



Bildungsergebnisse und Qualitätsentwicklung

- I 1 Ländervergleich zur Überprüfung der Bildungsstandards für den mittleren Schulabschluss in den Fächern Deutsch, Englisch und Französisch
- I 2 Fremdevaluation
 - I 2.1 Fremdevaluation an allgemein bildenden Schulen
 - I 2.1.1 Grundlagen
 - I 2.1.2 Ergebnisse
 - I 2.1.3 Vertiefende Analysen
 - I 2.2 Fremdevaluation an beruflichen Schulen
 - I 2.2.1 Grundlagen
 - I 2.2.2 Ergebnisse
 - I 2.2.3 Vertiefende Analysen
- I 3 Diagnose- und Vergleichsarbeiten
 - I 3.1 Diagnosearbeiten an Grundschulen (VERA 3)
 - I 3.2 Vergleichsarbeiten an weiterführenden Schulen (DVA)

I Bildungsergebnisse und Qualitätsentwicklung

Die in Baden-Württemberg eingeführten Maßnahmen zur Qualitätssicherung von Schule und Unterricht zielen entsprechend der im Juni 2006 von der Kultusministerkonferenz (KMK) beschlossenen *Gesamtstrategie zum Bildungsmonitoring*¹ darauf ab, Leistungsergebnisse des Bildungssystems systematisch und wissenschaftlich abgesichert darzustellen. Der KMK-Maßnahmenkatalog umfasst vier konzeptionell miteinander verbundene Bereiche: Die Teilnahme an internationalen Schulleistungsuntersuchungen, die zentrale Überprüfung des Erreichens der Bildungsstandards im Ländervergleich, Vergleichsarbeiten zur landesweiten Überprüfung der Leistungsfähigkeit einzelner Schulen

innerhalb der Länder sowie die gemeinsame Bildungsberichterstattung von Bund und Ländern. Darüber hinaus hat Baden-Württemberg – wie auch andere Bundesländer – ein System der schulischen Qualitätssicherung implementiert, in dem Selbstevaluation, Fremdevaluation und mit der Schulaufsicht zu treffende Zielvereinbarungen als verpflichtende Elemente der Schul- und Unterrichtsentwicklung eine zentrale Rolle einnehmen.

Umsetzungswege und Ergebnisse dieser Maßnahmen sind Inhalte der folgenden Kapitel: **I 1** stellt die Resultate des Ländervergleichs zur Überprüfung der Bildungsstandards 2009 für Baden-Württemberg dar. Die zentralen Ergebnisse der verpflichtenden Fremdevaluation sind Thema von Kapitel **I 2**. Im Abschnitt **I 3** werden wichtige Erkenntnisse der Diagnosearbeiten an Grundschulen und der Vergleichsarbeiten an weiterführenden Schulen berichtet.

1 Vgl. <http://www.kmk.org/bildung-schule/qualitaets-sicherung-in-schulen/bildungsmonitoring/ueberblick-gesamtstrategie-zum-bildungsmonitoring.html>; Zugriff am 16.02.2011

I 1 Ländervergleich zur Überprüfung der Bildungsstandards für den mittleren Schulabschluss in den Fächern Deutsch, Englisch und Französisch

In den Jahren 2003 und 2004 hat die Kultusministerkonferenz für alle 16 Bundesländer verbindliche Bildungsstandards für die Grundschule sowie für das Ende der Sekundarstufe I verabschiedet. Bildungsstandards sind als Leistungsstandards zu verstehen, die Ziele der pädagogischen Arbeit in den Schulen definieren und darüber hinaus festlegen, welche Kompetenzen Schüler bis zu einer bestimmten Jahrgangsstufe erworben haben sollen. Sie formulieren Basisqualifikationen, die für die weitere schulische und berufliche Ausbildung von Bedeutung sind und anschlussfähiges Lernen ermöglichen. Bildungsstandards liegen derzeit für den Primarbereich (Jahrgangsstufe 4, in den Fächern Deutsch und Mathematik), für den Hauptschulabschluss (Jahrgangsstufe 9, in den Fächern Deutsch, Mathematik sowie Englisch oder Französisch als erste Fremdsprachen) und für den Mittleren Schulabschluss (Jahrgangsstufe 10, in den Fächern Deutsch, Mathematik, erste Fremdsprache Englisch/ Französisch, Biologie, Chemie und Physik) vor. Bildungsstandards für die gymnasiale Oberstufe (Fächer Deutsch, Mathematik, Englisch, Französisch, Biologie, Chemie und Physik) befinden sich derzeit in Erarbeitung.

An die Beschlüsse zur Einführung von Bildungsstandards sind weitreichende Maßnahmen der Qualitätssicherung gekoppelt. Gemäß der KMK-Gesamtstrategie zum Bildungsmonitoring wird das Erreichen der Standards innerhalb der Länder durch Vergleichsarbeiten und länderübergreifend mittels stichprobenbasierter Leistungsvergleiche überprüft. Ziel der Testungen ist es, frühzeitig zu erkennen, in welchem Umfang die in den Bildungsstandards formulierten Leistungserwartungen erfüllt werden. Durch die Einführung abschlussbezogener Bildungsstandards kann die schulische Qualitätsentwicklung aller Bundesländer zum ersten Mal an einem gemeinsam vereinbarten Maßstab ausgerichtet werden.

Baden-württembergische Bildungspläne integrieren KMK-Standards

Die am 21. Januar 2004 veröffentlichten baden-württembergischen Bildungspläne für die Hauptschule, die Realschule und das Gymnasium gehen in vielen Fällen über die KMK-Standards hinaus. In einigen Fällen bedurften sie jedoch der Konkretisierung anhand von

Begrifflichkeiten, Inhalten und Themen, die durch die KMK-Standards explizit ausgewiesen werden. Ende 2007 übermittelte daher das Kultusministerium den betreffenden Schularten *fachspezifische Hinweise zur Umsetzung der Bildungspläne unter Beachtung der KMK-Standards*, die zu einer besseren Einbindung der KMK-Standards in das System der baden-württembergischen Bildungsstandards beitragen.

Damit vollzog sich ein Wechsel von einer vorrangigen Inputsteuerung des Schulsystems hin zu einer Outputsteuerung. Während frühere Lehrpläne in erster Linie die zu unterrichtenden Inhalte auswiesen, schreiben die neuen Bildungspläne vor, welche fachlichen, personalen, sozialen und methodischen Kompetenzen Kinder und Jugendliche bis zu einem bestimmten Zeitpunkt erwerben müssen. Bei der Umsetzung erhalten die Schulen durch die Reduzierung von Inhalten größere Freiräume. Sie erstellen ein spezifisches Schulcurriculum, für das circa ein Drittel der Unterrichtszeit zur Verfügung steht.

Leistungstests ermöglichen das Messen der erworbenen fachlichen Kompetenzen

Die Überprüfung der sprachlichen Kompetenzen im Ländervergleich im Jahr 2009 basiert auf den KMK-Bildungsstandards für den mittleren Schulabschluss in den Fächern Deutsch, Englisch und Französisch. Verantwortlich für die Durchführung der Testung ist das Institut zur Qualitätsentwicklung im Bildungswesen (IQB).

Mit den Erhebungen zum Ländervergleich waren zwei Anliegen verbunden: Zum einen sollten die Tests länderspezifisch darüber Auskunft geben, in welchem Umfang Schüler, die den mittleren Schulabschluss anstreben, bereits ein Jahr vor Abschluss der Sekundarstufe I die Ziele der Bildungsstandards für die Fächer Deutsch und die erste Fremdsprache erreichen. Zum anderen sollte es die Erhebung im Jahr 2009 für die Fächer Deutsch und Englisch gestatten, alle 16 Länder der Bundesrepublik Deutschland hinsichtlich der in der 9. Jahrgangsstufe des allgemein bildenden Schulsystems erzielten Leistungen zu vergleichen.²

In Deutsch wurden Teile der Kompetenzbereiche »Sprechen und Zuhören«, »Lesen – mit Texten und Medien umgehen« sowie »Schreiben« unter den Aspekten *Zuhören, Lesen sowie Orthografie* getestet.

2 Vgl. zu den Ausführungen in diesem Kapitel Köller, O.; Knigge, M.; Tesch, B. (Hrsg.) (2010): Sprachliche Kompetenzen im Ländervergleich; Münster.

Nicht alle in den Bildungsstandards aufgeführten Teilkompetenzen können untersucht werden (zum Beispiel der Kompetenzbereich »Sprechen«) – eine Verallgemeinerung dieser Teilkompetenzen zu einer »Deutschkompetenz« ist daher nicht möglich.

Grundlage für die Testung in den ersten Fremdsprachen Englisch und Französisch ist die Anbindung der betreffenden Bildungsstandards an den Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER) des Europarats. Im Bereich der Sprachen stellt der GER eine gemeinsame Basis für die explizite Beschreibung von Zielen, Inhalten und Methoden zur Verfügung. Insgesamt werden sechs Niveaustufen unterschieden, welche die elementare Sprachverwendung (A1 und A2), selbstständige Sprachverwendung (B1 und B2) sowie die kompetente Sprachverwendung (C1 und C2) umfassen.³

Im GER werden kommunikative Kompetenzen – bestehend aus kommunikativen Fertigkeiten und den zur Verfügung stehenden sprachlichen Mitteln – sowie interkulturelle und methodische Kompetenzen unterschieden. Die Testung im Rahmen des Ländervergleichs beschränkt sich auf die kommunikativen Fertigkeiten und die darin enthaltenen Teilbereiche *Hör- und Leseverstehen*.

Kompetenzstufenmodelle für den Mittleren Schulabschluss

Die für die Testung eingesetzten Aufgaben wurden entsprechend ihres Schwierigkeitsgrads auf einer linearen Skala fünf Kompetenzstufen zugeordnet. Durch die bei der Normierung der Aufgaben und ihrer fachdidaktischen Analyse gewonnenen Daten konnten auf dieser Skala Grenzwerte eingetragen werden, die das jeweilige Kompetenzspektrum in jeweils fünf Kompetenzstufen unterteilen. Teilaufgaben (Items) mit vergleichbarem Anspruchsniveau liegen somit in einer gemeinsamen Kompetenzstufe.

Die Anwendung komplexer Testverfahren erlaubt es, dass auf derselben Skala, auf der die Items ihrer Schwierigkeit entsprechend angeordnet sind, auch die jeweilige mittlere Leistungsfähigkeit der getesteten Schüler eingeordnet und damit einer bestimmten Kompetenzstufe zugerechnet werden. Neben der Zuordnung zu einer der fünf Kompetenzstufen erfolgt eine alle Kompetenzbereiche umfassende Klassifizierung in Mindest-

3 Die höchste Niveaustufe C2 entspricht einer nahezu muttersprachlichen Sprachbeherrschung und blieb bei der Testung unberücksichtigt, da davon auszugehen war, dass nur sehr wenige Schüler dieses Niveau erreichen.

I 1 (T1)

Zusammenhang von Niveauanforderungen der Bildungsstandards, Kompetenzstufen und GER-Niveaustufen

Klassifizierung nach Niveauanforderungen	Deutsch	Erste Fremdsprache Englisch/ Französisch
	(Zuhören, Lesen, Orthografie; Kompetenzstufenmodell des IOB)	(Lese- und Hörverstehen nach GER-Niveaustufe)
Verfehlung von Mindeststandards	Kompetenzstufe I	Niveau A2.1 (A2-) und niedriger
Erreichung von Mindeststandards	Kompetenzstufe II	Niveau A2.2 (A2+) und B1.1 (B1-)
Erreichung von Regelstandards	Kompetenzstufe III	Niveau B1.2 (B1+)
Erreichung der Regelstandards plus	Kompetenzstufe IV	Niveau B2.1 (B2-)
Erreichung von Maximalstandards	Kompetenzstufe V	Niveau B2.2 (B2+) und höher

Datenquelle: IOB Ländervergleich, 2010

Regel- und Maximalstandards gemäß der KMK-Bildungsstandards (Tabelle I 1 (T1)).

Mindeststandards beziehen sich auf ein definiertes Minimum an Kompetenzen, das alle Schüler zu einem vorher festgelegten Zeitpunkt in ihrer Schullaufbahn erreicht haben müssen. Ein Unterschreiten des definierten Minimalniveaus am Ende der Schullaufbahn würde mit erheblichen Schwierigkeiten dieser Schüler beim Übergang ins Berufsleben einhergehen. *Regelstandards* beschreiben Kompetenzen, die in der Regel von den Schülern einer Jahrgangsstufe erreicht werden sollen. Am Ende der Schullaufbahn würden Regelstandards entsprechend das Ausmaß an Kompetenz und Wissen kennzeichnen, über das zum Beispiel ein durchschnittlicher Zehntklässler verfügen sollte. Als höchste Niveaustufe der Bildungsstandards werden *Maximalstandards* definiert. Die im oberen Leistungsniveau angesiedelten Kompetenzen beziehen sich darauf, was die besten Schüler der jeweiligen Jahrgangsstufen können sollten.

In Baden-Württemberg wurden im Rahmen des Ländervergleichs 115 830 Schüler der 9. Jahrgangsstufe getestet. Jeweils knapp 31 % davon besuchten Hauptschulen oder Gymnasien, 38 % Realschulen und 0,5 % Schulen besonderer Art.

Die Testskalen des Ländervergleichs in den Fächern Deutsch und Englisch sind so konzipiert, dass – analog zu den PISA-Studien – der gesamtdeutsche Mittelwert für alle Neuntklässler im allgemein bildenden Schulsystem in den Kompetenzbereichen jeweils 500 Punkte mit einer Standardabweichung von 100 Punkten beträgt. Lediglich für den Kompetenzbereich Lesen wurde ein Mittelwert von 496 Punkten zugrunde gelegt, um das Kompetenzstufenmodell für diesen Bereich möglichst präzise an die Befunde der PISA-Studie des Jahres 2000 anlehnen zu können.

Fach Deutsch: Leistungen der Neuntklässler in Baden-Württemberg liegen in allen Kompetenzbereichen über dem Bundesdurchschnitt

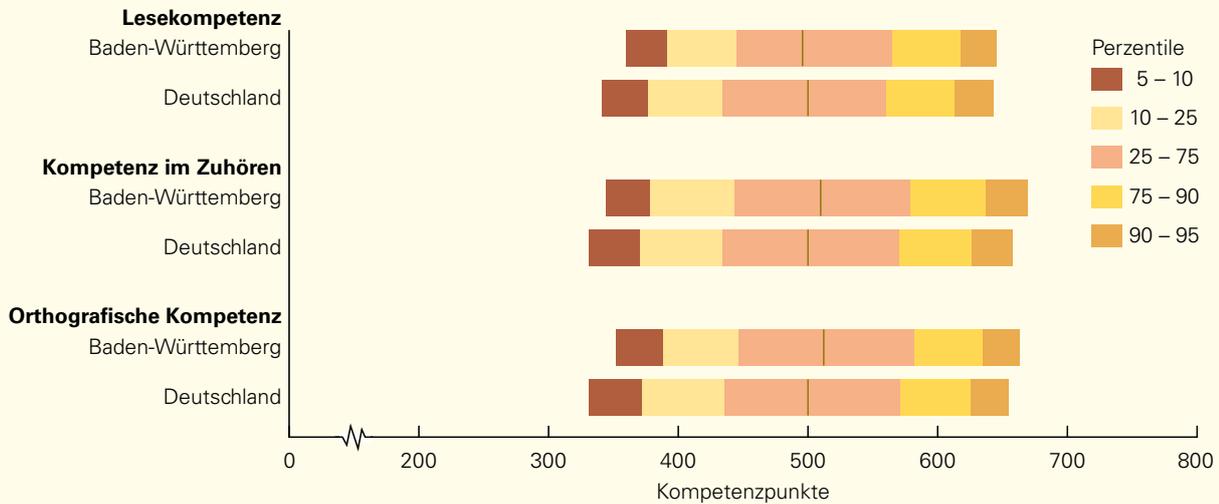
Die durchschnittlichen Testleistungen der baden-württembergischen Neuntklässler aller allgemein bildenden Schularten liegen in den untersuchten Kompetenzbereichen des Fachs Deutsch signifikant über dem für Gesamtdeutschland ermittelten Wert (Tabelle I 1 (T2) im Anhang). Für den Bereich *Lesekompetenz* ergibt sich ein mittlerer Leistungswert von 504 Punkten (Deutschlandwert 496 Punkte), für den Bereich *Zuhören* 510 Punkte und für die *orthografische Kompetenz* 512 Punkte (Deutschlandwerte jeweils 500 Punkte). Die Mittelwerte liefern allerdings nur einen relativ groben Überblick über die Testleistungen der Gesamtschülerschaft. Für eine vertiefende Betrachtung erforderlich ist eine Analyse der Leistungsstreuung und Leistungsbandbreite der jeweils leistungsschwächsten und -stärksten Schüleranteile. Ist der Abstand zwischen den Leistungsstärksten und den leistungsschwächsten besonders groß, deutet dies auf eine ausgeprägte Ungleichheit des Leistungsniveaus hin.

Die Verteilung der Schüleranteile mit unterschiedlicher Leistungsstärke auf der Kompetenzskala erfolgt in Form von Perzentilbändern (Grafik I 1 (G1)). Ein Perzentilwert ist ein Prozentwert auf der definierten Punkteskala, der einen bestimmten Anteil der Untersuchungsgruppe abschneidet. So finden sich zum Beispiel unterhalb des Perzentilwerts 5 die 5 % leistungsschwächsten Schüler, zwischen den Perzentilwerten 25 und 75 die Hälfte der untersuchten Schülerschaft und oberhalb des Perzentilwerts 95 die 5 % Leistungsstärksten.

