



Olaf Köller
Michel Knigge
Bernd Tesch
(Hrsg.)



Sprachliche Kompetenzen im Ländervergleich

WAXMANN

Sprachliche Kompetenzen im Ländervergleich

Olaf Köller
Michel Knigge
Bernd Tesch (Hrsg.)

Sprachliche Kompetenzen im Ländervergleich



Waxmann 2010
Münster/New York/München/Berlin

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-8309-2339-8

© 2010, Waxmann Verlag GmbH,
Postfach 8603, D-48046 Münster

Waxmann Publishing Co.,
P. O. Box 1318, New York, NY 10028, USA.

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Christian Aeverbeck, Münster
Titelfoto: Michael Mieth, Berlin
Satz: Stoddart Satz- und Layoutservice, Münster
Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706

Alle Rechte vorbehalten.
Printed in Germany

Inhalt

Vorwort des Präsidenten der Kultusministerkonferenz.....	9
---	----------

Kapitel 1

Die Bildungsstandards im allgemeinbildenden Schulsystem.....	13
---	-----------

Claudia Pöhlmann, Daniela Neumann, Bernd Tesch und Olaf Köller

1	Bildungsstandards im allgemeinbildenden Schulsystem	13
1.1	Definition von Bildungsstandards	14
1.2	Entwicklung von Testverfahren zur Überprüfung der Bildungsstandards	15
1.3	Standards und Tests in einer Gesamtstrategie zur Qualitätssicherung	16
	Literatur.....	17

Kapitel 2

Die im Ländervergleich untersuchten sprachlichen Kompetenzen.....	19
--	-----------

2.1	Beschreibung der im Ländervergleich im Fach Deutsch untersuchten Kompetenzen.....	19
	Katrin Böhme, Daniela Neumann und Stefan Schipolowski	
2.2	Beschreibung der in der ersten Fremdsprache (Englisch/Französisch) untersuchten Kompetenzen.....	25
	Raphaela Porsch und Bernd Tesch	
	Literatur.....	29

Kapitel 3

Kompetenzstufenmodelle für den Mittleren Schulabschluss	33
--	-----------

3.1	Einleitung.....	33
	Katrin Böhme und Olaf Köller	
3.2	Politische und inhaltliche Rahmenbedingungen bei der Setzung von Kompetenzstufen	35
	Olaf Köller	
3.3	Kompetenzstufenmodelle für das Fach Deutsch.....	37
	Albert Bremerich-Vos, Ulrike Behrens, Katrin Böhme, Michael Krelle, Daniela Neumann, Alexander Robitzsch, Stefan Schipolowski und Olaf Köller	
3.4	Kompetenzstufenmodelle für Lese- und Hörverstehen in der ersten Fremdsprache	50
	Bernd Tesch, Raphaela Porsch, Eynar Leupold, Rüdiger Grotjahn, Karin Kleppin, Jenny Frenzel, Claudia Harsch, Michael Leucht, Hans Anand Pant, Andre A. Rupp, Simon P. Tiffin-Richards und Olaf Köller	
	Literatur.....	63

Kapitel 4

Anlage und Durchführung des Ländervergleichs 65

Katrin Böhme, Michael Leucht, Stefan Schipolowski, Raphaela Porsch,
Michel Knigge und Olaf Köller

4.1	Beteiligte Institutionen und Personen	66
4.2	Testinstrumente, Testablauf und Testdesign	69
4.3	Verknüpfung des Ländervergleichs 2009 mit PISA 2009	74
4.4	Definition der Zielpopulationen und Stichprobenziehung	75
4.5	Aufbereitung und Analyse der Daten	82
4.6	Qualitätsmonitoring im Ländervergleich 2008/2009	83
	Literatur	84

Kapitel 5

Die sprachlichen Kompetenzen in den Ländern 87

5.1	Der Ländervergleich im Fach Deutsch	87
	Stefan Schipolowski und Katrin Böhme	
5.2	Der Ländervergleich im Fach Englisch	97
	Michael Leucht, Jenny Frenzel und Claudia Pöhlmann	
5.3	Zusammenfassung	105
	Michael Leucht und Stefan Schipolowski	
	Literatur	105

Kapitel 6

Der Blick in die Länder 107

Olaf Köller, Michel Knigge, Stefan Schipolowski, Michael Leucht, Simon P. Tiffin-Richards,
Raphaela Porsch, Bernd Tesch, Miriam Vock, Katrin Böhme, Jenny Frenzel und
Jessica Andrade

6.1	Baden-Württemberg	110
6.2	Bayern	116
6.3	Berlin	119
6.4	Brandenburg	124
6.5	Bremen	127
6.6	Hamburg	131
6.7	Hessen	134
6.8	Mecklenburg-Vorpommern	139
6.9	Niedersachsen	143
6.10	Nordrhein-Westfalen	147
6.11	Rheinland-Pfalz	151
6.12	Saarland	157
6.13	Sachsen	161
6.14	Sachsen-Anhalt	165
6.15	Schleswig-Holstein	168
6.16	Thüringen	171
6.17	Fazit	175
	Literatur	175

Kapitel 7

Geschlechterdisparitäten..... 177

Henrik Winkelmann und Imke Groeneveld

7.1 Empirische Befunde zu Geschlechterdifferenzen in sprachlichen Leistungen.....	177
7.2 Ergebnisse.....	178
7.3 Zusammenfassung.....	183
Literatur.....	184

Kapitel 8

Soziale Disparitäten im Spracherwerb 185

Michel Knigge und Michael Leucht

8.1 Einleitung.....	185
8.2 Indikatoren primärer und sekundärer sozialer Disparitäten.....	186
8.3 Überblick.....	187
8.4 Ergebnisse.....	188
8.5 Zusammenfassung und Diskussion.....	198
Literatur.....	200

Kapitel 9

Migrationsbedingte Disparitäten bei sprachlichen Kompetenzen 203

Katrin Böhme, Simon P. Tiffin-Richards, Stefan Schipolowski und Michael Leucht

9.1 Einleitung.....	203
9.2 Erfassung des Migrationshintergrundes: Empirische Befunde und Erklärungsansätze für migrationsbezogene Disparitäten.....	204
9.3 Datengrundlage.....	207
9.4 Kompetenzstände der Jugendlichen mit unterschiedlichem Migrationsstatus.....	207
9.5 Kompetenzstände verschiedener Herkunftsgruppen im Vergleich.....	210
9.6 Leistungen von Schülerinnen und Schülern mit und ohne Migrationshintergrund in den Ländern.....	213
9.7 Diskussion und Zusammenfassung.....	223
Literatur.....	224

Kapitel 10

Effekte der sozialen Zusammensetzung der Schülerschaft 227

Michel Knigge und Olaf Köller

10.1 Einleitung.....	227
10.2 Mehrebenenanalysen zur Bestimmung von Kompositionseffekten.....	228
10.3 Kompositionseffekte in der Gesamtstichprobe.....	230
10.4 Kompositionseffekte in den Ländern.....	232
Literatur.....	244

Dank

Viele haben zum Entstehen dieses Berichts beigetragen. Die Herausgeber möchten an dieser Stelle Christiane Penk ihren besonderen Dank für die Unterstützung bei der Erstellung von zahlreichen Abbildungen und Tabellen für den vorliegenden Bericht aussprechen. Auch bei Beate Plugge und ihren Kolleginnen und Kollegen vom Waxmann Verlag bedanken wir uns für die produktive und freundliche Zusammenarbeit bei der Erstellung dieses Berichts.

