	x						x	x		x			x			0	
4.2b				x				x					x			0	
4.2c	x							x		x			x			0	Der zwischen dem Angebot und dem tatsächlichen Kauf liegende Zeitraum erscheint unrealistisch. Des Weiteren ist nicht klar, in welchem Zeitraum die beim Angebot 3 vereinbarten Raten zu zahlen sind.

Aufgabe	Leitidee					Kompetenzen						Anforderungsbereich			Bewertung	Bemerkung
	L1	L2	L3	L4	L5	K1	K2	K3	K4	K5	K6	A1	A2	A3		
4.2d				x				x		x	x			x	0	Die beiden Angebote des Vaters sind zu konstruiert. Eine einmalige Unterstützung oder eine regelmäßige Zahlung in Form eines Taschengeldes wäre in diesem Fall realistischer.
4.3a1			x						x	x				x	1	
4.3a2		x							x	x				x	1	
4.3b		x						x	x	x				x	2	
4.3c1		x						x		x	x			x	2	
4.3c2		x						x		x	x			x	2	
Brandenburg Gymnasium																
2a				x					x	x				x	2	
2b				x						x				x	2	
2c				x			x			x				x	2	
2d				x			x		x	x				x	2	
2e		x					x		x	x				x	2	
2f		x							x	x				x	2	
3a1		x						x	x	x		x			1	
3a2		x						x	x	x				x	1	
3b	x							x		x				x	1	
3c		x					x	x		x				x	2	
3d		x						x	x	x				x	1	
3e		x					x	x	x	x				x	1	

Institut für Schulqualität der Länder Berlin und Brandenburg e. V.

www.isq-bb.de