Die Gruppe der leistungsschwächsten untersuchten Neuntklässler aus Baden-Württemberg weist in allen Teilkompetenzbereichen des Fachs Deutsch einen höheren Wert auf als der Bundesdurchschnitt. Der Wert für

I 1 (G1)

Leistungen der Schüler allgemein bildender Schulen der Jahrgangsstufe 9 in Baden-Württemberg und Deutschland nach Kompetenzbereichen



Kompetenzbereich	Mittelwert ¹⁾	Standardabweichung ¹⁾	Perzentilwert						Spannweite 95 – 5
			5	10	25	75	90	95	
Deutsch – Lesekompetenz									
Baden-Württemberg	504	87	360	391	445	565	617	645	285
Deutschland	496	92	341	376	434	560	613	643	301
Deutsch – Kompetenz im Zuhören									
Baden-Württemberg	510	99	344	378	443	579	637	669	325
Deutschland	500	100	331	370	434	570	626	658	327
Deutsch – orthografische Kompetenz									
Baden-Württemberg	512	96	352	388	446	582	634	663	311
Deutschland	500	100	331	372	435	571	625	655	325

1) Vgl. erläuternde Hinweise am Ende des Kapitels.
Die Perzentilwerte geben an, wieviel Prozent der Schüler unter dem jeweiligen Punktwert liegen; zum Beispiel: 5 % der Neuntklässler aus Baden-Württemberg erreichen unter 360 Punkte im Bereich Lesekompetenz, 95 % erzielen darüber liegende Werte, davon 5 % Werte zwischen 360 und 391 Punkten etc.
Fettdruck = statistisch signifikante Abweichung vom Mittelwert für Deutschland.

Datenquelle: IQB Ländervergleich, 2010

Statistisches Landesamt Baden-Württemberg

434 11

die *Lesekompetenz* liegt um 19 Punkte höher, der Wert für die Kompetenz *Zuhören* 13 Punkte und der Wert für die *orthografische Kompetenz* um 21 Punkte über dem gemittelten Wert für Deutschland. Ebenfalls gilt dies für die Gruppe der Leistungsstärksten, die auf der Kompetenzskala – insbesondere in den Bereichen

Zuhören und Orthografie – höhere Werte als der Bundesdurchschnitt erreichen.

Für die ersten beiden Schuljahre der Sekundarstufe I kann von einem Kompetenzzuwachs im Bereich Lesen von rund 35 Punkten pro Schuljahr ausgegangen wer-

den, für die beiden Jahrgangsstufen 7 und 8 von lediglich 25 Punkten insgesamt. Gegen Ende der Sekundarstufe I sinken die jährlichen Kompetenzzuwächse deutlich ab.⁴ Die festgestellte Punktedifferenz in der Lesekompetenz zwischen den leistungsschwächsten Schülern aus Baden-Württemberg und den leistungsschwächsten im Bundesdurchschnitt entspricht somit in etwa dem Punktezuwachs, der zwischen Mitte und Ende der Sekundarstufe I erzielt wird.

Die Spannweite der Perzentilbänder ist für die Schüler aus Baden-Württemberg im Vergleich zum gesamtdeutschen Wert etwas kleiner (Grafik I 1 (G1)). Die Leistungsunterschiede zwischen den leistungsstärksten und der leistungsschwächsten Schülergruppe sind somit geringer ausgeprägt, das Leistungsniveau homogener bei insgesamt höherem Niveau.

Fach Englisch: Signifikant bessere Leistungen nur im Kompetenzbereich Hörverstehen

Die Neuntklässler in Baden-Württemberg erzielen im Fach Englisch bessere Testwerte als der Bundesdurchschnitt (Tabelle I 1 (T2) im Anhang). Für die Teilkompetenz *Leseverstehen* beträgt der Wert für die Testleistung 507 Punkte, für den Bereich *Hörverstehen* 511 Punkte (Werte für Deutschland jeweils 500 Punkte). Im statistischen Sinne signifikant⁵ ist jedoch nur der Unterschied für den Kompetenzbereich *Hörverstehen*.

Wie im Fach Deutsch weisen auch in den Teilkompetenzbereichen Englisch die leistungsschwächsten Schüler aus Baden-Württemberg im Vergleich zum Bundeswert bessere Testleistungen auf (Grafik I 1 (G2)). Der Abstand im Bereich *Leseverstehen* beträgt 22 Punkte, im *Hörverstehen* 24 Punkte. Die Leistungsspitze in Baden-Württemberg liegt im *Leseverstehen* allerdings 6 Punkte unterhalb des Bundesdurchschnitts der Gruppe der Leistungsstärksten und im *Hörverstehen* mit 2 Punkten nur leicht darüber. Die Leistungsspannweite zwischen der stärksten und schwächsten Gruppe ist wiederum deutlich kleiner als im gesamtdeutschen Vergleich. Die Neuntklässler in Baden-Württemberg sind in Bezug auf ihre Leistungen in den getesteten Bereichen weniger heterogen.

Kein Bundesvergleich bei Französisch als erster Fremdsprache

Das Fach Französisch als erste Fremdsprache wird nur in sechs Bundesländern – darunter Baden-Württemberg – von einer ausreichenden Anzahl von Schülern

gewählt, um repräsentative Auswertungen erstellen zu können. Zudem unterscheidet sich die Zusammensetzung dieser Schülerschaft entsprechend ihrer sozialen Herkunft in anderen Bundesländern teilweise deutlich von derjenigen mit erster Fremdsprache Englisch. Ein Durchschnittswert für das Bundesgebiet wurde in der Studie zum Ländervergleich deshalb nicht berechnet, ebenfalls wurde auf das Ausweisen von Perzentilrängen verzichtet.

Die festgestellte Leistungsstreuung ist relativ groß, in den beiden getesteten Kompetenzbereichen Lese- und Hörverstehen allerdings wenig unterschiedlich (Grafik I 1 (G3)). Rund 27 % der Schüler erreichen in Jahrgangsstufe 9 das Niveau A2.2/B1.1. Im Bereich *Leseverstehen* entsprechen die Testleistungen jeweils etwa eines Fünftels der Schüler den Niveaus B1.2, B2.1 sowie B2.2 und höher. Beim Kompetenzbereich *Hörverstehen* fällt die Leistungsspitze mit knapp 19 % etwas kleiner aus, die Niveaustufen B1.2 und B2.1 mit 27 % bzw. 25 % etwas größer. Ein mit 5 % bzw. 2 % geringer Teil bleibt mit seinen Leistungen unterhalb der Stufe A2.2.

Mehrzahl erreicht oder übertrifft bereits in der 9. Jahrgangsstufe die Regelstandards für Jahrgangsstufe 10

Für die Ermittlung der Kompetenzstufen in Bezug auf das Erreichen der KMK-Bildungsstandards Sekundarstufe I wurden nur die Schülergruppen betrachtet, die mindestens einen mittleren Bildungsabschluss anstreben. Unter diesen Schülern erreicht bzw. übertrifft eine große Mehrheit bereits die für das Ende der Klasse 10 gesetzten Regelstandards des Fachs Deutsch. Für die Bildungsziele im Kompetenzbereich *Zuhören* ist dies bei 90 % der Fall, im Kompetenzbereich *Lesen* bei 73 % und für den Bereich *Orthografie* bei 87 % (Tabelle I 1 (T3) im Anhang). Innerhalb der drei getesteten Kompetenzbereiche nimmt die *Lesekompetenz* eine Sonderrolle ein. Einerseits ist der Anteil derer, die hier die Regelstandards noch nicht erreichen, mit knapp 27 % deutlich größer als bei den Bereichen *Zuhören* (10 %) und *Orthografie* (13 %). Andererseits erreichen mit 34 % deutlich weniger die über dem Regelstandard liegenden Kompetenzstufen IV und V. In den Bereichen *Zuhören* und *Orthografie* erreichen 57 % bzw. 51 % die beiden höchsten Stufen.

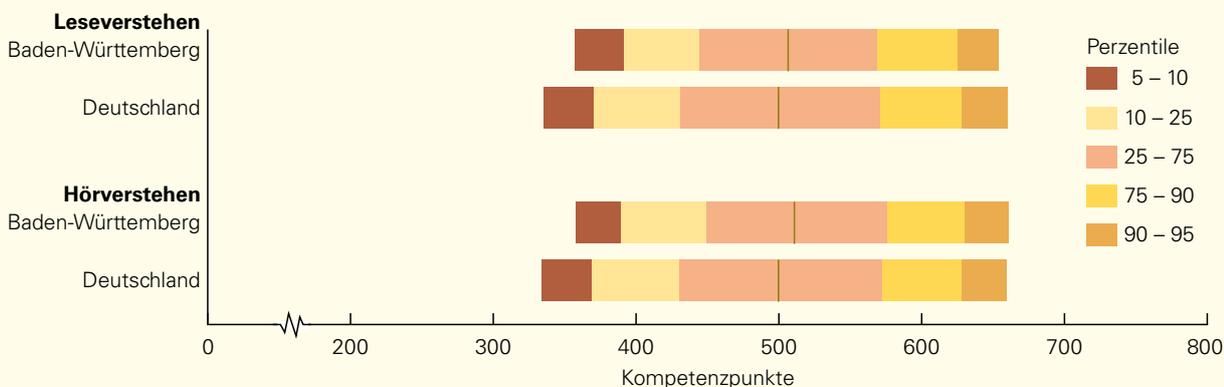
Im Fach Englisch als erste Fremdsprache, Kompetenzbereich *Hörverstehen*, erreichen bzw. übertreffen 46 % der mindestens einen mittleren Abschluss anstrebbenden Schüler die Regelstandards der 10. Jahrgangsstufe bereits im Schuljahr zuvor. Im Kompetenzbereich *Leseverstehen* sind es 41 % (Tabelle I 1 (T4) im Anhang).

4 Vgl. Köller, O.; Knigge, M.; Tesch, B. (Hrsg.) (2010), S. 93.

5 Siehe erläuternde Hinweise am Ende des Kapitels.

I1 (G 2)

Leistungen der Schüler allgemein bildender Schulen der Jahrgangsstufe 9 in Baden-Württemberg und Deutschland im Fach Englisch als erster Fremdsprache 2009 nach Kompetenzbereichen



Kompetenzbereich	Mittelwert ¹⁾	Standardabweichung ¹⁾	Perzentilwert						Spannweite 95 – 5
			5	10	25	75	90	95	
Deutsch – Leseverstehen									
Baden-Württemberg	507	90	357	391	444	569	625	645	297
Deutschland	500	100	335	370	431	571	628	660	325
Deutsch – Hörverstehen									
Baden-Württemberg	511	92	358	389	449	449	630	661	304
Deutschland	500	100	334	369	430	572	628	659	325

1) Vgl. erläuternde Hinweise am Ende des Kapitels.
Die Perzentilwerte geben an, wieviel Prozent der Schüler unter dem jeweiligen Punktwert liegen; zum Beispiel: 5 % der Neuntklässler aus Baden-Württemberg erreichen unter 360 Punkte im Bereich Lesekompetenz, 95 % erzielen darüber liegende Werte, davon 5 % Werte zwischen 360 und 391 Punkten etc.
Fettdruck = statistisch signifikante Abweichung vom Mittelwert für Deutschland.

Datenquelle: IQB Ländervergleich, 2010

Statistisches Landesamt Baden-Württemberg

433 11

Neuntklässler mit Mindestbildungsziel Mittlere Reife, die Französisch als erste Fremdsprache belegten, erreichen bzw. übertreffen zu 70 % vorzeitig die erst zum Ende der Sekundarstufe I geltenden Regelstandards im Bereich *Hörverstehen*. Im Kompetenzbereich *Leseverstehen* sind es 67 % (Tabelle I 1 (T4) im Anhang).

Unterschiedliche Kompetenzzuwächse in Jahrgangsstufe 10 je nach Fach zu erwarten

Die im Rahmen des Ländervergleichs durchgeführte Normierungsstudie kam zu dem Ergebnis, dass im Verlauf der 10. Jahrgangsstufe für das Fach Deutsch lediglich ein mittlerer Kompetenzgewinn von 10 Punkten gegenüber dem Ende der 9. Jahrgangsstufe

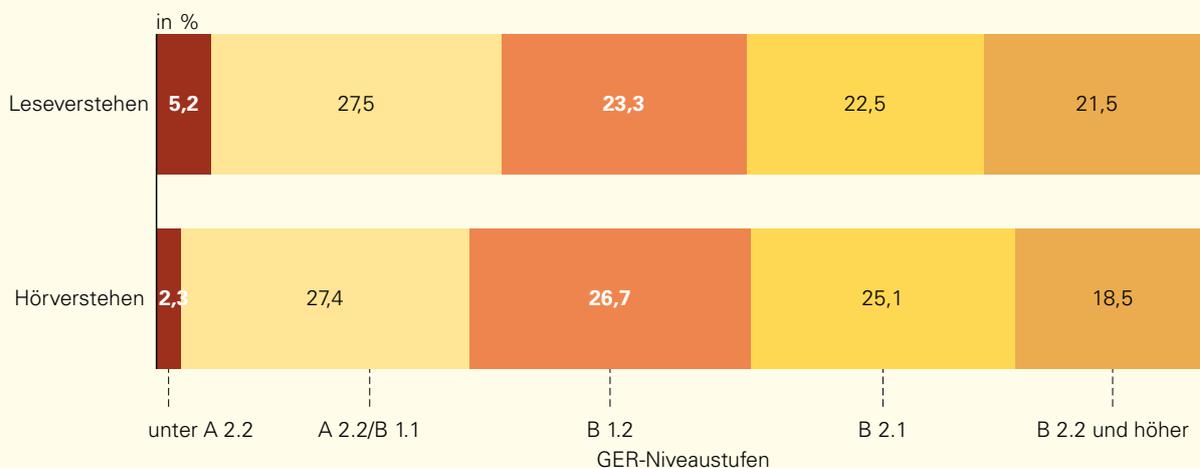
erworben wird.⁶ Die Verteilung der Schüleranteile auf die einzelnen Kompetenzstufen zum Ende der 10. Jahrgangsstufe wird sich aufgrund dieses relativ geringen Zuwachses nur wenig gegenüber der für die Jahrgangsstufe 9 festgestellten Werte verändern. Somit kann abgeschätzt werden, welcher Anteil der einen mittleren Abschluss anstrebenden Schüler die Regelstandards für diese Abschlussart zum Ende der Sekundarstufe I nicht erreichen oder aber übertreffen wird.

Unter der Annahme eines geringfügigen Kompetenzgewinns ist abschätzbar, dass im Bereich *Zuhören* weniger

6 Vgl. Köller, O.; Knigge, M.; Tesch, B. (Hrsg.) (2010), S. 111.

I 1 (G3)

Leistungen der Schüler allgemein bildender Schulen der Jahrgangsstufe 9 in Baden-Württemberg im Fach Französisch als erster Fremdsprache 2009 nach GER-Niveaustufen



435 11

als 9 % der Schüler lediglich nur die Mindeststandards erreichen werden, im Bereich *Lesen* rund 20 % und im Bereich *Orthografie* weniger als 11 %. Nur ein sehr geringer Teil wird in diesen Kompetenzbereichen die Mindeststandards auch am Ende der 10. Jahrgangsstufe nicht erfüllen. Im Kompetenzbereich *Lesen* wird dieser Anteil allerdings größer sein und knapp 5 % erreichen. Ein erheblicher Teil der Schüler wird die Regelstandards übertreffen. Im Kompetenzbereich *Zuhören* ist von einem Anteil über 56 % auszugehen (bei den Gymnasien über 80 %), im Bereich *Lesen* von über 34 % (bei den Gymnasien über 58 %) und im Bereich *Orthografie* mehr als 52 % (bei den Gymnasien über 84 %).

Trotz eines für das Fach Englisch zu erwartenden deutlichen Kompetenzzuwachses von 40 Punkten im Verlauf der Jahrgangsstufe 10 ist abschätzbar, dass rund ein Drittel der einen mittleren Abschluss anstrebenden Schüler nicht die Regelstandards, sondern lediglich die Mindeststandards erreichen wird. Weitere 2 % für den Bereich *Hörverstehen* bzw. 6 % für den Bereich *Leseverstehen* werden hinter den Mindeststandards zurückbleiben. Andererseits wird über ein Drittel aller Schüler der Jahrgangsstufe 10 – bei den Gymnasien rund zwei Drittel – die KMK-Regelstandards übertreffen.

Für das Fach Französisch wird ein Kompetenzzuwachs von 47 Punkten im letzten Schuljahr der Sekundarstufe I angenommen. Somit ist zu erwarten, dass prognostizierte 14 % lediglich die Mindeststan-

dards erreichen und nur ein sehr geringer Teil die für den mittleren Abschluss geltenden Mindeststandards verfehlt. Über 60 % werden die Regelstandards übertreffen, wobei es kaum Unterschiede zwischen den Gymnasien und den sonstigen Schularten (im Fach Französisch sind dies die Realschulen) geben wird.

Erwartungsgemäß liegt in allen Fächern und allen getesteten Teilkompetenzen der Anteil der Gymnasiasten, der die Regelstandards erreicht bzw. übertrifft, deutlich über dem Wert der sonstigen Schularten (Tabellen I 1 (T3, T4) im Anhang).

Mädchen erzielen bessere Leistungen in den getesteten Fächern

Zahlreiche Schulleistungsstudien belegen fachspezifisch unterschiedliche Leistungen von Jungen und Mädchen. Während Jungen im mathematischen Bereich besser abschneiden, erzielen die Mädchen im sprachlichen Bereich signifikant höhere Kompetenzwerte (vgl. Kapitel I 3). Auch der Ländervergleich der sprachlichen Kompetenzen bestätigt diese kontinuierlich feststellbaren Leistungsunterschiede.

Die größte Leistungsdifferenz findet sich im Kompetenzbereich *Orthografie* im Fach Deutsch. Hier beträgt der Abstand in Baden-Württemberg knapp 52 Punkte zugunsten der Mädchen. Weniger deutlich, jedoch

ebenfalls statistisch bedeutsam, fällt der Leistungsvorsprung mit 26 bzw. 22 Punkten in den Bereichen *Zuhören* und *Lesen* aus. In den Fremdsprachen sind die Unterschiede geringer. Der Vorsprung der Schülerinnen des Landes beträgt im Bereich *Leseverstehen* – sowohl in Englisch wie in Französisch – rund 17 Punkte. Der Abstand beim *Hörverstehen* ist in Englisch mit 19 Punkten etwas größer, in Französisch mit knapp 12 Punkten geringer.