Vorwort des Präsidenten der Kultusministerkonferenz

Mit der Verabschiedung von gemeinsamen Bildungsstandards steht den Ländern seit 2004 neben den internationalen Erhebungen PISA, IGLU und TIMSS erstmals auch ein bundesweit geltender Bezugsrahmen für das Bildungsmonitoring zur Verfügung. Die Bildungsstandards beschreiben die fachbezogenen Kompetenzen, die Schülerinnen und Schüler bis zu einem gewissen Abschnitt ihrer Schullaufbahn erreicht haben sollen. Die Standards leisten damit auch einen wichtigen Beitrag, um die Gleichwertigkeit der schulischen Ausbildung und der Schulabschlüsse in den Ländern sicherzustellen.

Alle Länder haben sich verpflichtet, ab dem Schuljahr 2004/2005 beziehungsweise 2005/2006 die von der Kultusministerkonferenz beschlossenen Bildungsstandards zu übernehmen und umzusetzen. Bundesweit geltende Bildungsstandards gibt es für die Fächer

- Deutsch und Mathematik für den Primarbereich (Jahrgangsstufe 4),
- Deutsch, Mathematik und Erste Fremdsprache (Englisch, Französisch) für den Hauptschulabschluss (Jahrgangsstufe 9),
- Deutsch, Mathematik und Erste Fremdsprache (Englisch, Französisch) für den Mittleren Schulabschluss (Jahrgangsstufe 10) sowie
- Biologie, Chemie und Physik für den Mittleren Schulabschluss (Jahrgangsstufe 10).

Standards zu setzen hat aber nur dann einen Sinn, wenn auch konsequent überprüft wird, ob sie erreicht werden. Hierzu ist im Juni 2004 von der Kultusministerkonferenz das von den Ländern gemeinsam getragene „Institut zur Qualitätsentwicklung im Bildungswesen – Wissenschaftliche Einrichtung der Länder an der Humboldt-Universität zu Berlin“ (IQB) gegründet worden. Das IQB entwickelt empirisch validierte Kompetenzstufenmodelle zu den Bildungsstandards. Diese Kompetenzstufenmodelle geben Lehrkräften, Lernenden und Eltern Auskunft darüber, wie groß die Schüleranteile sind, die die Standards erreichen beziehungsweise verfehlen. Man erhält so einen Vergleichsmaßstab für individuelle Schülerleistungen. Mit dem nun vorliegenden vom IQB 2009 durchgeführten Ländervergleich wird erstmalig zentral überprüft, inwieweit in den einzelnen Ländern die in den Bildungsstandards formulierten Kompetenzanforderungen tatsächlich erreicht werden. Der nächste Ländervergleich wird 2011 im Primarbereich durchgeführt werden.

Die zentrale Überprüfung des Erreichens der nationalen Bildungsstandards in Ländervergleichen hat vor allem drei Ziele: Sie soll

- für mehr Transparenz und eine größere Objektivität der Leistungsfeststellung über Ländergrenzen hinweg und für eine höhere Verbindlichkeit im Hinblick auf die gesetzten Anforderungen sorgen,
- die Rechenschaftslegung im Bildungsbereich verbessern und
- Hinweise für konkrete Maßnahmen liefern, die geeignet sind, das Bildungswesen und den Unterricht weiterzuentwickeln.

Die vom IQB durchgeführten Ländervergleiche lösen die bisherigen Ländervergleiche im Rahmen von PISA (PISA-E) und PIRLS/IGLU (IGLU-E) ab. Organisatorisch werden die Ländervergleiche aber weiterhin in Anknüpfung an die internationalen Schulleistungsstudien durchgeführt. So fand der Ländervergleich 2009 vom IQB im zeitlichen Kontext der PISA-2009-Erhebungen statt. Dadurch wird die Anbindung an internationale Standards gewährleistet. Im Gegensatz zu den internationalen Erhebungen orientieren sich die Ländervergleiche zur Überprüfung des Erreichens der KMK-Bildungsstandards jedoch stärker an der Lehrplanwirklichkeit in den deutschen Ländern. Im Gegensatz zur internationalen PISA-Studie werden beim Ländervergleich 2009 nicht nur die 15-jährigen Schülerinnen und Schüler, sondern die gesamte Bandbreite der Jahrgangsstufe 9 betrachtet. Zudem wurden zusätzlich zum Fach Deutsch auch die Fächer Englisch und Französisch getestet.

Die Kultusministerkonferenz hat sich dafür ausgesprochen, dass die Ergebnisse von Ländervergleichen stets *vor* den Ergebnissen der internationalen Schulleistungsstudien veröffentlicht werden, um der Politik und den Schulen eine rasche Rückmeldung geben zu können. Auf diese Weise wird auch deutlich gemacht, dass die Bildungsstandards der Kultusministerkonferenz der entscheidende Referenzrahmen für unsere Schulen sind – gleichwohl aber eine enge Anknüpfung an die internationalen Anforderungen und Entwicklungen gewährleistet wird.

Ich bin zuversichtlich, dass die von der Kultusministerkonferenz beschlossenen Bildungsstandards neben ihrer Überprüfungsfunktion auch die erforderlichen neuen Impulse zur Schul- und Unterrichtsentwicklung setzen werden. Indem wir die vorliegenden empirischen Daten systematisch nutzen, können wir maßgeblich dazu beitragen, den Unterricht weiterzuentwickeln. Vor diesem Hintergrund hat die Kultusministerkonferenz im Dezember 2009 eine „Konzeption zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung“ verabschiedet.

Bildungsstandards zu setzen und ihr Erreichen zu überprüfen, ist aber nur die eine Seite der bildungspolitischen Steuerungsaufgabe. Ebenso bedeutend ist es, die Schülerinnen und Schüler adäquat auf diese Anforderungen vorzubereiten. Hier gilt es vor allem, den Anteil der Schülerinnen und Schüler zu reduzieren, welche die Schule verlassen, ohne die für eine selbstbestimmte Lebensführung und die Aufnahme einer beruflichen Ausbildung oder Tätigkeit notwendigen Kompetenzen aufzuweisen. Deshalb hat die Kultusministerkonferenz in diesem Frühjahr in einer „Förderstrategie für leistungsschwächere Schülerinnen und Schüler“ die Maßnahmen der Länder zur nachhaltigen Förderung dieser Schülerinnen und Schüler zusammenfassend dargestellt und gemeinsame Leitlinien vereinbart, die zu den vorhandenen Maßnahmen Schwerpunkte setzen und Orientierung für eine verbesserte Förderung bieten.

Auf der Grundlage der jetzt vorliegenden Daten gilt es, die länderspezifisch identifizierten Stärken und Schwächen im Rahmen der bildungspolitischen Zuständigkeit der einzelnen Länder anzugehen. Zudem stellt sich in einem föderal organisierten Bildungswesen die Frage nach der gesamtstaatlichen Verantwortung der Ländergemeinschaft: Die Bildungssysteme der deutschen Länder müssen einen gemeinsamen Grundstock von Bildungsinhalten schaffen und die Vergleichbarkeit von Anforderungen in den Kernbereichen zentraler Fächer gewährleisten. Dies ist ein komplexer Entwicklungsprozess. Mit der Verabschiedung von gemeinsamen Bildungsstandards als Grundlage für die fachbezogenen Anforderungen an Schulen sind wir auf dem richtigen Weg. Mit deren Überprüfung in Ländervergleichen setzen wir diesen Weg konsequent fort.

Berlin, im Juni 2010

Staatsminister Dr. Ludwig Spaenle
Präsident der Ständigen Konferenz der Kultusminister der Länder
in der Bundesrepublik Deutschland

Kapitel 1

Die Bildungsstandards im allgemeinbildenden Schulsystem

Claudia Pöhlmann, Daniela Neumann, Bernd Tesch und Olaf Köller

1 Bildungsstandards im allgemeinbildenden Schulsystem

Infolge der Publikationen zu den Ergebnissen der PISA-Studie¹ 2000 (Baumert et al., 2001, 2002) haben sich in Deutschland erhebliche Veränderungen in der Steuerung des allgemeinbildenden Schulsystems ergeben, die häufig als Paradigmenwechsel von der Input- zur Outputsteuerung beschrieben werden (vgl. Klieme et al., 2003). Die erreichten fachlichen Kompetenzniveaus der Schülerinnen und Schüler sollen Rückschlüsse auf die Effizienz des Bildungssystems zulassen. Ausdruck findet diese Outputsteuerung unter anderem in der langfristigen Beteiligung an internationalen *Large-scale Assessments* wie PISA (z. B. Prenzel et al., 2007) oder PIRLS/IGLU² (vgl. z. B. Bos et al., 2007), aber auch in der Verabschiedung von Bildungsstandards, in denen für ausgewählte Fächer definiert ist, über welche Kompetenzen Schülerinnen und Schüler zu festen Zeitpunkten ihrer schulischen Karriere verfügen sollten.

Mit ihren Beschlüssen von 2003 und 2004 hat die Ständige Konferenz der Kultusminister (KMK) für alle 16 Länder der Bundesrepublik Deutschland solche verbindlichen Bildungsstandards für die Grundschule und das Ende der Sekundarstufe I verabschiedet. Mit den KMK-Beschlüssen vom Oktober 2007 ist der Auftrag dahin gehend erweitert worden, auch Standards für das Ende der gymnasialen Oberstufe zu erarbeiten, und zwar erneut für die Fächer Deutsch, Fremdsprachen (Englisch/Französisch), Mathematik und die drei Naturwissenschaften (Biologie, Chemie, Physik). Damit werden in absehbarer Zeit in den so genannten Kernfächern abschlussbezogene Standards für das gesamte allgemeinbildende Schulsystem in Deutschland vorliegen.

An die Verabschiedung der Bildungsstandards hat die KMK in ihren Beschlüssen vom Juni 2006 (KMK, 2006) ein breites System der Qualitätssicherung gekoppelt, das neben stichprobenbasierten internationalen und nationalen Vergleichen auch regelmäßige flächendeckende Vergleichsarbeiten vorsieht. Dieses System dient dazu frühzeitig rückzumelden, ob Schülerinnen und Schüler die in den Standards formulierten Leistungserwartungen einlösen können.

Dieses erste Kapitel soll dazu dienen, die Bildungsstandards der KMK zu beschreiben, die Erstellung von standardbasierten Testinstrumenten zu skizzieren und einen kurzen Überblick über die Gesamtstrategie der KMK zur Qualitätssicherung zu geben. Wir verzichten dabei auf eine Diskussion darüber, inwieweit Standards zur Qualitätsentwicklung im Bildungssystem beitragen können und wie sich Standards bildungstheoretisch verorten lassen. Dies ist andern-

1 PISA: Programme for International Student Assessment.

2 PIRLS: Progress in International Reading Literacy Study; IGLU: Internationale Grundschul-Lese-Untersuchung.

orts hinreichend geschehen (z.B. Köller, 2009). Im Folgenden liegt der Fokus deutlich auf der intendierten Rolle der Standards für die Qualitätssicherung im Bildungssystem.

1.1 Definition von Bildungsstandards

Bildungsstandards sind das Resultat eines Aushandlungsprozesses zwischen Politik, Wissenschaft, Verbänden und Schulpraxis. Sie definieren, welche Qualitätskriterien ein Bildungssystem erfüllen soll. An ihre Definition ist die Erwartung gebunden, dass sie zur Qualitätssicherung und Qualitätsentwicklung in Bildungssystemen beitragen. Hierzu ist es nötig, Standards so zu formulieren, dass sie sich operationalisieren lassen und ihre Erreichung überprüfbar wird.

In der internationalen Diskussion um Bildungsstandards werden drei Formen unterschieden, die sich entweder auf den Unterricht beziehen oder den Blick zu den erreichten Leistungen der Schülerinnen und Schüler wenden (vgl. NCTM, 2000; NRC, 1995). Inhaltliche Standards (*Content Standards*) definieren für ein Schulfach die Leistungen, die in den jeweiligen Stoffgebieten erreicht werden sollen. Leistungsstandards (*Performance Standards*) definieren dagegen prozessbezogene Kompetenzen (kognitive Prozesse) als Ziele schulischen Unterrichts, welche in ganz unterschiedlichen inhaltlichen Gebieten des Faches angewendet werden müssen. Schließlich skizzieren Unterrichtsstandards (*Opportunity-to-learn Standards*) Bedingungen und Merkmale gelingenden Fachunterrichts.

Mit Bezug auf die zunehmende Eigenverantwortung von Schulen und die gleichzeitig notwendige Orientierung an überprüfbaren und vergleichbaren Lernergebnissen (*Outcomes*) (vgl. Scheerens, 2004) haben Klieme et al. (2003) argumentiert, dass Bildungsstandards dann am ehesten steuerungsrelevant werden, wenn sie als Leistungsstandards formuliert werden. Die Autoren verstehen darunter erwünschte fachspezifische Lernergebnisse der Schülerinnen und Schüler zu bestimmten Zeitpunkten ihrer Bildungsbiografien. Im Einklang mit diesen Empfehlungen hat die Ständige Konferenz der Kultusminister der Länder (KMK) mit ihren Beschlüssen 2003 und 2004 Leistungsstandards für das Ende der 4. Jahrgangsstufe in der Primarstufe und das Ende der Sekundarstufe I beschlossen. Berücksichtigt wurden die Fächer Deutsch, Mathematik, die erste Fremdsprache (Englisch/Französisch) und die Naturwissenschaften (Biologie, Chemie, Physik). Im Grundschulbereich fand dabei eine Beschränkung auf Deutsch und Mathematik statt. In der Sekundarstufe I wurde noch zwischen Standards für den Hauptschulabschluss, der typischerweise nach der 9. Jahrgangsstufe erreicht wird, und Standards für den Mittleren Schulabschluss, der üblicherweise nach der 10. Jahrgangsstufe erreicht wird, unterschieden.