Die generell besseren Leistungen der Mädchen zeigen sich auch gegenüber dem gemittelten Wert der geschlechtsspezifischen Leistungsunterschiede bezogen auf Deutschland insgesamt (Tabelle I 1 (T5) im Anhang). Dabei ist auffallend, dass in allen getesteten Kompetenzbereichen der Fächer Deutsch und Englisch der Vorsprung der Mädchen aus Baden-Württemberg über dem bundesdeutschen Durchschnitt liegt – wie bereits ausgeführt allerdings bei insgesamt stärkeren Leistungen beider Geschlechter. Im Fach Französisch liegt die Leistungsdifferenz zwischen den Geschlechtern geringfügig unterhalb des gesamtdeutschen Werts.

Einfluss der sozialen Herkunft liegt in Baden-Württemberg auf dem Niveau des Bundesdurchschnitts

Um den Einfluss sozialer Disparitäten auf die Testleistungen zu untersuchen, wurden in der Studie zum Ländervergleich der sprachlichen Kompetenzen drei Indikatoren herangezogen. Der »soziale Gradient« (1), die Differenz von mittleren Leistungswerten unterschiedlicher sozialer Schichten (2) sowie die relative Chance für den Gymnasialbesuch für Viertklässler aus unterschiedlichen sozialen Schichten (3).

(1) Der »soziale Gradient« gibt an, wie viele Testpunkte ein Schüler durchschnittlich mehr erreicht hätte, wenn der sozioökonomische Status seiner Herkunftsfamilie um den Wert einer Standardabweichung⁷ höher wäre – je größer der soziale Gradient, desto stärker ist die soziale Ungleichheit ausgeprägt. Als Orientierungsmaß für den sozioökonomischen Status wird der HISEI⁸ herangezogen, der Werte zwischen 16 und 90 Punkten annehmen kann. Baden-Württemberg liegt mit einem mittleren HISEI von 50,1 Punkten etwas über dem Bundesdurchschnitt von 48,7 Punkten. In Baden-Württemberg betragen

die Werte der sozialen Gradienten im Fach Deutsch 32 Punkte für den Kompetenzbereich *Lesen*, 36 Punkte für *Zuhören* und 31 Punkte für *Orthografie*. Sie unterscheiden sich damit nicht bzw. nicht signifikant von den entsprechenden Gradientenwerten für Deutschland, die 32, 35 und 31 Punkte betragen. Im Fach Englisch liegen die Landeswerte mit 27 Punkten für den Bereich *Leseverstehen* signifikant und mit 31 Punkten für *Hörverstehen* nicht signifikant unterhalb der Werte für Deutschland, die 34 bzw. 36 Punkte hoch sind (Tabelle I 1 (T6) im Anhang). Für Französisch kann aufgrund geringer Fallzahlen kein Vergleich erfolgen.

(2) Als weiteres Maß für soziale Disparitäten wird die soziale Schicht herangezogen. Als Orientierungsmaß dient hierfür die Einteilung der Herkunftsfamilien in sieben EGP-Klassen.⁹ Die Schüler aus Familien der Klassen I und II wurden zusammengefasst, ebenso die Schüler aus Familien der Klassen V bis VII. Um den Einfluss sozialer Disparitäten auf die Testleistungen zu betrachten wurden die Leistungsunterschiede dieser beiden Extremgruppen miteinander verglichen. Je größer die Differenz der Gruppenmittelwerte, desto größer ist die soziale Ungleichheit. Die berechneten Differenzen liegen in den Kompetenzbereichen Deutsch mit 68, 81 und 63 Punkten zwar leicht, jedoch nicht statistisch signifikant über den entsprechenden bundesdeutschen Mittelwerten. Die Differenzen für die beiden Englisch-Kompetenzbereiche liegen mit 59 und 68 Punkten unterhalb des gesamtdeutschen Durchschnitts (Tabelle I 1 (T6) im Anhang).

(3) Als Indikator für soziale Disparitäten bei der Bildungsbeteiligung zieht der Ländervergleich die relative Chance eines Kindes, ein Gymnasium zu besuchen, in Abhängigkeit von der sozialen Schichtzugehörigkeit heran. Das Orientierungsmaß bilden wiederum die EGP-Klassen. Die relative Chance für den Gymnasialbesuch wird in Form von »Wettquotienten« ausgedrückt. Betrachtet wird dabei nur der Einfluss der sozialen Schichtzugehörigkeit, nicht die Leseleistungsfähigkeit des Kindes. Über statistische Verfahren kann gewährleistet werden, dass etwaige Leistungsunterschiede in der *Lesekompetenz* der betrachteten Schüler die Ergebnisse nicht beeinflussen. Die Werte geben somit Auskunft, um welchen Faktor die relative Chance eines Kindes auf den Gymnasialbesuch aus einer Familie einer bestimmten EGP-Klasse gegenüber einem Kind aus der Referenz-EGP-Klasse bei gleicher

7 Vgl. die Erläuterungen am Ende des Kapitels.

8 ISEI: International Socio-Economic Index of Occupational Status. Je höher der Wert, desto höher ist der sozioökonomische Status; HISEI: höchster ISEI-Wert der Eltern.

9 »Erikson-Goldthorpe-Portocarero-Klassen« als Indikator für die soziale Herkunft; EGP-Klasse I umfasst die »oberste Dienstleistungsklasse«, EGP-Klasse VII »ungelernte Arbeiter«.

Leseleistung höher ist.¹⁰ Für Baden-Württemberg ergibt diese Analyse, dass die Chance für den Besuch eines Gymnasiums für ein Kind aus der EGP-Klasse I um den Faktor 6,6, für ein Kind aus EGP-Klasse II um den Faktor 3,0 gegenüber einem im Lesen gleich starken Kind aus einer Facharbeiterfamilie (EGP-Klassen V und VI) höher liegt. Die Werte für Deutschland liegen mit Faktoren von 4,5 bzw. 2,5 darunter.

Dieses Modell gibt allerdings nur Auskünfte über die Übergänge von der Primarstufe in die Sekundarstufe I. Einer objektiven Analyse der Bildungsgerechtigkeit in Baden-Württemberg wird es nicht gerecht, da es die verschiedenen alternativen Möglichkeiten zum Erwerb einer Hochschulzugangsberechtigung nicht adäquat abbildet. Ein großer Teil der Hochschulzugangsberechtigungen wird nicht durch Abschlüsse an allgemein bildenden Gymnasien erworben, sondern an Bildungsgängen in beruflichen Schulen. Absolventen von Realschulen, Hauptschulen/Werkrealschulen und Berufsschulen können unter bestimmten Voraussetzungen die Fachhochschulreife und die allgemeine Hochschulreife an beruflichen Gymnasien, Berufskollegs oder Berufsoberschulen erlangen. So erhielten 2009 rund 30 % der Abiturienten ihr Zeugnis an einem beruflichen Gymnasium. Insgesamt wurde in diesem Jahr knapp die Hälfte aller Hochschulzugangsberechtigungen an beruflichen Schulen erworben (vgl. **Kapitel F 1**). Diese Möglichkeit des Erwerbs der Hochschulreife wird von Schülern aus Familien mit niedrigerem sozialem Status stärker genutzt.¹¹ Die Bedeutung des Übergangs von der Grundschule auf das Gymnasium wird hinsichtlich der Verteilung von Bildungschancen überschätzt, da in Baden-Württemberg alternative

Wege zum Erwerb einer Studienberechtigung eingeschlagen werden können.

Deutliche migrationsbedingte Disparitäten, jedoch weniger stark ausgeprägt als im Bundesschnitt

Für die Fächer Deutsch und Englisch wurden die Testleistungen der Schüler mit Migrationshintergrund¹² mit den Leistungen der anderen Schüler verglichen. In allen Kompetenzbereichen zeigt sich ein deutlicher Leistungsunterschied zu Ungunsten der Jugendlichen mit Migrationshintergrund.

Im Fach Deutsch, Kompetenzbereich *Lesen*, beträgt die Leistungsdifferenz 60 Punkte, im *Zuhören* 71 Punkte und in der *Orthografie* 47 Punkte (Tabelle I 1 (T7) im Anhang). Für das Fach Englisch fallen die Unterschiede weniger deutlich aus. Sie betragen im *Leseverstehen* 40 Punkte und im *Hörverstehen* 44 Punkte. Der Leistungsmittelwert der Schüler mit Migrationshintergrund liegt in Baden-Württemberg in allen getesteten Kompetenzbereichen über dem Durchschnittswert für Deutschland.

Die in der Studie zum Ländervergleich festgestellten Leistungsunterschiede von Jugendlichen mit Migrationshintergrund betreffen diese Schülergruppe nicht gleichermaßen. Deutliche Unterschiede sind zwischen Schülern unterschiedlicher Herkunftsnationen festzustellen. So liegen die mittleren Leistungen von Schülern aus den Herkunftsländern Polen und den Staaten der ehemaligen Sowjetunion in den Deutsch- und Englisch-Kompetenzbereichen fast durchweg mehr als 50 Punkte über den Leistungen von Schülern aus der Türkei (vgl. **Kapitel D 5**).

10 Ein alternatives Auswertungsmodell ermöglicht die Betrachtung des Effekts, der sich ohne Kontrolle der Leseleistung ergibt. Dieses Modell wird aufgrund seiner geringeren Aussagekraft hier jedoch nicht dargestellt.

11 Köller, O.; Watermann, R.; Trautwein, U.; Lüdtke, O. (Hrsg.) (2004): Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung der allgemein bildenden und beruflichen Gymnasien in Baden-Württemberg.

12 Entsprechend der Definition der Studie zum Ländervergleich liegt ein Migrationshintergrund vor, wenn beide Elternteile bzw. die Eltern und der Jugendliche im Ausland geboren wurden.

i

Standardabweichung und Standardfehler

Standardabweichung (SD): Maß für die Streuung der Messwerte. Wird die Standardabweichung zu beiden Seiten des Mittelwerts aufgetragen, liegen etwa 68 % der Werte innerhalb dieses Intervalls.

Standardfehler (SE): Dieses Maß gibt an, wie weit ein Stichprobenwert um den wahren Wert der Grundgesamtheit streut. Der Wert der Grundgesamtheit liegt mit einer Wahrscheinlichkeit von etwa 68 % innerhalb des Intervalls, das durch den Wert des Standardfehlers bestimmt wird.

Signifikanz: Ein empirisch festgestellter Unterschied zwischen zwei Untersuchungsgruppen gilt als signifikant, wenn die Annahme berechtigt ist, dass der vorgefundene Unterschied nicht als zufälliger, durch die mit der Stichprobenziehung verbundene Unschärfe entstandener, erklärt werden kann. Zur Überprüfung der Signifikanz werden teststatistische Verfahren herangezogen. In der Regel wird ein Unterschied als ausreichend signifikant akzeptiert, wenn er nur in 5 % aller Stichprobenziehungen rein zufällig auftreten würde, obwohl er in der Grundgesamtheit gar nicht besteht.

I 2 Fremdevaluation

Selbst- und Fremdevaluation sind zentrale Instrumente der Schul- und Unterrichtsentwicklung an öffentlichen Schulen in Baden-Württemberg. Sie gewährleisten über systematische und kriteriengeleitete Verfahren der Datengewinnung Rückmeldungen zur Art und Weise des professionellen pädagogischen Handelns. Die Evaluation erfolgt in einem zweistufigen, abgestimmten Verfahren: Mittels kontinuierlicher Selbstevaluationen nehmen die Schulen interne Maßnahmen der Qualitätsentwicklung eigenverantwortlich in den Blick. Diese Standortbestimmungen werden durch eine umfangreiche, die zentralen schulischen Prozesse betrachtende Fremdevaluation ergänzt, die in regelmäßigen Abständen durchgeführt wird.

Nach vorangegangenen Pilotphasen wurde die Selbstevaluation im Schuljahr 2007/08 für die allgemein bildenden Schulen und seit 2008/09 für die beruflichen Schulen eingeführt. An den beruflichen Schulen erfolgt der Einstieg in die Prozesse der Konzeption *Operativ Eigenständige Schule (OES)* stufenweise über vier Jahre hinweg. Seit dem Schuljahr 2008/09 ist die Fremdevaluation an allen öffentlichen Schulen in Baden-Württemberg verbindlich. Für die Organisation und Durchführung der Fremdevaluation ist das Landesinstitut für Schulentwicklung zuständig.

Selbst- und Fremdevaluation bieten eine datengestützte Basis für die Analyse und Weiterentwicklung schulischer Qualität. Im allgemein bildenden Bereich werden diese beiden Elemente der Qualitätssicherung durch verbindlich eingeführte zentral erstellte Lernstandserhebungen, den Diagnosearbeiten an Grundschulen (VERA 3, vgl. **Kapitel I 3.1**) und Vergleichsarbeiten an weiterführenden Schulen (DVA, vgl. **Kapitel I 3.2**) ergänzt.

I 2.1 Fremdevaluation an allgemein bildenden Schulen

I 2.1.1 Grundlagen

Die Evaluation an allgemein bildenden Schulen erfolgt auf Basis eines verbindlichen Referenzsystems zur Schulqualität in Baden-Württemberg. Für die Selbstevaluation ist dies der »Orientierungsrahmen zur Schulqualität« und für die Fremdevaluation der darauf abgestimmte »Qualitätsrahmen Fremdevaluation«.

Die Fremdevaluation wird vom Landesinstitut für Schulentwicklung mit für diese Tätigkeit abgeordneten Lehrkräften aus allen allgemein bildenden Schularten durchgeführt. Dabei werden die Schulen – je nach Größe – von einem aus zwei bis vier Mitgliedern bestehenden Evaluationsteam evaluiert. Mindestens ein Teammitglied verfügt über die Lehrbefähigung für die zu evaluierende Schulart. Auf Wunsch der Schule kann eine von dort benannte Vertrauensperson als sogenanntes »assoziertes Mitglied« das Evaluationsteam begleiten und dessen Wahrnehmungen durch eine weitere Sicht ergänzen.

Jede Fremdevaluation erfolgt nach einem standardisierten Vorgehen, in dem verschiedene qualitative Methoden der Datenerhebung zum Einsatz kommen. Sämtliche am Schulleben beteiligten Gruppen werden in die Evaluation einbezogen. Der Ablauf gliedert sich in folgende Schritte:

- Vorgespräch der Teamleitung mit der Schule,
- Analyse der von der Schule übersandten Unterlagen zur schulischen Qualitätsentwicklung, des sogenannten Schulportfolios,
- Evaluationsplanung in Abstimmung mit der Schule,
- Organisatorische Vorbereitung der Fremdevaluation durch die Schule,
- Datenerhebung und -dokumentation an der Schule im Rahmen von leitfadengestützten Interviews mit Schulleitung, Lehrkräften, Schülern und Eltern, Beobachtungen von Unterrichtssituationen bei mindestens der Hälfte aller Klassen bzw. Kurse sowie einem Schulhausrundgang,
- Auswertung der Daten,
- Erstellung des vorläufigen Evaluationsberichts,
- Ergebnispräsentation in der Gesamtlehrerkonferenz,
- Fertigstellung und Versand des Evaluationsberichts an die Schule.

Die Datenerhebung und -auswertung erfolgt in Bezug auf die im *Qualitätsrahmen Fremdevaluation* festgelegten Evaluationsbereiche und den dort formulierten Standards. Der *Qualitätsrahmen* ist thematisch analog zum

Orientierungsrahmen zur Schulqualität aufgebaut, der den Referenzrahmen für die Selbstevaluation bildet. Er gliedert sich insgesamt in sechs *Qualitätsbereiche*: *QB I Unterricht, QB II Professionalität der Lehrkräfte, QB III Schulführung und Schulmanagement, QB IV Schul- und Klassenklima, QB V Innerschulische und außerschulische Partnerschaften sowie QB Q Qualitätssicherung und Qualitätsentwicklung* (Tabelle I 2.1 (T1)). Jeder dieser Qualitätsbereiche ist in *Kriterien* unterteilt, die wiederum in 34 *Merkmale* differenziert sind. Diese umfassen einen obligatorischen Bereich im Umfang von 22 Qualitätsmerkmalen, der an allen evaluierten Schulen betrachtet wird, und einen sogenannten Wahlpflichtbereich, aus dem die Schule vier weitere Qualitätsmerkmale auswählt. Mit der Ausgestaltung in obligatorische und Wahlpflichtbereiche gewährleistet die Fremdevaluation Aussagen zu einem breiten Spektrum verbindlicher Kriterien und erlaubt gleichzeitig eine Rückmeldung zu schulspezifischen Schwerpunktsetzungen.

Um den Schulen eine konkrete und differenzierte Rückmeldung zu ihrem momentanen Entwicklungsstand geben zu können, ordnet das Evaluationsteam anhand der erhobenen Daten schulische Qualität auf Ebene der *Merkmale* ein. Die Einordnungen werden als *Entwicklungsstufen* bezeichnet. Diese lauten:

- Noch zu entwickelnde Praxis
(Die auf das jeweilige Merkmal bezogenen Ziele sind im Sinne einer funktionsfähigen Praxis noch nicht erreicht.)
- Gute Praxis
(Zielstufe: Die Ziele sind im Sinne einer funktionsfähigen Praxis erreicht.)
- Ausgezeichnete Praxis
(Die Ziele werden durch herausragende Praxis weit übertroffen; oft handelt es sich dabei um besondere Schwerpunkte der Schule.)

Dabei ist zu berücksichtigen, dass die drei Entwicklungsstufen in der schulischen Realität ein Kontinuum sind.