Mit den Bildungsstandards wurde die Verbindung der traditionellen Steuerungsinstrumente mit einer Output-Steuerung eingeleitet, und zwar in dem Sinne, dass Bildungsstandards „Anforderungen an das Lehren und Lernen in der Schule [formulieren]. Sie benennen Ziele für die pädagogische Arbeit, ausgedrückt als erwünschte Lernergebnisse der Schülerinnen und Schüler. Damit konkretisieren Standards den Bildungsauftrag, den allgemeinbildende Schulen zu erfüllen haben [...]. Die Bildungsstandards legen fest, welche Kompetenzen die Kinder oder Jugendlichen bis zu einer bestimmten Jahrgangsstufe erworben haben sollen“ (Klieme, 2004, S. 258). Kompetenzen werden als die bei Schülerinnen und Schülern verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten verstanden. Sie umfassen darüber hinaus die damit verbun-

denen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können (vgl. Weinert, 2001). Kompetenzen sind freilich nicht unmittelbar beobachtbare, sondern hypothetische Konstrukte. „Solche Konstrukte können erst mit Hilfe von Messinstrumenten der Beobachtung zugänglich gemacht werden [...]“ (Köller, 2008, S. 165). In diesem Sinne beschreibt der Ländervergleich einen Weg, ausgewählte Kompetenzen, für die eine reliable und valide Messung möglich ist, zu erfassen und Wissen darüber zu generieren, ob die von der KMK in den Standards formulierten Leistungserwartungen von den Schülerinnen und Schülern erreicht werden. Im ersten standardbasierten Ländervergleich 2009 wurden die Kompetenzen der Schülerinnen und Schülern der 9. Jahrgangsstufen in allen Bildungsgängen des allgemeinbildenden Systems in den Fächern Deutsch und Englisch (Durchführung der Untersuchung im Jahr 2009) sowie Französisch (Durchführung der Untersuchung bereits im Jahr 2008) erfasst. Eine genauere Beschreibung der untersuchten Teilkompetenzen enthält Kapitel 2.

1.2 Entwicklung von Testverfahren zur Überprüfung der Bildungsstandards

Damit Bildungsstandards im Rahmen des Ländervergleichs als Instrumente der Qualitätssicherung verwendet werden können, müssen reliable und valide Messinstrumente zur Verfügung gestellt werden. Die Testkonstruktion wurde für die Fächer Deutsch und die erste Fremdsprache vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin koordiniert (vgl. Köller, 2009). In Vorarbeiten des IQB wurde bereits fachdidaktisch und lerntheoretisch präzisiert, was unter den jeweiligen Kompetenzen der einzelnen Fächer zu verstehen ist. Weiterhin konnte für die Testkonstruktion auf Materialien zurückgegriffen werden (in Anlehnung an Rost, 2004), mit denen Aufgabenentwickler trainiert wurden, wie gute Testaufgaben konstruiert werden können beziehungsweise welche Fehler bei der Itementwicklung vermieden werden sollten. Die Aufgabenentwicklung erfolgte auf Grundlage dieser Arbeiten in den im Folgenden beschriebenen fünf Schritten für die Fächer Deutsch, Englisch und Französisch:

1. *Initiierung und Aufrechterhaltung der Itementwicklung durch erfahrene Lehrkräfte:* Lehrkräfte aus allen 16 Bundesländern wurden zunächst geschult und entwickelten dann in regionalen Arbeitsgruppen unter fachdidaktischer Anleitung Testaufgaben für die Fächer Deutsch, Englisch und Französisch. Die entwickelten Aufgaben wurden an dieser Stelle durch die Lehrkräfte selbst bei Schülerinnen und Schülern erprobt (Präpilotierungen), um authentische Schülerlösungen für die Weiterentwicklung der Items zu erhalten. Im Fach Französisch konnten neben diesen Items zusätzlich 360 Items eingesetzt werden, die vom Centre International d'Etudes Pédagogiques (CIEP) in Sèvres konstruiert worden waren.
2. *Begutachtung der Items:* Expertinnen und Experten aus Bildungsforschung, Psychometrie und Fachdidaktik beurteilten und kommentierten in einem zweiten Schritt die Items hinsichtlich ihrer psychometrischen und fachdidaktischen Güte. Diese Bewertungen wurden für die weitere Optimierung der Aufgaben genutzt.
3. *Empirische Erprobung der entwickelten Aufgaben (Pilotierung):* Alle Items wurden dann in großen Schülerstichproben erprobt. Die Stichprobengröße

wurde dabei so gewählt, dass hinreichend robuste Schätzungen der Itemparameter mit Modellen der *Item-Response-Theory* möglich waren. Auf der Basis entsprechender Auswertungen wurden Items mit problematischen Kennwerten identifiziert, überarbeitet oder aus dem Itempool entfernt. Damit stand als Ergebnis dieses Schrittes ein erprobter beziehungsweise überarbeiteter Aufgabenpool zur Verfügung. Gleichzeitig wurden die Angaben der Schülerinnen und Schüler für die Erstellung und Optimierung der Auswertungsanleitungen und Kodiermanuale genutzt.

4. *Normierung der Aufgaben:* In diesem Schritt wurden die pilotierten Items einer bundesweit repräsentativen Stichprobe von Schülerinnen und Schülern der 9. und 10. Jahrgangsstufe vorgelegt und anhand von Kennwerten der klassischen und probabilistischen Testtheorie weitere problematische Items aus dem Itempool entfernt. Auf der Basis dieser Stichproben wurden die Items kalibriert und anschließend nationale Skalen für die verschiedenen Kompetenzen definiert.
5. *Definition von Kompetenzstufen:* In konsensuellen, so genannten *Standard-setting*-Verfahren erarbeiten Mitarbeiterinnen und Mitarbeiter der Fachdidaktik, Erziehungswissenschaft, Schulverwaltung sowie Lehrkräfte durch die Festlegung von Trennlinien Niveaustufen auf den Skalen. Auf diesen kontinuierlichen Skalen markieren die Trennlinien bezogen auf die empirisch ermittelten Schwierigkeitswerte von Aufgaben die Übergänge zwischen Mindest-, Regel- und Maximalstandards. Für die erste Fremdsprache waren die Schwellen bezogen auf die Niveaus des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GER; Europarat, 2001) für den Übergang von A1 zu A2, von A2 zu B1, von B1 zu B2 und von B2 zu C1 festzulegen. Die Kompetenzstufenmodelle mit ihren konkreten Grenzen und inhaltlichen Anforderungen werden für Deutsch und die erste Fremdsprache in Kapitel 3 dargestellt. Mit Hilfe dieser Kompetenzstufenmodelle ist es einerseits möglich, Testergebnisse Niveaus zuzuordnen und zu vergleichen, andererseits können den Stufen auch Testaufgaben zugeordnet werden. So werden Aussagen darüber möglich sein, welches konkrete Aufgabenmaterial von Schülerinnen und Schülern auf einer bestimmten Stufe erfolgreich bearbeitet werden sollte.

1.3 Standards und Tests in einer Gesamtstrategie zur Qualitätssicherung

Am Ende der gerade beschriebenen fünf Schritte steht ein umfassender und kontinuierlich ergänzter Aufgabenpool, der für standardorientierte Testungen verwendet werden kann, wie sie die KMK (2006) in ihren „Plöner Beschlüssen“ zur Gesamtstrategie zur Qualitätssicherung im allgemeinbildenden Schulsystem beschlossen hat. Auf der Basis von landesweit repräsentativen Stichproben soll demnach in allen Ländern Deutschlands festgestellt werden, welche Anteile der Schülerpopulation die Standards erreichen. Der in Plön beschlossene Zeitplan und die Verzahnung internationaler und nationaler Schulleistungstudien sind in Tabelle 1.1 wiedergegeben.

Tabelle 1.1: Internationale Vergleiche und Ländervergleiche in den nächsten Jahren als Folge der Plöner Beschlüsse der KMK

Jahr	Stufe	Internationale Studie	Fächer im Ländervergleich
2009	Sekundarstufe I	PISA	Deutsch, erste Fremdsprache
2011	Primarstufe	IGLU/TIMSS	Deutsch, Mathematik
2012	Sekundarstufe I	PISA	Mathematik, Biologie, Chemie, Physik
2015	Sekundarstufe I	PISA	Deutsch, erste Fremdsprache
2016	Primarstufe	IGLU	Deutsch, Mathematik
2018	Sekundarstufe I	PISA	Mathematik, Biologie, Chemie, Physik

PISA: Programme for International Student Assessment; IGLU: Internationale Grundschul-Lese-Untersuchung; TIMSS; Trends in Mathematics and Science Study

Im Grundschulbereich soll die Überprüfung der Bildungsstandards im Fünf-Jahres-Rhythmus in der 4. Jahrgangsstufe geschehen, und zwar zeitlich gekoppelt an die Internationale Grundschul-Lese-Untersuchung (IGLU/PIRLS; vgl. Bos et al., 2003, 2007), welche ebenfalls einem Fünf-Jahres-Rhythmus folgt. Neben diesem nationalen Monitoring beteiligt sich Deutschland weiterhin an den internationalen Studien (PIRLS, TIMSS). Letztmalig in IGLU 2006 wurde der Ländervergleich auf der Basis internationaler Instrumente durchgeführt, erstmalig 2011 mit der nächsten PIRLS- und TIMSS-Erhebung wird es auf der Basis der Bildungsstandards zu einem Ländervergleich in den Fächern Deutsch und Mathematik kommen. Dies wiederholt sich dann passend zum IGLU-Rhythmus alle 5 Jahre.

Im Bereich der Sekundarstufe I sehen die Plöner Beschlüsse vor, dass letztmalig 2006 ein nationaler Vergleich zwischen den Ländern auf der Basis der PISA-Instrumente erfolgte. Seit dem Jahr 2009 wird der nationale Vergleich zwischen den Ländern auf Grundlage der Bildungsstandards durchgeführt, wobei die Erhebung zeitgleich mit PISA erfolgt und die Studien auch aneinander angebundnen sind. Im Jahr 2009 wurde der Test für Deutsch und die erste Fremdsprache (Englisch, Französisch bereits 2008) durchgeführt, 2012 folgt der Vergleich in Mathematik und den Naturwissenschaften, 2015 folgen wieder die Sprachen, 2018 Mathematik und die Naturwissenschaften. Die Ländervergleiche in der Sekundarstufe I passen sich damit einem sechsjährigen Rhythmus an, getrennt für die Sprachen und Mathematik und die Naturwissenschaften. Berücksichtigt werden dabei 9. Jahrgangsstufen. Die Ergebnisse, die in diesem Berichtsband dargestellt werden, beziehen sich auf die Untersuchungen im Ländervergleich im Jahr 2009 für die Sprachen Deutsch, Englisch und Französisch.

Literatur

- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, J. & Weiß, M. (Hrsg.). (2002). *PISA 2000. Die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, J. & Weiß, M. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.

- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K., Valtin, R. (2007). *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Europarat (2001). *Gemeinsamer Europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Klieme, E. (2004). Zur Entwicklung nationaler Bildungsstandards. Grundpositionen einer Expertise. In T. Fitzner (Hrsg.), *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform* (S. 256–265). Bad Boll: Evangelische Akademie.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E., Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards: Eine Expertise*. Berlin: Bundesministerium für Bildung und Forschung.
- KMK (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Luchterhand.
- Köller, O. (2008). Bildungsstandards – Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. *Zeitschrift für Pädagogik*, 54, 163–173.
- Köller, O. (2009). Bildungsstandards. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (2. Aufl., S. 529–548). Wiesbaden: VS Verlag für Sozialwissenschaften.
- National Council of Teachers of Mathematics (NCTM) (2000). *Professional standards for school mathematics*. Reston, VA: NCTM.
- National Research Council (NRC) (Hrsg.). (1995). *National science education standards*. Washington, DC: NRC.
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (Hrsg.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.
- Rost, J. (2004). *Lehrbuch Testtheorie/Testkonstruktion* (2. Aufl.). Bern: Huber.
- Scheerens, J. (2004). The formative implications of standards. In T. Fitzner (Hrsg.), *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform* (S. 202–219). Bad Boll: Evangelische Akademie.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim: Beltz.

Kapitel 2

Die im Ländervergleich untersuchten sprachlichen Kompetenzen

Die KMK hat mit ihren Plöner Beschlüssen vom 2.6.2006 (314. Sitzung der KMK) festgelegt, dass der Ländervergleich 2009 in der Sekundarstufe I auf der Grundlage der Bildungsstandards in den Fächern Deutsch und der ersten Fremdsprache (Englisch beziehungsweise Französisch) erfolgen soll. Im Rahmen der 324. KMK-Sitzung vom Dezember 2008 wurde weiterhin beschlossen, den ersten Ländervergleich zur Erreichung der Bildungsstandards auf die Kompetenzbereiche Leseverstehen und Hörverstehen in der ersten Fremdsprache sowie Lesen, Hören und Orthografie im Fach Deutsch zu beschränken.

2.1 Beschreibung der im Ländervergleich im Fach Deutsch untersuchten Kompetenzen

Katrin Böhme, Daniela Neumann und Stefan Schipolowski

Im Rahmen dieses Kapitels werden die im Fach Deutsch getesteten Kompetenzen genauer beschrieben. Hierbei werden zunächst die Bildungsstandards für den Mittleren Schulabschluss im Fach Deutsch vorgestellt. Nachfolgend wird thematisiert, wie die Struktur der sprachlichen Kompetenzen im Deutschen konzipiert werden kann. Abschließend wird geklärt, welche der in den Standards beschriebenen Teilbereiche in Testaufgaben überführt und einer Überprüfung zugänglich gemacht werden konnten. Hierbei wird vertiefend auf die verwendeten Testinstrumente eingegangen.