In einem ausführlichen Evaluationsbericht erhalten die Schulen zu jedem einzelnen evaluierten Merkmal eine spezifische Rückmeldung auf Grundlage der erhobenen Daten. Dieser beschreibende Teil macht die Einstufung transparent und nachvollziehbar. Der Bericht schließt mit konkreten Empfehlungen für die Schulentwicklung, die sich auf einzelne Qualitätsmerkmale bzw. -kriterien beziehen. Die Schule leitet den Evaluationsbericht an den Schulträger sowie an die zuständige Schulaufsichtsbehörde weiter. Letztere trifft auf Grundlage des Berichts und weiterer Ziele der Schule und Vorgaben des Landes Zielvereinbarungen mit der Schule.

Organisation der Fremdevaluation an allgemein bildenden Schulen

Die in einem bestimmten Schuljahr an der Fremdevaluation teilnehmenden Schulen werden durch ein kriteriengeleitetes Zufallsverfahren ermittelt. Das Verfahren der anonymisierten Ziehung beruht im Wesentlichen auf folgenden Schritten:

- Berechnung des Anteils der Schulen und des entsprechenden Anteils einer jeden Schulart
- Ermittlung der Anteile jeder Schulart auf Ebene der Stadt- und Landkreise innerhalb eines Regierungsbezirks
- Ziehung der berechneten Schulanzahl pro Schulart in den Stadt-/ Landkreisen nach dem Zufallsprinzip für einen bestimmten Zeitraum

Dieses Verfahren ermöglicht eine repräsentative, jedoch zufällige Verteilung der evaluierten Schulen über das ganze Land bei ausgewogener Berücksichtigung aller Schularten (Grafik I 2.1 (G1)).

I 2.1.2 Ergebnisse

In den Schuljahren 2008/09 und 2009/10 wurden 476 allgemein bildende Schulen evaluiert. Davon sind

- 163 Grundschulen,
- 156 Grund- und Hauptschulen/ Hauptschulen/ Hauptschulen mit Werkrealschule,
- 57 Realschulen,
- 59 Gymnasien,
- 35 Sonderschulen sowie
- 6 Schulen sonstiger Schulverbünde, welche in der Auswertung unberücksichtigt bleiben.

Häufigkeitsverteilungen der Einordnungen für den obligatorischen Bereich

Eine differenzierte Sichtweise auf die Evaluationsergebnisse leisten Häufigkeitsverteilungen. In der Visualisierung in Form von Balkendiagrammen werden die relativen Verteilungen der Einstufungen abgebildet. Somit kann abgelesen werden, welcher prozentuale Anteil der Schulen einer Schulart eine bestimmte Einordnung erhielt (Länge des Balkensegments). Die prozentualen Anteile ermöglichen darüber hinaus eine schularten-übergreifende Übersicht von Qualitätsmerkmalen, bei denen alle Schularten ähnliche Ergebnisse erzielen, und von Merkmalen, bei denen mehr oder weniger deutliche Einordnungsunterschiede vorliegen. (Fortsetzung auf Seite 295)

I 2.1 (T1) Qualitätsbereiche, Kriterien und Merkmale des Qualitätsrahmens Fremdevaluation im Überblick

Qualitätsbereich	Kriterium ¹⁾	Anzahl der Merkmale	Merkmale
QB I Unterricht	I 2 Gestaltung der Lehr-Lern-Prozesse	8	Nutzung der Lernzeit Strukturierung der Lernangebote Differenzierungsangebote Aktivierung der Schüler Förderung von methodischen Kompetenzen Förderung von sozialen und personalen Kompetenzen Kommunikationsstil im Unterricht Gestaltung der Klassenzimmer, Fachräume
	I 3 Praxis der Leistungsbeurteilung und Leistungsrückmeldung	2	Leistungsrückmeldung als Orientierungshilfe Abgestimmtheit der Kriterien
QB II Professionalität der Lehrkräfte	II 1 Kooperation	1	Nutzen der kollegialen Zusammenarbeit
	II 2 Praxis der Weiterqualifizierung	2	Bedarfsorientierte Fortbildungsplanung Weitergabe und Nutzung von Wissen
	II 3 Umgang mit beruflichen Anforderungen	1	Verbesserung von innerschulischen Arbeitsbedingungen
QB III Schulführung und Schulmanagement	III 1 Führung	3	Entwicklung von Unterricht und Schule Umgang mit Konflikten und Problemen Möglichkeiten der Personalentwicklung
	III 2 Verwaltung und Organisation	2	Funktionale Arbeitsabläufe Einsatz von Ressourcen
QB IV Schul- und Klassenklima	IV 1 Schulleben	2	Gestaltung des Umgangs miteinander Zusätzliche Angebote für Schülerinnen und Schüler
	IV 2 Mitgestaltungsmöglichkeiten der Schülerinnen und Schüler	2	Einbeziehung in die Gestaltung des Schullebens Angebote zum Aufbau eines demokratischen Grundverständnisses
QB V Innerschulische und außerschulische Partnerschaften	V 1 Mitgestaltungsmöglichkeiten der Eltern	2	Einbeziehung in das schulische Leben Erziehungspartnerschaft
	V 2 Zusammenarbeit mit anderen Institutionen	2	Kooperation mit schulischen Partnern Kooperation mit außerschulischen Partnern
QB Q Qualitätssicherung und Qualitätsentwicklung	Q 1 Pädagogische Grundsätze	1	Handlungsleitende pädagogische Grundsätze
	Q 2 Strukturen der schulischen Qualitätsentwicklung	2	Funktionalität der Strukturen Nutzen der Dokumentation
	Q 3 Durchführung der Selbstevaluation	3	Orientierung an den schulischen Zielen und Ressourcen Systematisches Vorgehen unter Beachtung von Evaluationsgrundsätzen Konsequenzen ziehen
	Q 4 Individualfeedback	1	Vorbereitung und Durchführung von Individualfeedback

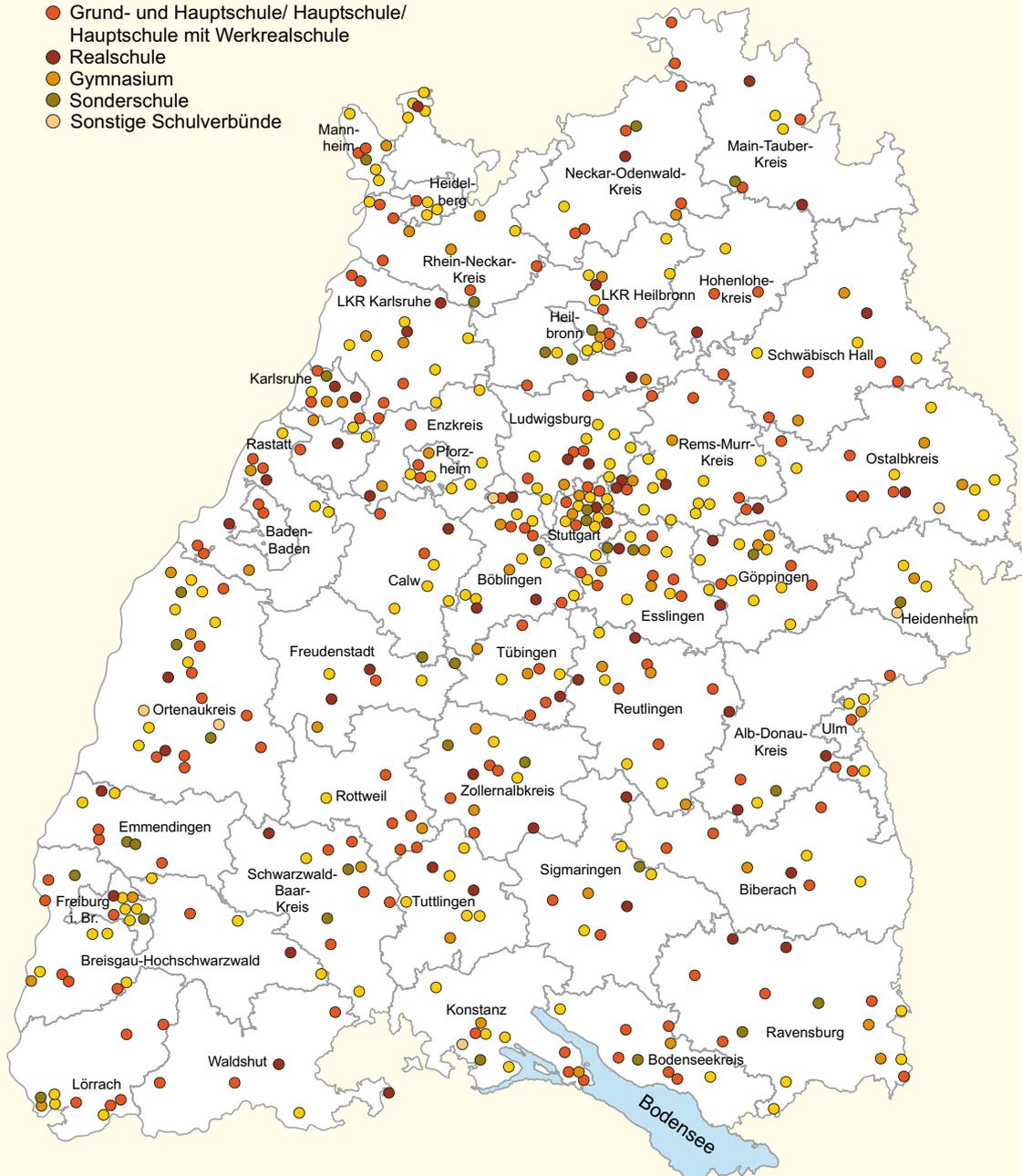
Text = Obligatorische Bereiche, Text = Wahlpflichtbereiche

1) Kriterium I 1 Schulinterne Umsetzung des Bildungsplans wird nicht in die Fremdevaluation einbezogen.

I 2.1 (G1)

Regionale Verteilung der evaluierten allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10

- Grundschule
- Grund- und Hauptschule/ Hauptschule/ Hauptschule mit Werkrealschule
- Realschule
- Gymnasium
- Sonderschule
- Sonstige Schulverbände



Datenquelle: Landesinstitut für Schulentwicklung.

Statistisches Landesamt Baden-Württemberg
Landesinformationssystem

22-22-11-015
© Kartengrundlage GfK GeoMarketing GmbH
Karte erstellt mit RegioGraph

Das Konzept der Fremdevaluation an allgemein bildenden Schulen

Das Konzept der Fremdevaluation an allgemein bildenden Schulen in Baden-Württemberg basiert auf aktuellen und wissenschaftlich fundierten Erkenntnissen der Schul- und Unterrichtsforschung, auf Vorgaben der Evaluationsverordnung, auf Standards des Bildungsplans sowie weiteren allgemeinen Vorgaben. Zusammenfassend werden die Grundlagen sowie die zu betrachtenden Merkmale für Schulqualität im *Orientierungsrahmen zur Schulqualität für allgemein bildende Schulen in Baden-Württemberg*¹ beschrieben und einzelnen Kriterien und Qualitätsbereichen zugeordnet. Zentrale Informationen über Konzeption und Verfahren der Fremdevaluation an allgemein bildenden Schulen sind im *Qualitätsrahmen Fremdevaluation*² zusammengestellt.

Aufbau des Qualitätsrahmens Fremdevaluation am Beispiel des Qualitätsbereichs Unterricht

Im *Qualitätsrahmen Fremdevaluation* werden sechs Bereiche, in denen sich die Qualität der schulischen Arbeit zeigen kann, ausgewiesen. Hierbei wird eine analytische Trennung von Bereichen vorgenommen, die in der schulischen Praxis ineinander wirken. Zentral ist dabei der Qualitätsbereich *Unterricht*. Innerhalb der Qualitätsbereiche wird je nach Komplexität eine unterschiedliche Anzahl von Kriterien unterschieden, denen wiederum mehrere Merkmale zugeordnet sind.

Entsprechend dem *Orientierungsrahmen zur Schulqualität für allgemein bildende Schulen in Baden-Württemberg* besteht der Qualitätsbereich Unterricht aus den drei Kriterien:

- I 1 *Schulinterne Umsetzung des Bildungsplans*,
- I 2 *Gestaltung der Lehr-Lern-Prozesse* und
- I 3 *Praxis der Leistungsbeurteilung und Leistungsrückmeldung* (vgl. Auszug aus dem Qualitätsbereich *Unterricht*).

Das Kriterium I 1 *Schulinterne Umsetzung des Bildungsplans* wird aufgrund des damit verbundenen nicht darstellbaren Ressourcenaufwands nicht in die Fremdevaluation einbezogen.

Kriterien sind inhaltliche Untergliederungen eines Qualitätsbereichs. Dabei handelt es sich um verbindliche Beschreibungen für die Qualität von Teilbereichen, die mittelfristig im Laufe der Schulentwicklung erreicht werden sollen. Die Kriterienbeschreibungen sind als verbindliche Qualitätsansprüche bzw. Qualitätsziele zu verstehen. Jedes Kriterium setzt sich wiederum aus mehreren Merkmalen zusammen, auf deren Ebene die Einschätzungen bei der Fremdevaluation vorgenommen werden.

Die Merkmale weisen das höchste Konkretisierungsniveau im Qualitätsrahmen auf und bieten daher praktische Anknüpfungspunkte für die Qualitätsarbeit an der Schule. Die Kriterien werden unterschiedlich stark in Merkmale aufgefächert und umfassen diejenigen Teilaspekte der Kriterien, die im Rahmen der Fremdevaluation mit einem vertretbaren Ressourceneinsatz extern evaluierbar sind. Im Beispiel sind die beiden Merkmale *Leistungsrückmeldung als Orientierungshilfe* und *Abgestimmtheit der Kriterien* sowie die inhaltliche Beschreibung der Entwicklungsstufen dargestellt.

Die Wirksamkeit einer »guten Praxis« im Bereich Unterricht hängt vom komplexen Bedingungsgefüge in den jeweiligen Lerngruppen, wie zum Beispiel den unterschiedlichen Lernvoraussetzungen oder dem sozialen und kulturellen Hintergrund der Schüler, ab. Die auf den Unterricht bezogenen Merkmale des *Orientierungsrahmens zur Schulqualität für allgemein bildende Schulen in Baden-Württemberg* berücksichtigen

1 Vgl. Ministerium für Kultus, Jugend und Sport Baden-Württemberg/ Landesinstitut für Schulentwicklung (2007).

2 Vgl. Landesinstitut für Schulentwicklung (2009).

Kriterium I3 – Auszug aus dem Qualitätsbereich *Unterricht*

Praxis der Leistungsbeurteilung und Leistungsrückmeldung

Die Leistungsbeurteilung ist mehr als ein reiner Leistungs- und Berechtigungsnachweis. Sie stellt für Schüler, Lehrkräfte und Eltern eine wichtige Orientierungshilfe für die Einordnung des Lernstandes und Lernfortschrittes dar. Die Schüler werden systematisch dahin geführt, ihre eigenen Leistungen einzuschätzen. Dies trägt zur Persönlichkeitsentwicklung bei. Die Abstimmung der Kriterien zur Leistungsbeurteilung erfolgt weitgehend kooperativ im Teilkollegium beziehungsweise in Fachschaften.

Merkmal: Leistungsrückmeldung als Orientierungshilfe	Entwicklungsstufen		
	<i>Noch zu entwickelnde Praxis</i>	<i>Gute Praxis</i>	<i>Ausgezeichnete Praxis</i>
	Leistungsrückmeldungen sind für die Schüler sowie deren Eltern wenig Orientierungshilfe für das weitere Lernen.	Leistungsrückmeldungen sind für die Schüler sowie deren Eltern überwiegend Orientierungshilfe für das weitere Lernen.	Leistungsrückmeldungen sind für die Schüler sowie deren Eltern durchgängig Orientierungshilfe für das weitere Lernen.
Merkmal: Abgestimmtheit der Kriterien	Entwicklungsstufen		
	<i>Noch zu entwickelnde Praxis</i>	<i>Gute Praxis</i>	<i>Ausgezeichnete Praxis</i>
	Die Kriterien zur Leistungsbeurteilung sind im Kollegium wenig abgestimmt.	Die Kriterien zur Leistungsbeurteilung sind im Kollegium überwiegend abgestimmt.	Die Kriterien zur Leistungsbeurteilung sind im Kollegium durchgängig abgestimmt.

Datenquelle: Landesinstitut für Schulentwicklung 2009, S. 29

sichtigen dies. Sie leiten sich einerseits aus den Standards des Bildungsplans für Baden-Württemberg aus dem Jahre 2004 bzw. 2008 ab und lassen sich andererseits auf die wissenschaftliche Arbeiten unter anderem von Helmke³, Dubs⁴, Fend⁵ und Oelkers⁶ zurückführen.

Wie erfolgt die Zuordnung eines Merkmals zu einer der drei Entwicklungsstufen?

Der Ablauf der Fremdevaluation für allgemein bildende Schulen in Baden-Württemberg ist in folgende fünf Phasen gegliedert: Planung, Vorbereitung, Datenerhebung vor Ort, Datenauswertung und Datenrückmeldung.⁷ An dieser Stelle soll auf die Datenauswertung genauer eingegangen werden, das heißt wie die Evaluatoren die Einordnung eines Merkmals zu einer der drei Entwicklungsstufen vornehmen.

Nach Abschluss des Schulbesuchs vor Ort werden die im Rahmen der Fremdevaluation erhobenen Daten aus allen Datenerhebungsverfahren (Dokumentenanalyse, Beobachtungen beim Schulhausrundgang, Beobachtungen von Unterrichtssituationen, Interviews mit Schülern, Eltern, Lehrkräften und Schulleitung) zusammengeführt. Die anschließende Datenauswertung erfolgt merkmalsweise durch eine »Triangulation« der Daten. Dabei werden die gewonnenen Daten im ersten Schritt aus einem Datenerhebungsverfahren und in einem zweiten Schritt aus unterschiedlichen Verfahren miteinander verglichen. Auf diese Weise wird sichergestellt, dass nur mehrmals bestätigte Aussagen Grundlage für die Einschätzung werden und nur solche in die Beschreibung und Rückmeldung über die Schule als System in den Evaluationsbericht einfließen.