2.1.1 Die Bildungsstandards für den Mittleren Schulabschluss (MSA) im Fach Deutsch

In der Verkehrssprache mündlich wie schriftlich zu kommunizieren sowie Texte und andere Medienbeiträge reflektierend zu verstehen und zu nutzen sind Kernkompetenzen, die für die Teilhabe am gesellschaftlichen Leben maßgeblich sind und insbesondere im Laufe der Schulzeit erworben und vertieft werden sollen. Die Bildungsstandards für den MSA im Fach Deutsch gehen von einer vorschulisch erworbenen Sprachhandlungskompetenz aus, die im Primar- und Sekundarbereich systematisch erweitert und ausgebaut wird. Sie benennen verschiedene Aspekte sprachlicher Kompetenzen, über die Schülerinnen und Schüler zum Ende der Sekundarstufe I verfügen sollen.

In der fachdidaktischen und sprachwissenschaftlichen Literatur werden sprachliche Kompetenzen häufig in verschiedene Teilbereiche, zumeist in die Subkompetenzen Lesen, Hören, Schreiben und Sprechen unterteilt (vgl. Felder, 2003). Innerhalb dieser Kompetenzbereiche können zum einen anhand der

Sprachmodalität mündliche Kompetenzen (Sprechen und (Zu-)Hören) von schriftsprachlichen Kompetenzen (Schreiben und Lesen) unterschieden werden. Zum anderen lassen sich anhand der beteiligten sprachlichen Prozesse produktive Kompetenzen (Sprechen und Schreiben) von rezeptiven Kompetenzen (Hören und Lesen) differenzieren.

In den Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss sind die folgenden vier Kompetenzbereiche ausgewiesen (KMK, 2004a):

1. *Sprechen und Zuhören*: zu anderen, mit anderen, vor anderen sprechen sowie Hörverstehen entwickeln
2. *Schreiben*: reflektierend, kommunikativ und gestalterisch schreiben; richtig schreiben¹
3. *Lesen – mit Texten und Medien umgehen*: lesen, Texte und Medien verstehen und nutzen, Kenntnisse über Literatur erwerben
4. *Sprache und Sprachgebrauch untersuchen*: Sprache zur Verständigung gebrauchen, fachliche Kenntnisse erwerben, über Verwendung von Sprache nachdenken und sie als System verstehen

Drei der vier Kompetenzbereiche können entsprechend Tabelle 2.1 den oben genannten Sprachmodalitäten zugeordnet werden:

Tabelle 2.1: Strukturierung sprachlicher Kompetenzen

	Mündlich	Schriftlich
Sprachproduktion	Sprechen	Schreiben/Rechtschreibung
Sprachrezeption	Zuhören	Lesen

Offen bleibt in dieser Zuordnung die Frage, wie der Kompetenzbereich „Sprache und Sprachgebrauch untersuchen“ verortet werden kann. Dieser hat sowohl rezeptive als auch produktive Aspekte und ist in beiden Modalitäten, also mündlich und schriftlich, relevant. Daher erscheint die in den Bildungsstandards vorgenommene Charakterisierung dieses Bereichs als übergreifende Kompetenz (vgl. KMK, 2004a, S. 7f.) mit Stützfunktion für alle übrigen Kompetenzen (Sprechen, Zuhören, Schreiben und Lesen) als sinnvolle Lösung.

2.1.2 Die Struktur sprachlicher Kompetenzen im Deutschen

Für die Überprüfung der (verkehrs-)sprachlichen Kompetenzen der Schülerinnen und Schüler im Rahmen des Ländervergleichs ist die Frage zentral, ob ein übergreifender Faktor sprachlicher Fähigkeiten existiert, auf den sämtliche sprachliche Leistungen wie Lesen und Schreiben zurückgeführt werden können (vgl. Oller, 1976), oder ob diese verschiedenen sprachlichen Leistungen als ein mehrdimensionales Konstrukt verstanden werden sollten, bei dem mehrere weitgehend voneinander unabhängige Teilbereiche zu unterscheiden sind. Je nachdem, ob von einer allgemeinen sprachlichen Kompetenz oder von relativ unabhängigen Teilkompetenzen ausgegangen wird, verändern sich Aufbau und Inhalt des eingesetzten Tests (vgl. hierzu Jude, 2008).

¹ Auf den besonderen Status der orthografischen Teilkompetenz wird später näher eingegangen.

Aus theoretischer Sicht scheint es naheliegend, insbesondere zwischen den Kompetenzbereichen Zuhören und Lesen beträchtliche Ähnlichkeiten zu vermuten, da es sich in beiden Fällen um rezeptive Fähigkeiten handelt. Den Fokus empirischer Studien zur Diagnostik rezeptiver sprachlicher Kompetenzen bildet die Erforschung des Leseverstehens. Es liegen gesicherte empirische Erkenntnisse zu verschiedenen individualdiagnostischen Verfahren (vgl. bspw. den HAMLET 3-4 für den Primarbereich: Lehmann, Peek & Poerschke, 2006) sowie für das *Large-scale Assessment* vor (vgl. bspw. IGLU im Primarbereich: Bos et al., 2007; PISA in der Sekundarstufe I: Deutsches PISA-Konsortium, 2001). Die Operationalisierung des verkehrssprachlichen Hörverstehens ist im deutschsprachigen Raum hingegen wenig untersucht (vgl. Behrens, Böhme & Krelle, 2009).

Zu der Frage, ob es sich beim verkehrssprachlichen Hör- und Leseverstehen um verschiedene Prozesse handelt, gibt es in der Literatur unterschiedliche Standpunkte. Da die kognitiven Verarbeitungsprozesse beim Lese- und Hörverstehen sehr ähnlich sind (vgl. bspw. Kürschner & Schnotz, 2008), wird mitunter angenommen, dass die Lese- und die Hörverstehenskompetenz zu einer rezeptiven Sprachkompetenz zusammengefasst werden können (Rindermann, 2006). Auch empirisch wurde ein enger Zusammenhang zwischen verkehrssprachlichem Hör- und Leseverstehen wiederholt untermauert (vgl. bspw. Rost & Hartmann, 1992; Marx & Jungmann, 2000). Andererseits werden bedeutsame Differenzen zwischen gesprochener und geschriebener (und damit zwischen gehörter und gelesener) Sprache anerkannt (vgl. Grotjahn & Tesch, 2010), was zu der gegenteiligen Annahme führt, dass von zwei verschiedenen Subdimensionen auszugehen ist (Buck, 1991, 2001). Ein wichtiges Unterscheidungsmerkmal bei der Abgrenzung von Hör- und Leseverstehen ist die Flüchtigkeit gesprochener Sprache. Gesprochenes muss beim einmaligen Hören in Echtzeit verarbeitet werden. Im Gegensatz dazu besteht beim Leseverstehen die Möglichkeit zum wiederholten Überprüfen verschiedener Informationen. Andererseits enthält die gesprochene Sprache deutlich mehr Redundanzen und Möglichkeiten, das Verständnis der Gesprächsbeteiligten zu bestätigen.