3 Vgl. Helmke, A. (2007), S. 111–131 und S. 283–294.

4 Vgl. Dubs, R. (2006), S. 1 220–1 230.

5 Vgl. Fend, H. (1998).

6 Vgl. Oelkers, J. (2003).

7 Vgl. Landesinstitut für Schulentwicklung 2009, S. 47ff.

Das Auswertungsverfahren der Fremdevaluation ist standardisiert und regelgeleitet, sodass unabhängig vom jeweils eingesetzten Evaluationsteam einheitliche Einschätzungen der evaluierten Merkmale gewährleistet werden. Die Einschätzung der evaluierten Merkmale des Qualitätsrahmens erfolgt in drei Kategorien (»Noch zu entwickelnde Praxis«, »Gute Praxis«, »Ausgezeichnete Praxis«), die als Entwicklungsstufen zu verstehen sind.

Diese drei Entwicklungsstufen sind in der schulischen Praxis ein Kontinuum, das heißt die Grenzen sind in Wirklichkeit fließend. Es ist damit Aufgabe des Evaluationsteams, für jedes evaluierte Merkmal der jeweiligen Schule zu entscheiden,

- ob hier ein sogenannter Verbesserungsbereich vorliegt (= »Noch zu entwickelnde Praxis«),
- die Schule einen für sie geeigneten Weg oder passende Maßnahmen gefunden hat, um im Alltag zur Zufriedenheit der Beteiligten »gut zu funktionieren« (= »Gute Praxis«) oder
- ob die Schule in dem Bereich über das übliche Maß hinaus Lösungen und Wege gefunden hat, die als Best Practice bezeichnet werden können und damit an der Schule »herausragen« (= »Ausgezeichnete Praxis«).

Für eine Einordnung in die Entwicklungsstufe »Gute Praxis« spielt es keine Rolle, auf welchem Weg und mit welchen Maßnahmen es der Schule gelingt, die Anforderungen zu erfüllen, da die Funktionalität bzw. der Nutzen an der Schule im Mittelpunkt stehen.

Die Übersicht der Evaluationsergebnisse im Qualitätsbereich »Unterricht« (QB I) bietet ein heterogenes Bild mit unterschiedlichen Ausprägungen sowohl zwischen den einzelnen Schularten als auch zwischen den einzelnen Qualitätsmerkmalen (Grafik I 2.1 (G2) und Web-Tabelle I 2.1 (T2)). Hierbei ist zu berücksichtigen, dass die fachliche Umsetzung des Bildungsplans nicht Gegenstand der Fremdevaluation ist und daher Merkmale wie die fachliche Qualität des Unterrichts oder die pädagogische Fähigkeit der Lehrkraft nicht erfasst werden. Eine fächer- und schulartspezifische externe Einschätzung ist für alle Fächer/ Fächerverbünde mit den eingesetzten Evaluationsteams (zwei bis vier Personen pro Team) nicht möglich.

Die Merkmale *Aktivierung der Schülerinnen und Schüler*, *Förderung von sozialen und personalen Kompetenzen* und *Kommunikationsstil im Unterricht* werden überwiegend mit »guter Praxis« eingestuft. Die Einordnung »noch zu entwickelnde Praxis« ist sehr selten vertreten, öfter hingegen, mit Anteilen bis zu knapp 50 % – insbesondere bei den Sonderschulen und den Grundschulen – »ausgezeichnete Praxis«.

Drei von vier Gymnasien, über 80 % der Grundschulen, Grund- und Hauptschulen/ Hauptschulen und Sonderschulen sowie 90 % aller Realschulen sind beim Merkmal *Förderung von methodischen Kompetenzen* »guter«, teilweise auch »ausgezeichnete Praxis« zugeordnet. Die Gymnasien weisen bei diesem Merkmal in knapp einem Viertel der Fälle noch Entwicklungs-

möglichkeiten auf. Der Anteil dieser Einstufung ist bei den anderen Schularten geringer.

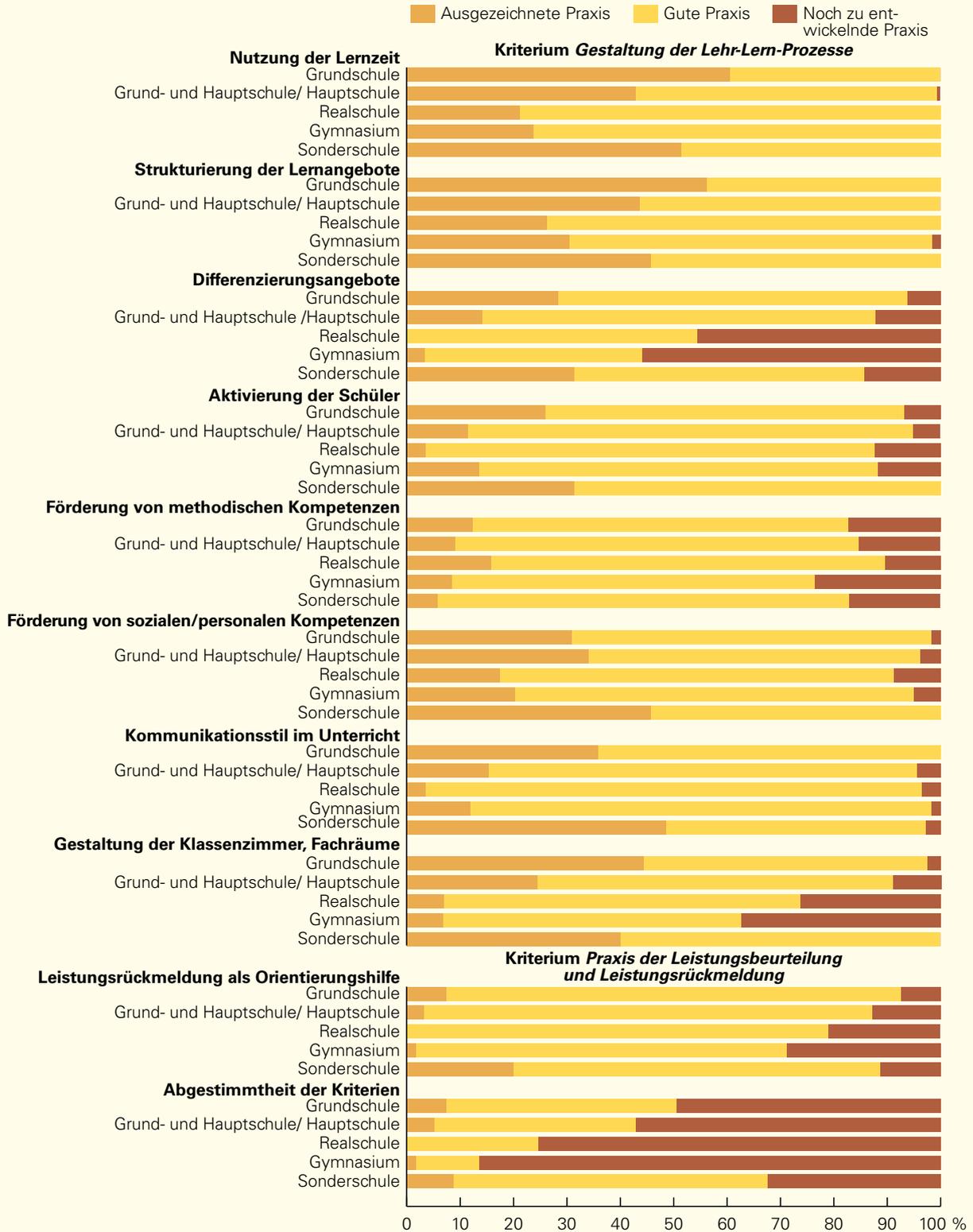
Unterschiede zwischen den Schularten sind beim Merkmal *Differenzierungsangebote* erkennbar. Die Darstellung der Häufigkeitsverteilungen verdeutlicht, dass jeweils rund ein Drittel der Grundschulen und Sonderschulen hier eine »ausgezeichnete Praxis« vorweisen können und die Mehrzahl der Schulen dieser Schularten eine »gute Praxis« erzielen. Bei den Grund- und Hauptschulen/ Hauptschulen erreichen 14 % diese höchste Einordungskategorie und 74 % eine »gute Praxis«. Bei den Realschulen und Gymnasien ist dieses Merkmal weniger ausgeprägt. Knapp jede zweite Realschule und über die Hälfte der Gymnasien sind hier einer »noch zu entwickelnden Praxis« zugeordnet.

Deutliche schulartspezifische Unterschiede weist auch das Merkmal *Gestaltung der Klassenzimmer/ Fachräume* auf. Über 40 % der Grundschulen und in etwas geringerem Umfang die Sonderschulen erzielen hier eine »ausgezeichnete Praxis«, die restlichen Schulen dieser Schularten mit wenigen Ausnahmen eine »gute Praxis«. Auch bei neun von zehn Grund- und Hauptschulen/ Hauptschulen wird hier »gute Praxis« oder sogar »ausgezeichnete Praxis« festgestellt. Ein Viertel der Realschulen und knapp 40 % der Gymnasien sind bei »noch zu entwickelnder Praxis« eingeordnet.

Das Qualitätsmerkmal *Leistungsrückmeldung als Orientierungshilfe* wird nahezu bei drei von vier Gymnasien mit mindestens »guter Praxis« bewertet, bei den

I 2.1 (G2)

Entwicklungsstufen der allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Qualitätsmerkmal und Schultart: QB I Unterricht



Datenquelle: Landesinstitut für Schulentwicklung

Statistisches Landesamt Baden-Württemberg

275 11

Realschulen und Grund- und Hauptschulen/ Hauptschulen zu rund 80 %. Noch positiver schneiden die Grundschulen und die Sonderschulen ab.

Schulartübergreifend konstatiert die Fremdevaluation beim Merkmal *Abgestimmtheit der Kriterien* für die Leistungsrückmeldung den insgesamt deutlichsten Entwicklungsbedarf. Jede zweite Grundschule und mehr als die Hälfte der Grund- und Hauptschulen/ Hauptschulen erhalten hier lediglich eine »noch zu entwickelnde Praxis«. An den Realschulen ist dies bei 75 %, bei den Gymnasien zu 86 % der Fall. Die *Nutzung der Lernzeit* und die *Strukturierung der Lernangebote* gelingen bei nahezu allen evaluierten Schulen in »guter«, häufig sogar in »ausgezeichneter Praxis«.

Im zweiten Pflichtbereich der Fremdevaluation – dem Qualitätsbereich »Schulführung und Schulmanagement« (QB III) – liegen über alle Schularten und Merkmale hinweg die evaluierten Schulen zu über 85 % entweder auf der Stufe »gute Praxis« oder »ausgezeichnete Praxis« (Grafik I 2.1 (G3) und Web-Tabellen I 2.1 (T3)). Beim Qualitätsmerkmal *Entwicklung von Unterricht und Schule* erzielen über die Hälfte der Grundschulen und Sonderschulen eine »ausgezeichnete Praxis«. Auch die Evaluation des Merkmals *Möglichkeiten der Personalentwicklung* verdeutlicht, dass diese Führungsaufgabe über alle Schularten hinweg an rund neun von zehn Schulen in »guter« oder gar »ausgezeichneter Praxis« umgesetzt wird.

Die Merkmale *Funktionale Arbeitsabläufe* und *Einsatz von Ressourcen* werden bei nahezu allen Schulen mit »guter«, bei vielen auch mit »ausgezeichneter Praxis« beurteilt. Lediglich in einzelnen Fällen liegt eine »noch zu entwickelnde Praxis« vor. Statistisch bedeutsame schulartspezifische Unterschiede sind nur beim Merkmal *Entwicklung von Unterricht und Schule* zu finden. Hier erzielen die Grundschulen und Sonderschulen geringfügig höhere Einordnungen.

Der Aufbau und die Weiterentwicklung von Verfahren im Bereich der systematischen »Qualitätssicherung und Qualitätsentwicklung« (QB Q) sind für Schulen ganz neue Aufgabenfelder. So ist wenig verwunderlich, dass die in diesem Evaluationsbereich erzielten Ergebnisse oftmals eine »noch zu entwickelnde Praxis« an den Schulen widerspiegeln (Grafik I 2.1 (G4) und Web-Tabellen I 2.1 (T4)). Die Evaluationsergebnisse verdeutlichen folglich den gegenwärtigen Einführungsprozess hinsichtlich der Planung und Umsetzung eines systematischen Qualitätsmanagements an den allgemein bildenden Schulen.

Beim Merkmal *Orientierung an den schulischen Zielen und Ressourcen* erzielten über 80 % der Schulen eine »gute« oder »ausgezeichnete Praxis«, wobei der Anteil der mit »zu entwickelnder Praxis« eingestuften Schu-

len ähnlich groß ist wie der Anteil derjenigen, die mit »ausgezeichneter Praxis« eingestuft wurden.

Bei den anderen Merkmalen dieses Qualitätsbereichs – *Handlungsleitende pädagogische Grundsätze, Funktionalität der Strukturen, Nutzen der Dokumentation, Systematisches Vorgehen unter Beachtung von Evaluationsgrundsätzen und Konsequenzen ziehen* aus Selbstevaluationen – liegen die Anteile der Schulen, die mit »noch zu entwickelnder Praxis« bewertet wurden, zwischen 20 % und über 40 % relativ hoch.

Der größte Entwicklungsbedarf der evaluierten Schulen besteht hinsichtlich der *Vorbereitung und Durchführung von Individualfeedback*. Diese noch sehr junge Maßnahme der Qualitäts- und Unterrichtsentwicklung wird bislang an rund 10 % der Grundschulen, Grund- und Hauptschulen/ Hauptschulen und Realschulen sowie an 20 % der Sonderschulen und Gymnasien in »guter Praxis« und vereinzelt in »ausgezeichneter Praxis« umgesetzt.

Häufigkeitsverteilungen der Einordnungen für den Wahlpflichtbereich

Die Qualitätsbereiche »Professionalität der Lehrkräfte« (QB II), »Schul- und Klassenklima« (QB IV) und »Innerschulische und außerschulische Partnerschaften« (QB V) sind Wahlpflichtbereiche. Die Schulen haben im Rahmen der Fremdevaluation aus zwei dieser Qualitätsbereiche jeweils zwei zu evaluierende Qualitätsmerkmale auszuwählen. Bislang wurde die Verbindung der Merkmale *Nutzen der kollegialen Zusammenarbeit* und *Verbesserung von innerschulischen Arbeitsbedingungen* aus QB II favorisiert. 375 von insgesamt 476 Schulen haben diese Option gewählt. Weitere häufige Kombinationen sind mit 291 Schulen die beiden Merkmale *Gestaltung des Umgangs miteinander* und *Zusätzliche Angebote für Schülerinnen und Schüler* aus QB IV (Kriterium *Schul生活*) sowie die Merkmale *Kooperation mit schulischen Partnern* und *Kooperation mit außerschulischen Partnern* aus QB V (Kriterium *Zusammenarbeit mit anderen Institutionen*) mit 123 Schulen. Am wenigsten, lediglich von 25 Schulen, wurden die Merkmale *Einbeziehung in die Gestaltung des Schullebens* und *Angebote zum Aufbau eines demokratischen Grundverständnisses* gewählt.

Für den Qualitätsbereich »Professionalität der Lehrkräfte« (QB II) ist festzuhalten, dass an allen Schularten fast sämtliche evaluierten Merkmale zu über 80 % die Einstufung »gute Praxis« bzw. »ausgezeichnete Praxis« erhalten (Grafik I 2.1 (G5) und Web-Tabellen I 2.1 (T5)). Bei den beiden Merkmalen *Nutzen der kollegialen Zusammenarbeit* und *Verbesserung von innerschulischen Arbeitsbedingungen* liegen schulartspe-

I 2.1 (G3)

Entwicklungsstufen der allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Qualitätsmerkmal und Schulart: QB III Schulführung und Schulmanagement



Datenquelle: Landesinstitut für Schulentwicklung

Statistisches Landesamt Baden-Württemberg

274 11

zifische Unterschiede vor. Einem geringfügig höheren Anteil der Realschulen wird eine »noch zu entwickelnde Praxis« zugeordnet.