In der jüngeren kognitionspsychologischen Forschung wird die Ansicht vertreten, dass es sich beim Hör- und Leseverstehen weder um völlig identische, noch um gänzlich verschiedene Prozesse handelt (vgl. bspw. Kürschner & Schnotz, 2008). Die kognitiven Verarbeitungsprozesse beim Hör- und Leseverstehen können demnach weder als ausschließlich modalitätsunabhängig, noch als ausschließlich modalitätsspezifisch gelten. Kürschner und Schnotz (2008) betonen, dass hierarchiehöhere Informationsverarbeitungsprozesse, wie die Konstruktion einer mentalen Repräsentation, beim Hör- und Leseverstehen sehr ähnlich ablaufen und Unterschiede insbesondere auf hierarchieniedrigeren Ebenen des kognitiven Systems vorliegen. Diese Unterschiede werden auf die modalitätsspezifischen Besonderheiten der akustischen beziehungsweise visuellen Information zurückgeführt.

Empirische Analysen zur Struktur der in den Bildungsstandards für das Fach Deutsch insgesamt beschriebenen Kompetenzen legten Bremerich-Vos, Böhme und Robitzsch (2009) für den Primarbereich vor. Dort zeigte sich in einer Hauptkomponentenanalyse der Teilkompetenzen Zuhören, Schreiben, Lesen, Sprachgebrauch, Orthografie sowie einer Skala zur Prüfung der verbalen Intelligenz ein dominanter erster Faktor, der als Globalfaktor sprachlicher Kompetenz interpretiert werden kann und der mit 66,4 Prozent einen erheblichen Teil der Varianz aufklärt. Ergänzend wurde eine explorative Faktorenanalyse mit

zwei Faktoren durchgeführt, die sich an der theoretischen Unterscheidung von produktiven und rezeptiven Sprachkompetenzen orientiert. Die Befunde ergaben plausible Anordnungen der verschiedenen Kompetenzbereiche entsprechend der Unterscheidung von produktiven und rezeptiven Fähigkeiten, wobei keine vollständige Trennung der entsprechenden Kompetenzen erzielt werden konnte. Erwartungskonform zeigte der produktive Faktor hohe Ladungen von Schreiben, Rechtschreibung und zusätzlich verbaler Intelligenz. Für den rezeptiven Faktor erzielte der Bereich Zuhören eine sehr hohe Ladung. Die Kompetenzbereiche Lesen und Sprachgebrauch luden auf beiden Faktoren und wiesen somit Merkmale beider Dimensionen auf. Insgesamt kann dieses Befundmuster so verstanden werden, dass das Hörverstehen eine sehr spezielle Teilfähigkeit innerhalb der Deutschkompetenz insgesamt darstellt und dass die durchgängige Trennung in rezeptive und produktive Fähigkeiten nicht überzeugend gelingt.

Folgt man diesen Befunden und der theoretischen Diskussion in der Deutschdidaktik, so liegt es nahe, die in den Bildungsstandards für das Fach Deutsch benannten Kompetenzbereiche nicht als eine einzige homogene Fähigkeit zu behandeln, sondern die Kompetenzen als jeweils separate Fähigkeitskonstrukte zu verstehen. Aus diesem Grund wird im Fach Deutsch auf die Ermittlung eines Gesamtscores im Sinne einer allgemeinen Deutschkompetenz verzichtet und es werden in den nachfolgenden Kapiteln stets gesonderte Ergebnisse für die einzelnen Kompetenzbereiche (Zuhören, Lesen sowie Orthografie) berichtet.

2.1.3 Konstruktbestimmung und Operationalisierung der im Ländervergleich überprüften sprachlichen Kompetenzen im Fach Deutsch

Die theoretischen Grundlagen der Testkonstruktion für die beiden rezeptiven Sprachkompetenzen sind unter anderem im Rahmen der großen Schulleistungsstudien der letzten Jahre, wie etwa PISA, intensiv aufgearbeitet worden (vgl. z.B. Beck & Klieme, 2007; Buck, 2001; Prenzel et al., 2007; Kintsch, 1998).

Vor diesem Hintergrund werden Zuhören und Lesen als kommunikative Fähigkeiten verstanden, die mit Hilfe von authentischem, altersangemessenem Stimulusmaterial in Testaufgaben überführt werden können. Die Lesekompetenz beschreibt die Fähigkeit, geschriebenes Textmaterial zu verstehen. Die Zuhörkompetenz bezieht sich auf die Fähigkeit, auditiv dargebotenes Stimulusmaterial wie Hörtexte, Radiobeiträge und ähnliches zu verstehen. Zuhören und Lesen erfordern Rezeptionsleistungen verschiedener Komplexität, die vom Detailverstehen bis hin zum Globalverstehen und Bewerten von Texten reichen.

Im Ländervergleich 2009 wurden Teile der Kompetenzbereiche „Sprechen und Zuhören“, „Lesen – mit Texten und Medien umgehen“ und „Schreiben“ getestet, und zwar die Facetten Zuhören, Lesen sowie Orthografie. Nicht alle der in den Bildungsstandards für diese Bereiche aufgeführten Teilkompetenzen lassen sich derzeit für das *Large-scale Assessment* operationalisieren. Somit sind Aussagen über die Kompetenzstände der Schülerinnen und Schüler auch nur für die getesteten Aspekte dieser Teilbereiche möglich. Eine Verallgemeinerung der ermittelten Kompetenzstände auf andere Kompetenzbereiche oder auf eine globale „Deutschkompetenz“ ist nicht sinnvoll (vgl. Abschnitt 2.1.2). Daher wird sich auch die Darstellung der empirischen Befunde im Rahmen dieses Berichts

stets auf die getesteten Kompetenzbereiche beziehungsweise Teilkompetenzen beschränken (vgl. Kapitel 5 bis 10 in diesem Band).

Die im Kompetenzbereich „Sprechen und Zuhören“ überprüfbaren Teilkompetenzen beziehen sich überwiegend auf das verstehende Zuhören, bei dem Gesprächsbeiträge anderer verfolgt und rezipiert werden sollen, wesentliche Informationen verstanden, gesichert und wiedergegeben sowie auch non-verbale Äußerungen (zum Beispiel Stimmführung) verstanden werden sollen (KMK, 2004a). Wesentlicher Grund hierfür ist, dass die standardisierte Leistungsmessung im Kompetenzbereich „Sprechen“ in Deutschland noch in ihren Anfängen steckt und Testungen in diesem Bereich bislang nur Pilotcharakter haben. Eine Überprüfung der Sprechkompetenz würde in Einzel- beziehungsweise Kleingruppentestungen erfolgen müssen, wobei die Sprechakte entweder direkt von Beobachtern eingeschätzt und bewertet oder auf Video aufgezeichnet und nachfolgend kodiert werden müssten. Diese Form der Leistungsmessung würde allerdings eine moderne technische Ausstattung in den Schulen voraussetzen beziehungsweise äußerst aufwändige logistische Erfordernisse mit sich bringen, die zum Zeitpunkt des Ländervergleichs 2009 nicht durchgängig erwartet werden konnten.