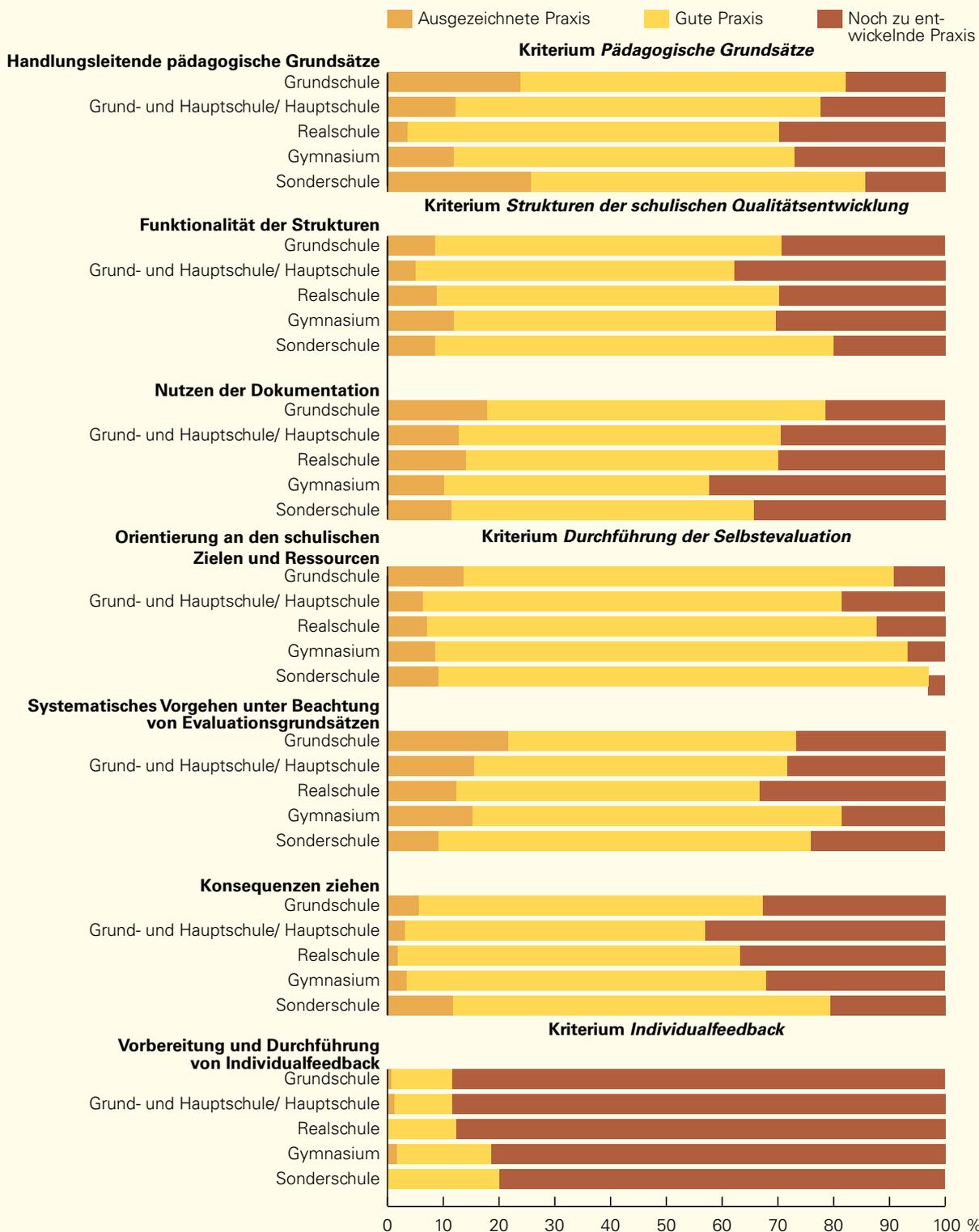
Über 90 % der evaluierten Schulen mit den beiden Merkmalen des Qualitätsbereichs »Schul- und Klassenklima« (QB IV) wurden mit »guter Praxis« bzw. »ausgezeichneter Praxis« beurteilt (Grafik I 2.1 (G6) und Web-Tabelle I 2.1 (T6)). Vor allem das Merkmal *Zusätzliche Angebote für Schülerinnen und Schüler*

hebt sich bei allen Schularten und insbesondere an den Gymnasien besonders positiv ab.

75 % bis 100 % der Schulen erreichen in den Merkmalen des Qualitätsbereichs »Innerschulische und außerschulische Partnerschaften« (QB V) eine »gute Praxis« bzw. »ausgezeichnete Praxis« (Grafik I 2.1 (G7) und Web-Tabelle I 2.1 (T7)). Die Merkmale *Kooperation mit außerschulischen Partnern* und *Einbeziehung in das schulische Leben* erhalten dabei eine besonders positive Einordnung.

I 2.1 (G4)

Entwicklungsstufen der allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Qualitätsmerkmal und Schultart: QB Q Qualitätssicherung und Qualitätsentwicklung



Datenquelle: Landesinstitut für Schulentwicklung

I 2.1.3 Vertiefende Analysen

Um Erkenntnisse aus der Fremdevaluation zu gewinnen, die über die Zuordnung in Entwicklungsstufen und die Darstellung von mittleren Bewertungen hinausgehen, wurden weiterreichende qualitative Analysen vorgenommen. Diese Auswertungen ermöglichen es, tieferegehende Informationen zur Qualität schulischer Prozesse und Maßnahmen zu gewinnen. Als Datenmaterial wurden die einzelnen Evaluationsberichte der allgemein bildenden Schulen herangezogen.

Ergänzend zur quantifizierenden Einordnung der Qualitätsmerkmale in drei Entwicklungsstufen enthalten die Berichte für jedes evaluierte Qualitätsmerkmal eine ausführliche Beschreibung der an der jeweiligen Schule vorgefundenen Situation. Sie dienen der Ausführung, Präzisierung und Begründung der resultierenden Einstufung. In diese Texte fließen die aus sämtlichen Verfahren und Datenquellen gewonnenen Erkenntnisse ein. Sie enthalten wichtige Hinweise zu den Maßnahmen der Schulen, die mit dem Ziel der Erfüllung dieses Qualitätsmerkmals erfolgreich praktiziert werden. Neben den Ausführungen zu den evaluierten

Merkmale enthalten die Berichte abschließende Empfehlungen für die Weiterentwicklung der Schule.

Bei der qualitativen Analyse kommen zwei unterschiedliche Verfahren zum Einsatz: Die Auswertung der im Evaluationsbericht vermerkten Empfehlungen an die Schule sowie die Analyse von Best Practice.

1. Analyse der Empfehlungen:

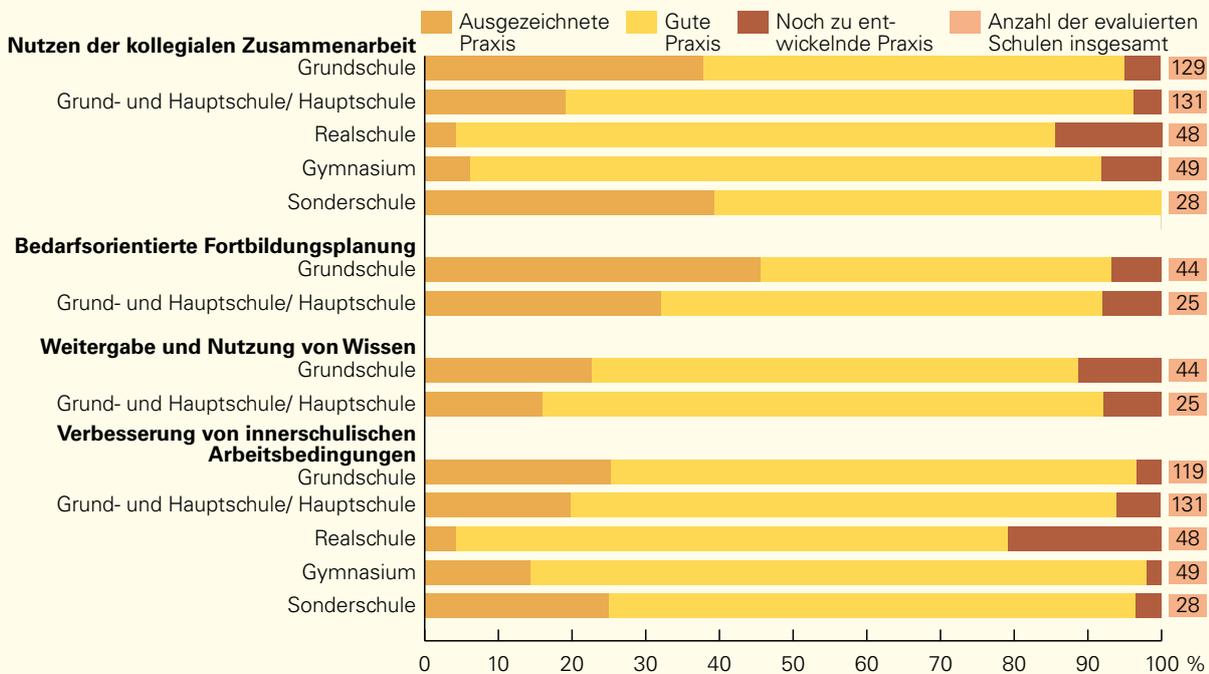
Die in den Berichten niedergelegten konkreten Empfehlungen der Evaluationsteams an die evaluierten Schulen zur Weiterentwicklung von Schule und Unterricht werden zunächst systematisiert und übergreifenden Kategorien zugeordnet. In einem zweiten Schritt wird ermittelt, zu welchen Anteilen die einzelnen Kategorien je Schulart genannt werden. Dies ergibt Hinweise auf die wichtigsten Entwicklungsbedarfe an den evaluierten Schulen.

2. Analyse von Best Practice:

An einzelnen Schulen vorgefundene vorbildhafte Maßnahmen und Aktivitäten für die Umsetzung der Qualitätsansprüche im Bereich Unterricht werden in einer inhaltsanalytischen Auswertung der Evaluationsberichte identifiziert und aufbereitet. Methodisch wird

I 2.1 (G5)

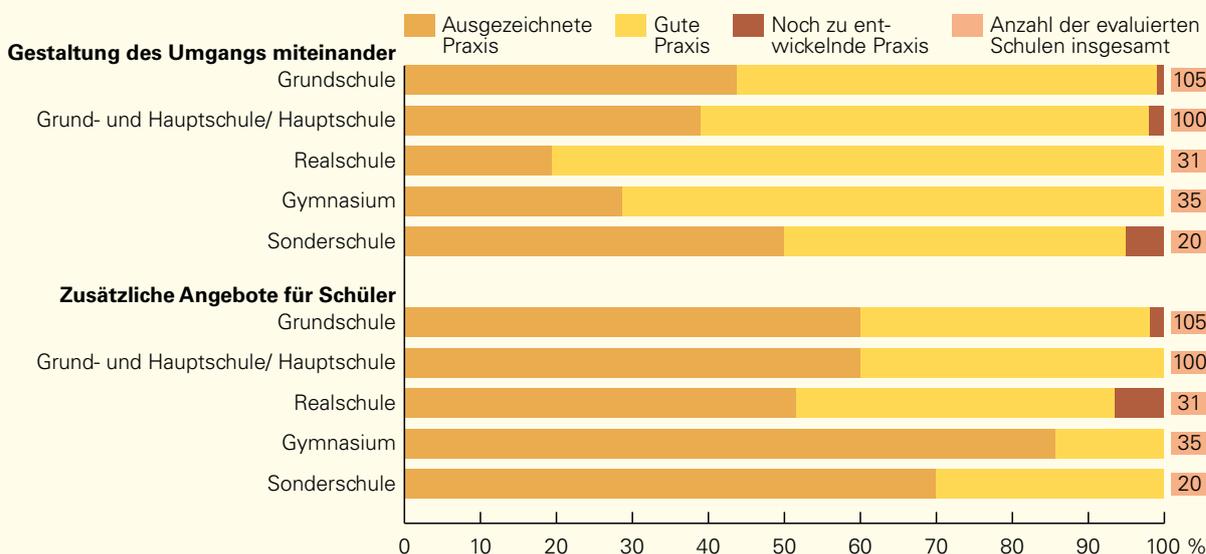
Entwicklungsstufen der allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Qualitätsmerkmal und Schulart: QB II Professionalität der Lehrkräfte



Anmerkung: Die grafische Darstellung berücksichtigt nur die jeweiligen Qualitätsmerkmale, die von mindestens zehn Schulen einer Schulart ausgewählt wurden.
Datenquelle: Landesinstitut für Schulentwicklung

I 2.1 (G6)

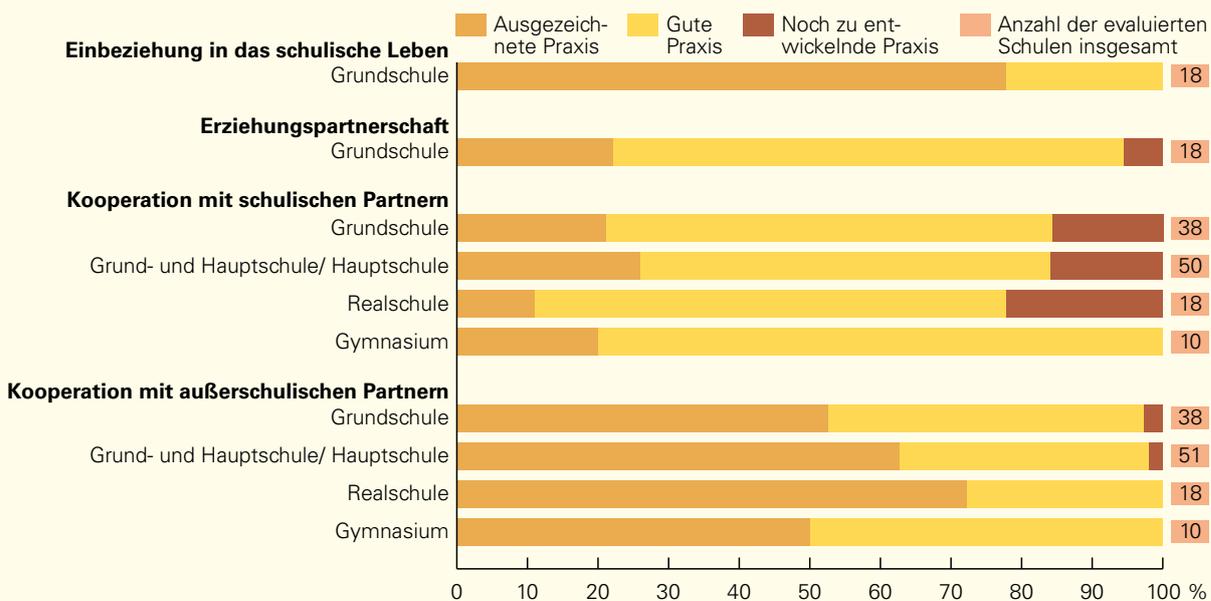
Entwicklungsstufen der allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Qualitätsmerkmal und Schulart: QB IV Schul- und Klassenklima



Anmerkung: Die grafische Darstellung berücksichtigt nur die jeweiligen Qualitätsmerkmale, die von mindestens zehn Schulen einer Schulart ausgewählt wurden.
 Datenquelle: Landesinstitut für Schulentwicklung

I 2.1 (G7)

Entwicklungsstufen der allgemein bildenden Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Qualitätsmerkmal und Schulart: QB V Inner- und außerschulische Partnerschaften



Anmerkung: Die grafische Darstellung berücksichtigt nur die jeweiligen Qualitätsmerkmale, die von mindestens zehn Schulen einer Schulart ausgewählt wurden.
 Datenquelle: Landesinstitut für Schulentwicklung

dabei auf folgende Weise vorgegangen: Zunächst werden Merkmale im Qualitätsbereich Unterricht herangezogen, bei denen die evaluierten Schulen bzw. einzelne Schularten im Mittel nur unterdurchschnittliche Bewertungen erzielten. Im nächsten Schritt erfolgt eine Identifizierung derjenigen Schulen, die bei diesen schwierig zu erfüllenden Qualitätsmerkmalen trotzdem eine »ausgezeichnete Praxis« erzielten. Diese Schulen haben offenbar Zugangsweisen entwickelt, die ungeachtet der in der Unterrichtspraxis anspruchsvoll umzusetzenden Anforderungen zu überdurchschnittlichen Erfolgen führten.

Ziel der qualitativen Untersuchung ist es, solche Praktiken zu identifizieren und in einem Kategoriensystem zu systematisieren. Dies erfolgt mittels einer Analyse der unterrichtlichen Aktivitäten der entsprechenden Schulen in Bezug auf das betrachtete Qualitätsmerkmal. Damit entsteht eine kategorisierte Darstellung an Best-Practice-Beispielen. Um abzuschätzen, in welchem Umfang bestimmte Aktivitäten positiv zur Schul- und Unterrichtsentwicklung beitragen, werden die Vorkommenshäufigkeiten der einzelnen Best-Practice-Kategorien ermittelt.

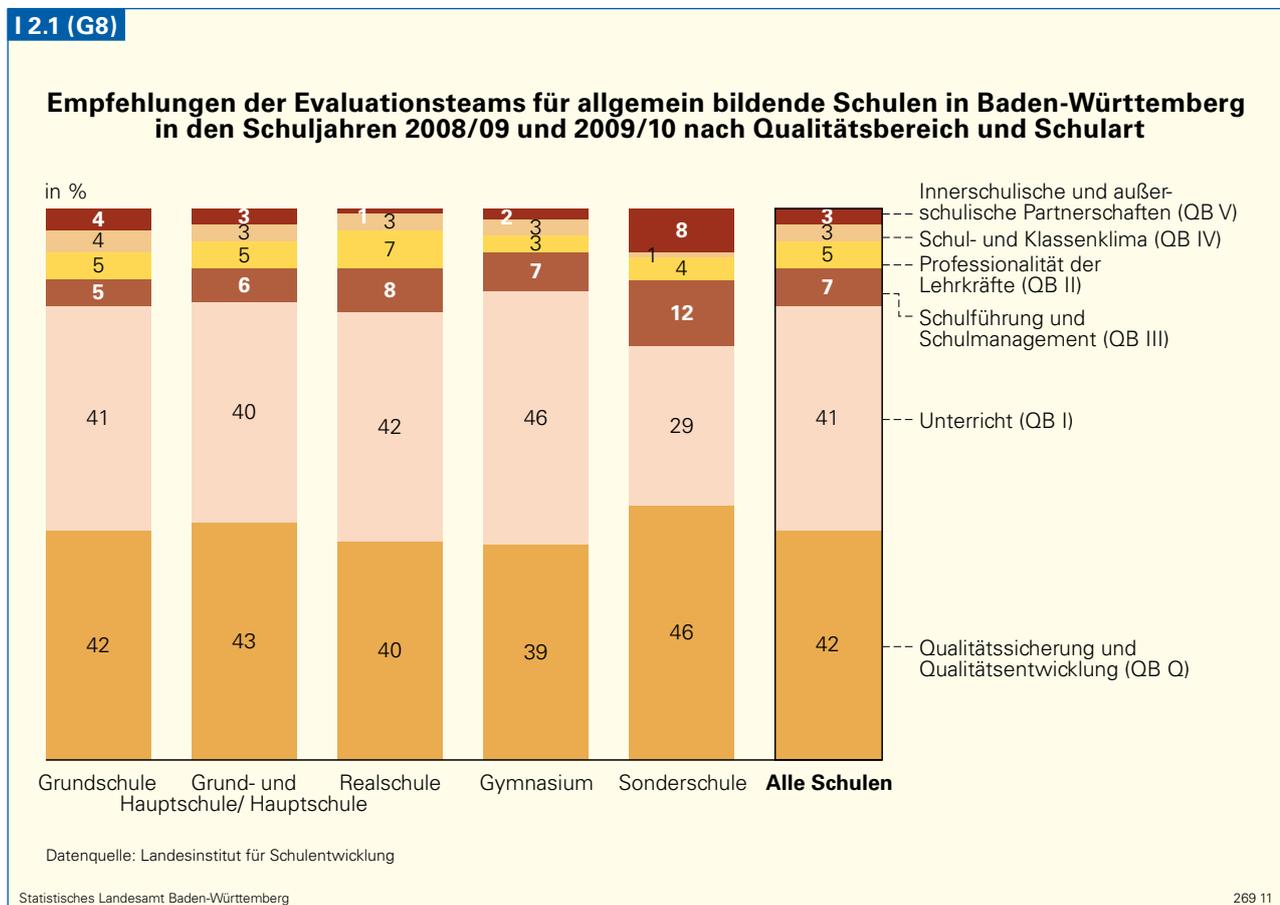
Die Analysen verdeutlichen, mit welchen Strategien und Maßnahmen es Schulen gelingen kann, in an-

spruchsvollen Bereichen der Unterrichtsentwicklung ein überdurchschnittliches Niveau zu erreichen.

Analyse der Empfehlungen

Die Evaluationsberichte führen am Ende spezifische Empfehlungen auf, die der Schule konkrete Anregungen und Vorschläge für ihre Weiterentwicklung anbieten. Die Empfehlungen beziehen sich dabei auf die einzelnen Bereiche und Merkmale des Qualitätsrahmens Fremdevaluation. Die Betrachtung der Art und Häufigkeit der vorgefundenen Empfehlung liefert Hinweise auf konkrete Entwicklungsbedarfe. Um feststellen zu können, ob schulartspezifische Unterschiede vorliegen, wurden entsprechend differenzierende Analysen vorgenommen. Hierfür wurden insgesamt 2 384 Empfehlungen an 476 evaluierten allgemein bildenden Schulen ausgewertet. Sechs Schulen sonstiger Schulverbände blieben hierbei unberücksichtigt (Tabelle I 2.1 (T8) im Anhang).

Die meisten Empfehlungen – mit Anteilen zwischen 39 % bei den Gymnasien und 46 % bei den Sonderschulen – beziehen sich auf den Bereich der »Qualitätssicherung und Qualitätsentwicklung« (QB Q) (Grafik I 2.1 (G8)). An zweiter Stelle rangieren mit Anteilen zwischen 29 %



und 46 % Empfehlungen, die auf die Unterrichtsentwicklung (QB I) abzielen. Lediglich weniger als ein Fünftel der Empfehlungen verweist auf die weiteren vier Qualitätsbereiche. Die geringe Zahl der Empfehlungen hinsichtlich der »Inner- und außerschulischen Partnerschaften« (QB V), des »Schul- und Klassenklimas« (QB IV) sowie der »Professionalität der Lehrkräfte« (QB II) erklärt sich allerdings auch dadurch, dass diese Qualitätsbereiche Wahlpflichtbereiche sind und entsprechend weniger diesbezügliche Rückmeldungen an die Schulen erfolgen können.

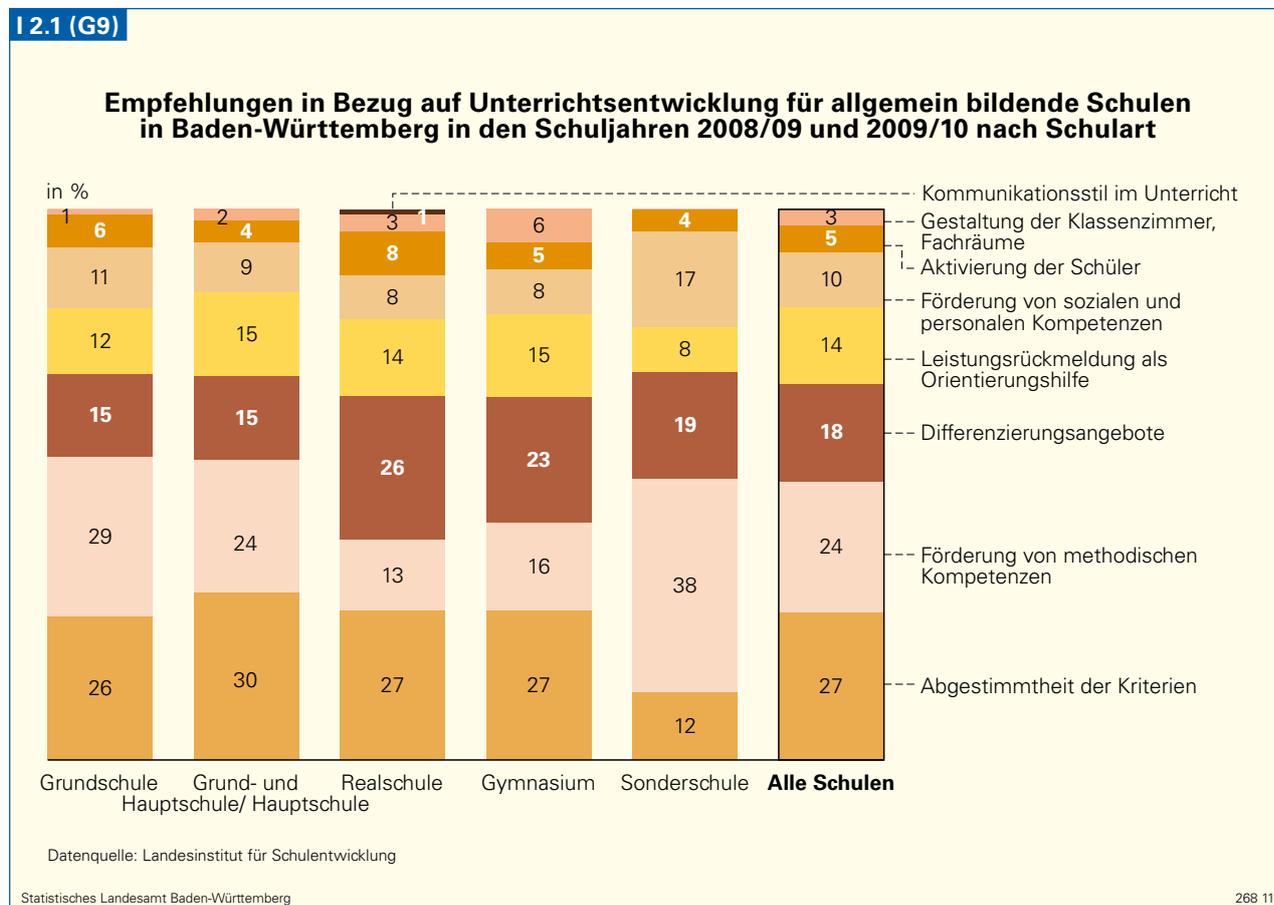
Eine nach Qualitätsmerkmalen differenzierte Betrachtung der Empfehlungen, die sich auf den Bereich »Unterricht« beziehen, weist Unterschiede zwischen den Schularten auf (Grafik I 2.1 (G9)).

Zwischen 26 % und 30 % der an Grundschulen, Grund- und Hauptschulen/ Hauptschulen, Realschulen und Gymnasien gerichteten Empfehlungen – und damit die häufigsten Nennungen – regen zu einer Optimierung bei der Abstimmung der Kriterien für die Leistungsrückmeldung an. An den Sonderschulen werden dagegen lediglich 12 % der Empfehlungen zu diesem Merkmal ausgesprochen. Die an den Grundschulen häufigste Empfehlungskategorie verweist auf die Förderung methodischer Kompetenzen bzw. die Erstellung eines Methodencurriculums, die ebenfalls bei den Grund- und Hauptschulen/ Hauptschulen und

vor allem an den Sonderschulen (38 %) oft ausgesprochen wird. An den Gymnasien und insbesondere an den Realschulen spielen Empfehlungen dieses Typs eine weniger bedeutende Rolle. An diesen beiden Schularten sind Anregungen, die auf eine Weiterentwicklung der Differenzierungsangebote zielen, häufiger vertreten. Zudem ist dies bei den Grundschulen in etwas geringerem Umfang der Fall – obwohl diese Schulart die besten Evaluationsergebnisse im Hinblick auf dieses Qualitätsmerkmal erzielt.

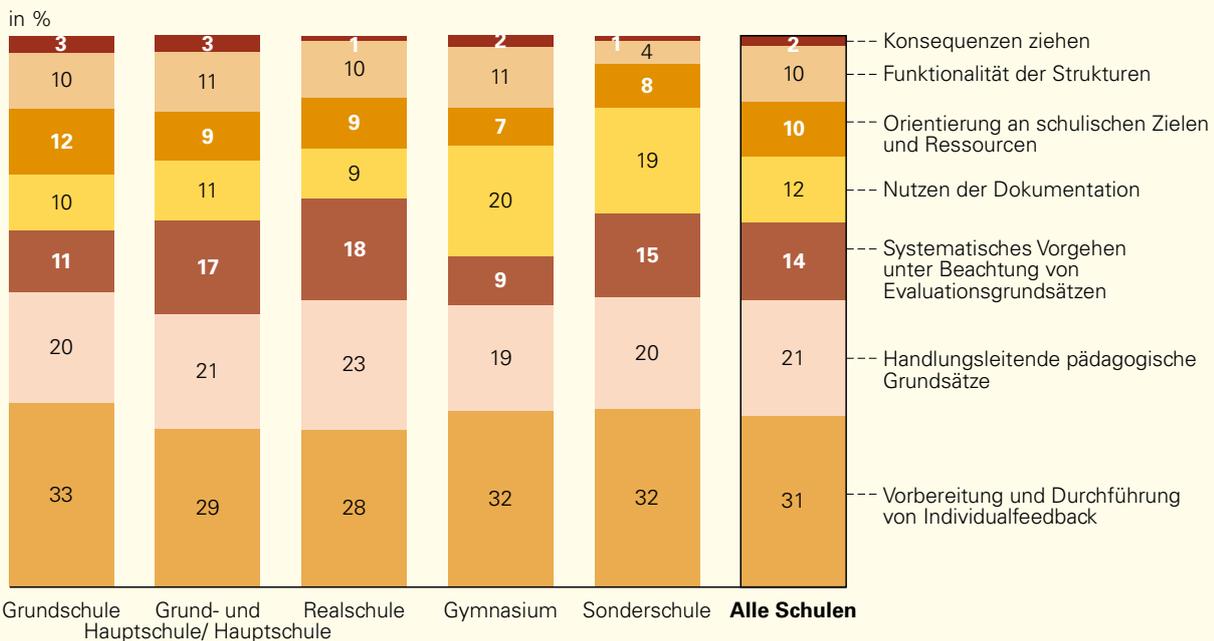
In Anteilen zwischen 8 % und 17 % der Nennungen richten sich Empfehlungen auf die Optimierung der Leistungsrückmeldung als Orientierungshilfe sowie die Förderung von sozialen und personalen Kompetenzen. Empfehlungen zur Aktivierung der Schülerinnen und Schüler sind lediglich bei den Realschulen mit 8 % noch etwas häufiger vertreten, Vorschläge in Bezug auf die Gestaltung der Klassenzimmer/ Fachräume mit 6 % in geringem Umfang bei den Gymnasien.

Rund ein Drittel aller Empfehlungen aus dem Bereich »Qualitätssicherung und Qualitätsentwicklung« bezieht sich – relativ gleichmäßig über die Schularten verteilt – auf die Vorbereitung und Durchführung von Individualfeedback (Grafik I 2.1 (G10)). Mit durchschnittlich 10 Prozentpunkten Abstand folgen Vorschläge zur Entwicklung handlungsleitender pädagogischer Grundsätze. Empfeh-



I 2.1 (G 10)

Empfehlungen in Bezug auf Qualitätssicherung und -entwicklung für allgemein bildende Schulen in Baden-Württemberg in den Schuljahren 2008/09 und 2009/10 nach Schulart



Datenquelle: Landesinstitut für Schulentwicklung

Statistisches Landesamt Baden-Württemberg

267 11

lungen zu einem *systematischen Vorgehen unter Beachtung von Evaluationsgrundsätzen* bei der Selbstevaluation richten sich mit Anteilen zwischen 15 % und 18 % an die Grund- und Hauptschulen/ Hauptschulen, Realschulen und Sonderschulen. Bei den Grundschulen und Gymnasien wird diese Empfehlung im Vergleich dazu seltener ausgesprochen (11 % bzw. 9 %). An den Gymnasien und den Sonderschulen werden, in größerem Umfang als bei anderen Schularten, häufiger Empfehlungen zu einer besseren *Nutzung der Dokumentation des schulischen Qualitätsmanagements* formuliert (20 % bzw. 19 %).

Empfehlungen in Bezug auf die *Orientierung an schulischen Zielen und Ressourcen* spielen an allen Schularten mit Anteilen von meist unter 10 % eine relativ geringfügige Rolle. Mit Ausnahme der Sonderschulen zielt etwa jede zehnte Empfehlung auf die *Funktionalität der Strukturen* der Qualitätssicherung und Qualitätsentwicklung. Das *Ziehen von Konsequenzen* aus durchgeführten Selbstevaluationen wird nur vereinzelt genannt.

Analyse von Best Practice

Die im Qualitätsbereich »Unterricht« bei der Evaluation erzielten Ergebnisse sind sehr unterschiedlich (Grafik I 2.1 (G2)). Entwicklungsbedarf besteht vor

allem hinsichtlich derjenigen Qualitätsmerkmale, die sich auf den Prozess der Leistungsrückmeldung beziehen und – bei den Realschulen und Gymnasien – auf die im Unterricht genutzten Möglichkeiten der Individualisierung und Differenzierung. Die in Bezug auf diese drei Merkmale vorgefundenen ausgezeichneten Praktiken werden im Folgenden analysiert.

Gestaltung der Lehr-Lern-Prozesse: Merkmal Differenzierungsangebote

Von allen Schularten wird erwartet, dass der Unterricht Differenzierungsangebote enthält, um den individuellen Voraussetzungen und Fähigkeiten der Schüler gerecht werden zu können. Für den Erwerb fachlicher und überfachlicher Kompetenzen erweisen sich insbesondere Lernarrangements günstig, die ein selbstgesteuertes Arbeiten ermöglichen.¹³

Die Ausprägung des Merkmals *Differenzierungsangebote* ist nach Schularten unterschiedlich. Während die Grundschulen und Sonderschulen im Mittel eine über

¹³ Vgl. Qualitätsrahmen zur Fremdevaluation, Landesinstitut für Schulentwicklung (2009), S. 26.

die »gute Praxis« hinausgehende Einordnung erzielen, liegen die durchschnittlichen Einstufungen der Realschulen und Gymnasien unterhalb dieser Entwicklungsstufe (Grafik I 2.1 (G2)). Insgesamt erreichen rund 17 % bzw. 81 von 476 evaluierten allgemein bildenden Schulen – darunter zwei Gymnasien – eine differenzierende Gestaltung der Lehr-Lern-Prozesse auf dem Niveau der »ausgezeichneten Praxis«.

Als zentrale Aktivitäten für die herausragende Umsetzung von Differenzierungsmaßnahmen im Unterricht erweisen sich das Bereitstellen *individualisierter Arbeitsmaterialien* und *Hilfsmittel* entsprechend dem Leistungsstand der Schüler sowie die Implementierung förderlicher *schulorganisatorischer Maßnahmen* wie beispielsweise Bildung von Lerngruppen oder Unterrichten im Tandem (Grafik I 2.1 (G11)). Mehr als ein Drittel aller Aktivitäten können diesen beiden Kategorien zugeordnet werden. Ein weiteres Drittel an Maßnahmen umfasst *individuelle Unterstützungsangebote*

(zum Beispiel Förderangebote für Kinder mit Lese-/ Rechtschreibschwäche oder Sprachhilfeangebote für Kinder mit Migrationshintergrund), *spezifische Unterrichtsarrangements* (zum Beispiel Einsatz von Wochen- oder Tagesplänen) und das *Ermöglichen von unterschiedlichen Lern tempi* (zum Beispiel Einrichten von Lerntheken mit unterschiedlichen Arbeits- und Übungsmaterialien). Dass erfolgreiche Differenzierungsangebote individueller Lösungen bedürfen, verdeutlicht die relativ große Kategorie *Sonstige Aktivitäten*.

Praxis der Leistungsbeurteilung und Leistungsrückmeldung: Merkmal *Leistungsrückmeldung als Orientierungshilfe*

Leistungsrückmeldungen dienen nicht nur dem reinen Nachweis der von einzelnen Schülern erbrachten Leistungen. Darüber hinaus sollen sie gleichermaßen den Schülern, Eltern und Lehrkräften eine wichtige Orientierungshilfe für die Einordnung des Lernstands und des Lernfortschritts bieten. Schüler sollen systematisch dahin geführt werden, eigene Leistungen selbstständig und kompetent einschätzen zu können.¹⁴ Die mittleren Bewertungen dieses Qualitätsmerkmals liegen für die Realschulen und die Gymnasien unterhalb der Stufe »gute Praxis« (Grafik I 2.1 (G2)). Lediglich die Grundschulen und die Sonderschulen erreichen hier eine »gute Praxis«. Insgesamt wurde das Merkmal *Leistungsrückmeldung als Orientierungshilfe* an rund 5 % bzw. 25 von 476 evaluierten allgemein bildenden Schulen – zumeist Grundschulen – mit einer »ausgezeichneten Praxis« bewertet.