Auch im Kompetenzbereich „Lesen“ sind einige der in den Bildungsstandards aufgeführten Teilkompetenzen bislang noch nicht im *Large-scale Assessment* überprüfbar. Die *de facto* eingesetzten Testaufgaben beziehen sich daher überwiegend auf die Teilkompetenzen „Literarische Texte verstehen und nutzen“ sowie „Sach- und Gebrauchstexte verstehen und nutzen“ (KMK, 2004a, S. 14f.). Dabei ist insbesondere die Fähigkeit relevant, auch längere sowie nichtlineare Texte zu verstehen und im Detail zu erfassen. Informationen sollen zielgerichtet entnommen, geordnet sowie hinsichtlich Aussagekraft und Wertung geprüft werden können. Weiterhin sollen die Schülerinnen und Schüler diese Informationen interpretieren, Schlussfolgerungen daraus ziehen und die Intention eines Textes erkennen können.

Bei der Aufgabenentwicklung für die Kompetenzbereiche „Zuhören“ und „Lesen“ wurde darauf geachtet, dass die den Aufgaben zugrunde liegenden Stimuli zu ungefähr gleichen Anteilen auf literarische und nicht-literarische sowie kontinuierliche und nicht-kontinuierliche Texte verteilt sind. Des Weiteren wurden verschiedene Itemformate eingesetzt, die nach dem Grad des von den Schülerinnen und Schülern eigenständig zu leistenden produktiven Gehalts in geschlossene, halboffene und offene Formate eingeteilt werden. Geschlossene Items bieten Antwortmöglichkeiten an, die angekreuzt oder unterstrichen werden müssen. Konkret wurden im Ländervergleich als geschlossene Itemformate Multiple-Choice-Items (MC-Items), Richtig-Falsch-Items sowie Um- und Zuordnungsaufgaben verwendet. MC-Items umfassen zumeist eine Frage oder einen zu vervollständigenden Satzanfang und vier Antwortalternativen, von denen genau eine Option korrekt ist. Bei (halb-)offenen Antwortformaten werden den Schülerinnen und Schülern keine Antwortoptionen angeboten. Sie verlangen eine eigenständig formulierte Antwort, die aus einem Wort oder wenigen Wörtern (halboffen) bis hin zu mehreren Sätzen (offen) bestehen kann.

Die Operationalisierung der orthografischen Kompetenz im Rahmen von großen nationalen und internationalen Schulleistungsstudien hat weit weniger Tradition als die Überprüfung des Hör- und Leseverstehens. Auch ist die empirische Durchdringung dieses Kompetenzbereichs – insbesondere für den Sekundarstufenbereich – weit weniger fortgeschritten. Die deutschsprachige und internationale Fachliteratur bezieht sich vorrangig auf die Entwicklung der or-

thografischen Kompetenz im Vor- und Grundschulalter. Auch aktuellere individualdiagnostische Instrumente zur Erfassung der Rechtschreibkompetenz Jugendlicher und junger Erwachsener sind rar und berücksichtigen zum Teil noch nicht die in der Rechtschreibreform neu fixierten Notierungskonventionen („neue Rechtschreibung“; vgl. Kersting & Althoff, 2004; s. aber auch: Bulheller, Ibrahimovic & Häcker, 2005). In Hinblick auf den Entwicklungsstand der orthografischen Kompetenz der Schülerinnen und Schüler der 9. Jahrgangsstufe gibt es nur wenige empirische Befunde. Es besteht allerdings Anlass zu der Vermutung, dass alphabetische Elemente mit großer Sicherheit und auch orthografische Aspekte in der überwiegenden Mehrzahl der Fälle beherrscht werden. Im Zentrum der Diagnostik sollten daher anspruchsvolle orthografische Phänomene wie etwa die Groß- und Kleinschreibung, Getrennt- und Zusammenschreibung sowie die Zeichensetzung stehen (vgl. Kersting & Althoff, 2004).

Bei der Überführung der Bildungsstandards in Testaufgaben wurde der Bereich Orthografie separat operationalisiert, wiewohl die orthografische Kompetenz in den Bildungsstandards für den MSA in erster Linie als ein Teilbereich des Kompetenzbereichs Schreiben („richtig schreiben“) aufgeführt ist und dort wie folgt spezifiziert wird:

- „Grundregeln der Rechtschreibung und Zeichensetzung sicher beherrschen und häufig vorkommende Wörter, Fachbegriffe und Fremdwörter richtig schreiben,
- individuelle Fehlerschwerpunkte erkennen und mit Hilfe von Rechtschreibstrategien abbauen, insbesondere Nachschlagen, Ableiten, Wortverwandtschaften suchen, grammatisches Wissen anwenden.“ (KMK, 2004a, S. 11)

Ferner findet sich unter den Erläuterungen der Teilkompetenz „Texte überarbeiten“ im Kompetenzbereich „Schreiben“ die Formulierung „Strategien zur Überprüfung der sprachlichen Richtigkeit und Rechtschreibung anwenden“ (KMK, 2004a, S. 13). Zusätzlich wird unter den hier aufgeführten Methoden und Arbeitstechniken der Aspekt „Einhaltung orthografischer und grammatischer Normen kontrollieren“ benannt, welcher der orthografischen Kompetenz zugeordnet werden kann (KMK, 2004a, S. 13). Auch innerhalb des Kompetenzbereichs „Sprache und Sprachgebrauch untersuchen“ wird eine Teilfähigkeit aufgeführt, die sich der orthografischen Kompetenz zuordnen lässt: „wichtige Regeln der Aussprache und der Orthografie kennen und beim Sprachhandeln berücksichtigen“ (KMK, 2004a, S. 16). Unter den zugehörigen Methoden und Arbeitstechniken finden sich zusätzlich die Punkte „Rechtschreibstrategien anwenden: zum Beispiel Ableitung vom Wortstamm, Wortverlängerung, Ähnlichkeitsschreibung“ sowie „Nachschlagewerke nutzen“ (KMK, 2004a, S. 17).

Die Entscheidung, die orthografische Kompetenz als eigenständigen Bereich zu behandeln und unabhängig von der Schreibkompetenz zu operationalisieren, wurde auch aus messtheoretischen Erwägungen getroffen, um die hier relevanten Teilkompetenzen durch den Einsatz spezifischer Aufgabenformate so präzise und passgenau wie möglich erfassen zu können und eine Konfundierung mit anderen Fähigkeitsaspekten, insbesondere des Kompetenzbereichs „Schreiben“, zu vermeiden.

Die Entscheidung der Kultusministerkonferenz, im Rahmen des Ländervergleichs im Fach Deutsch neben dem Hör- und Leseverstehen zusätzlich auch die orthografische Kompetenz der Schülerinnen und Schüler zu testen, wurde zu einem Zeitpunkt gefällt, zu dem bereits Details des Testdesigns und der Testheftzusammenstellung in Planung waren. Um die Durchführung des Länder-