Die qualitative Analyse verdeutlicht, dass Maßnahmen, die es den Schülern ermöglichen *Lernimpulse aus der Leistungsrückmeldung* – zum Beispiel mittels einer kriteriengeleiteten Rückmeldung – zu entnehmen, die *Bewertungskriterien den Schülern* bzw. den *Eltern bekanntzumachen* und den Schülern *Selbsteinschätzungen zu ermöglichen* – zum Beispiel mithilfe von Selbsteinschätzungsbögen oder mittels Einschätzung durch die Klassengemeinschaft bei Präsentationen – einen zentralen Beitrag für diese anspruchsvolle Funktion der Leistungsrückmeldung leisten (Grafik I 2.1 (G13)). An den betreffenden Schulen wurden zu jeder dieser Kategorien spezifische Praktiken konzipiert und erfolgreich umgesetzt (ein Beispiel zeigt Grafik I 2.1 (G12)). Dabei verfolgt jede in diesem Aspekt als »ausgezeichnet« bewertete Schule mehrere für sie passende Maßnahmen, die parallel eingesetzt werden. Diese vier Praktiken umfassen insgesamt zwei Drittel aller durchgeführten Maßnahmen, die zu einer ausge-



¹⁴ Vgl. Qualitätsrahmen zur Fremdevaluation, Landesinstitut für Schulentwicklung (2009), S. 29.

I 2.1 (G12)

Beispiel für eine Leistungsrückmeldung im Fach Deutsch am Robert-Bosch-Gymnasium in Wendlingen

RBG Wendlingen Deutsch (Lh) Klasse

NAME:

1. Klassenarbeit (28.10.10): TEXTINTERPRETATION

„Hauptsache weit“ von Sybille Berg

Einleitung

Die Einleitung gibt Auskunft über Titel, Autor, Textsorte und Entstehungszeit des zu interpretierenden Textes.

Die Textsorte fehlt.

In der Einleitung wurde eine erste Deutungshypothese formuliert.

Deine Deutungshypothese fängt gut an, aber du müsstest einen Schritt weiter gehen: Sein Traum ist nicht realistisch und zerplatzt während der Reise.

Die Kurzgeschichte wurde kurz zusammengefasst.

Nein.

Interpretation

Die wesentlichen inhaltlichen Gesichtspunkte wurden verstanden und aus dem Text herausgearbeitet.

Du hast die Geschichte ganz gut verstanden. Schade, dass du nicht näher auf das Pensionszimmer eingehst und auch nicht erklärst, warum er seine Freunde in den E-Mails belügt.

Inhalt, Sprache und Form wurden richtig analysiert (→ Vorarbeit).

Du hast einige Aspekte zur Sprache und Form herausgefunden, ...

In der Interpretation wurden Form, Sprache und Inhalt gedeutet und aufeinander bezogen (→ Vorarbeit).

... die du sinnvoll deuten konntest.

Die Überschrift wurde mit in die Interpretation einbezogen (→ Vorarbeit).

Nein!

Schluss

Du hast in deinem Schluss Stellung zum Text und zum übergeordneten Thema genommen.

In Ordnung.

Aufbau und Form

Der Aufbau der Interpretation ist klar (→ Schreibplan).

Du hast die Ergebnisse deiner Analyse sinnvoll in Absätze gegliedert, die Reihenfolge ist jedoch nicht immer logisch.

Aussagen wurden am Text belegt, Zitate sinnvoll angewendet und gegebenenfalls erläutert.

Du hast die eingefügten Zitate nicht erläutert.

Einzelaspekte wurden miteinander verknüpft.

Zu wenig.

Sprache

Deine Ausdrucksweise ist genau und korrekt, du hast Hauptsätze und Satzgefüge verwendet.

Du drückst dich nicht immer korrekt aus.

Die Darstellung der Aussagen ist gedanklich klar aufgebaut.

In Ordnung.

Rechtschreibung und Zeichensetzung sind korrekt

Ein paar Fehler.

Äußere Form

Der Aufsatz wurde sinnvoll in Abschnitte gegliedert und leserlich/leserfreundlich geschrieben.

In Ordnung.

NOTE: 3+

I 2.1 (G 13)

Von Schulen mit ausgezeichneter Praxis* in Baden-Württemberg umgesetzte Maßnahmen zum Merkmal *Leistungsrück- meldung als Orientierungshilfe* in den Schuljahren 2008/09 und 2009/10



*) Allgemein bildende Schulen, die in den Schuljahren 2008/09 und 2009/10 an einer Fremdevaluation teilgenommen haben.
Lesehilfe: Von insgesamt 141 Maßnahmen zum Merkmal *Leistungsrückmeldung als Orientierungshilfe* an 25 Schulen mit >>ausgezeichneter Praxis<< wurden 13 % der Maßnahme *Hinweise für Eltern* zugeordnet.

Datenquelle: Landesinstitut für Schulentwicklung.

zeichneten Praxis dieses Qualitätsmerkmals führten. Weniger häufig wurden die Kategorien, bei denen die Leistungsrückmeldung eine Orientierungshilfe anhand von *Hinweisen für Eltern* liefert, identifiziert. Ebenfalls seltener wurden systematische *Verfahren zur Qualitätssicherung und -entwicklung* der Leistungsrückmeldung – zum Beispiel das Führen von Lernentwicklungsheften mit Kompetenzbögen durch die Lehrkräfte- und *Reflexion der Lernergebnisse/-prozesse* im Unterricht sowie *Sonstige Maßnahmen* durchgeführt.

Praxis der Leistungsbeurteilung und Leistungsrückmeldung: Merkmal *Abgestimmtheit der Kriterien*

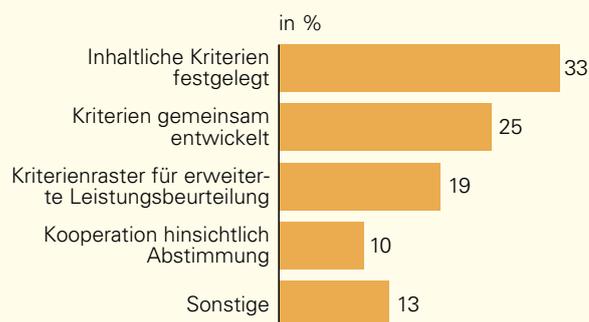
Leistungsbeurteilungen und Leistungsrückmeldungen liefern Schülern, Eltern und Lehrkräften eine notwendige Rückmeldung für die Beurteilung des momentan erreichten Lernstands und Lernfortschritts. Sie bilden eine wichtige Lern- und Orientierungshilfe, sofern ihr Zustandekommen objektiv, kriteriengestützt und möglichst transparent erfolgt. Damit diese lernförderliche Funktion nicht nur in einzelnen Klassen, sondern innerhalb der gesamten Schule wirksam werden kann, ist es erforderlich, die Beurteilungskriterien

kooperativ innerhalb des Kollegiums zu entwickeln, abzustimmen und systematisch zu implementieren.¹⁵

Von allen evaluierten Aspekten im Bereich *Unterricht* ist das Qualitätsmerkmal *Abgestimmtheit der Kriterien* dasjenige mit dem größten Entwicklungsbedarf. Die gemittelten Werte befinden sich über alle Schularten hinweg deutlich unterhalb der »guten Praxis« (Grafik I 2.1 (G2)). Mit Ausnahme der Sonderschulen wird bei über der Hälfte der evaluierten Schulen eine »noch zu entwickelnde Praxis« konstatiert. Rund 5 % bzw. 24 von 476 evaluierten allgemein bildenden Schulen etablierten beispielgebende Aktivitäten, die zu einer »ausgezeichneten Praxis« führten. Jede dieser Schulen definiert und dokumentiert *inhaltliche Kriterien*, an denen sich die Leistungsbeurteilungen orientieren müssen – dieser Aspekt bildet mit einem Anteil von 33 % somit die häufigste Aktivität (Grafik I 2.1 (G14)). An den meisten Schulen vollzog sich dieser Prozess kooperativ innerhalb des Kollegiums bzw. der Fachschaft. Viele dieser Schulen nutzten auch die lernprozessförderlichen Möglichkeiten eines *Kriterienrasters für erweiterte Leistungsbeurteilungen*, zum Beispiel in Bezug auf inhaltliche Anforderungen um mündliche oder schriftliche Leistungen. 13 % der Aktivitäten sind *Sonstige Maßnahmen*, die sich aus den individuellen Begebenheiten der einzelnen Schulen heraus entwickelt haben.

I 2.1 (G14)

Von Schulen mit ausgezeichneter Praxis* in Baden-Württemberg umgesetzte Maßnahmen zum Merkmal *Abgestimmtheit der Kriterien* in den Schuljahren 2008/09 und 2009/10



*) Allgemein bildende Schulen, die in den Schuljahren 2008/09 und 2009/10 an einer Fremdevaluation teilgenommen haben.
Lesehilfe: Von insgesamt 84 Maßnahmen zum Merkmal *Abgestimmtheit der Kriterien* an 24 Schulen mit >>ausgezeichneter Praxis<< wurden 25 % der Maßnahme *Kriterien gemeinsam entwickelt* zugeordnet.

Datenquelle: Landesinstitut für Schulentwicklung.

Statistisches Landesamt Baden-Württemberg

262 11

15 Vgl. Qualitätsrahmen zur Fremdevaluation, Landesinstitut für Schulentwicklung (2009), S. 29.

I 2.2 Fremdevaluation an beruflichen Schulen

I 2.2.1 Grundlagen

Die Fremdevaluation an beruflichen Schulen wurde im Rahmen der Konzeption *Operativ Eigenständige Schule (OES)*¹⁶ auf der Grundlage des in der Schweiz entwickelten Verfahrens der externen Evaluation nach *Q2E – Qualität durch Evaluation und Entwicklung*¹⁷ eingerichtet und an das Qualitätssystem der beruflichen Schulen Baden-Württembergs angepasst. Dabei wird das schulische Qualitätsmanagement betrachtet: Nicht einzelne Projekte und Maßnahmen auf der operativen Ebene stehen im Mittelpunkt, sondern die von der Schule auf Ebene der Steuerung und Organisation vorgenommene systemische Verankerung der Qualitätsentwicklung entsprechend den im Verfahren Q2E formulierten Bereichen und Standards.

Das Evaluationsteam betrachtet damit die Ergebnisse der schulinternen Selbstevaluationen und die daraus abgeleiteten Maßnahmen dahingehend, ob diese leitbildkonform sind, ob die Qualitätsaktivitäten zielgerichtet gesteuert werden und ob Akzeptanz bei den Beteiligten hergestellt werden konnte.

Die Kriterien für die Betrachtung der systematischen Qualitätsentwicklung beziehen sich auf die vier Bereiche

- Qualitätsentwicklung im Bereich Unterricht,
- Schulleitung und Steuerung der Qualitätsprozesse,
- Individualfeedback und persönliche Qualitätsentwicklung,
- Selbstevaluation und Qualitätsentwicklung der Schule.

Die Evaluation nimmt diese Bereiche dabei unter den Aspekten

- Durchführung,
- Wirkung und Wirksamkeit,
- organisatorische Einbindung

in den Blick und beschreibt vor diesem Hintergrund die an der Schule vorgefundenen Aktivitäten und Maßnahmen.

¹⁶ Vgl. Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2010a).

¹⁷ Qualitätsmanagementmodell *Qualität durch Evaluation und Entwicklung (Q2E)*; vgl. Landwehr, N./ Steiner, P. (2007a).

Jede Fremdevaluation erfolgt nach einem standardisierten Vorgehen, in dem verschiedene qualitative Methoden der Datenerhebung zum Einsatz kommen. Schüler, Lehrkräfte und Schulleitung werden in die Evaluation einbezogen. Der Ablauf gliedert sich in folgende Schritte:

- Vorgespräch der Teamleitung mit der Schule,
- Analyse der von der Schule übersandten Unterlagen zur schulischen Qualitätsentwicklung, des sogenannten Qualitätshandbuchs,
- Evaluationsplanung: Hypothesen, Gruppen, Instrumente,
- Planungsgespräch an der Schule,
- Organisatorische Vorbereitung der Fremdevaluation durch die Schule,
- Datenerhebung und -dokumentation an der Schule im Rahmen von Interviews und Ratingkonferenzen mit Schulleitung, Lehrkräften und Schülern sowie einem Schulhausrundgang,
- Auswertung der Daten,
- Erstellung des vorläufigen Evaluationsberichts,
- Ergebnispräsentation in der Gesamtlehrerkonferenz,
- Fertigstellung und Versand des Evaluationsberichts an die Schule.

Der von der Fremdevaluation festgestellte Entwicklungsstand wird in vier verschiedenen, qualitativ definierten Entwicklungsstufen abgebildet: »Die Praxis ist wenig entwickelt« (Stufe 1), »die Praxis entspricht grundlegenden Anforderungen« (Stufe 2), »die Praxis ist auf gutem Niveau entwickelt« (Stufe 3) und »die Praxis ist exzellent« (Stufe 4). Stufe 2 entspricht dabei einer für den schulischen Alltag grundlegend funktionsfähigen Praxis und Stufe 3 einer für alle beruflichen Schulen angestrebten Zielstufe. Stufe 4 hingegen verweist auf ein darüber hinausreichendes Exzellenzniveau, dessen Erreichen einen mehrjährigen Entwicklungsprozess im betreffenden Qualitätsbereich voraussetzt.

Die Schule leitet den Evaluationsbericht an den Schulträger sowie an die zuständige Schulaufsichtsbehörde weiter. Letztere trifft auf Grundlage des Berichts und weiterer Ziele der Schule und Vorgaben des Landes Zielvereinbarungen mit der Schule. Im Unterschied zu den allgemein bildenden Schulen findet im beruflichen

Bereich keine repräsentative Zufallsauswahl der zu evaluierenden Schulen statt. Die beruflichen Schulen können, gesteuert durch die Regierungspräsidien im

Rahmen der OES-Konzeption, innerhalb eines größeren Zeitraums ein Schuljahr ihrer Wahl für die Fremdevaluation festlegen.

Das Konzept der Fremdevaluation an beruflichen Schulen

Grundlage der Qualitätsentwicklung von Schule und Unterricht an den beruflichen Schulen Baden-Württembergs bildet das Konzept *Operativ eigenständige Schule (OES)*, welches im Handbuch OES¹ und den darin enthaltenen 14 Handreichungen umfassend dargestellt wird. Als Grundlage diente das Qualitätsmanagementsystem *Q2E – Qualität durch Evaluation und Entwicklung*² aus der Nordwestschweiz. Das Konzept OES wurde in den Jahren 2003 bis 2006 in Zusammenarbeit von Kultusministerium und 18 beruflichen Schulen entwickelt und erprobt. Am Ende des Modellvorhabens durchlief das Verfahren eine Evaluation. Jeweils etwa ein Viertel pro Schuljahr der rund 300 öffentlichen beruflichen Schulen im Land begann in den Schuljahren 2007/08 bis 2010/11 mit OES zu arbeiten.

Die Fremdevaluation durch das Landesinstitut für Schulentwicklung (LS) ist das externe Element der systematischen Qualitätsentwicklung von Schulen. Ein Evaluationsteam des LS betrachtet dabei auf der Grundlage des schulischen Leitbildes, der Selbstevaluation, der gegebenenfalls mit der Schulaufsicht vereinbarten Ziele und der schulischen Dokumentation die Prozesse, Verfahren, Strukturen und Methoden zur systematischen Qualitätsentwicklung der beruflichen Schulen. Hierbei gibt es zwei Betrachtungsfelder: Zum einen werden die Charakteristika der Schule, ihre Stärken, Schwächen und Besonderheiten erfasst, zum anderen bildet die systematische Qualitätsentwicklung den inhaltlichen Schwerpunkt der Fremdevaluation.

Für die Betrachtung der systematischen Qualitätsentwicklung der beruflichen Schulen werden Kriterien in den Bereichen

- Qualitätsentwicklung im Bereich Unterricht,
- Schulleitung und Steuerung der Qualitätsprozesse,
- Individualfeedback und persönliche Qualitätsentwicklung und
- Selbstevaluation und Qualitätsentwicklung der Schule evaluiert.

Die Einordnung erfolgt dabei in vier verschiedenen Qualitätsentwicklungsstufen:

- Entwicklungsstufe 1: Die Praxis ist wenig entwickelt,
- Entwicklungsstufe 2: Die Praxis entspricht den grundlegenden Anforderungen,
- Entwicklungsstufe 3: Die Praxis ist auf gutem Niveau entwickelt und
- Entwicklungsstufe 4: Die Praxis ist exzellent.

1 Vgl. Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2010b).

2 Vgl. Landwehr, N./ Steiner, P. (2007a).

Beispiel: Kriterium *Ausführungsqualität des Qualitätsbereichs Schulleitung und Steuerung der Qualitätsprozesse*

Entwicklungsstufe 1: <i>Die Praxis ist wenig entwickelt.</i>	Entwicklungsstufe 2: <i>Die Praxis entspricht grundlegenden Anforderungen.</i>	<i>Entwicklungsstufe 3: Die Praxis ist auf gutem Niveau entwickelt.</i>	Entwicklungsstufe 4: <i>Die Praxis ist exzellent.</i>
Ein Qualitätsmanagement ist nicht vorhanden. Einzelne Maßnahmen oder Elemente zur systematischen Qualitätssicherung oder -entwicklung (zum Beispiel Mitarbeitergespräch, Lehrerfortbildungsmaßnahmen usw.) sind eingeführt. Sie stehen unverbunden nebeneinander und wirken eher zufällig.	Qualitätsmanagement ist in Ansätzen vorhanden. Es dient der Förderung der Schul- und Unterrichtsqualität sowie der einzelnen Personen.	An der Schule wird ein funktionsfähiges und wirksames Qualitätsmanagement aufgebaut und umgesetzt. Qualitätsmanagement ist Schulführungsaufgabe.	Zusätzlich zu den Kriterien der Entwicklungsstufe 3: Die Schule verfügt über ein professionelles Qualitätsmanagement, das kreative Lösungen für schulspezifische Anliegen erlaubt.

Datenquelle: Landesinstitut für Schulentwicklung 2010, S. 15

Die Kriterien des Q2E-Konzepts wurden an das Qualitätsentwicklungssystem OES der beruflichen Schulen Baden-Württembergs angepasst.³ Sie werden innerhalb der Qualitätsbereiche jeweils unter den Aspekten

- Durchführung,
- Wirkung und Wirksamkeit sowie
- organisatorische Einbindung

betrachtet. Die Kriterien können außerdem verschiedenen Themenfeldern zugeordnet werden wie zum Beispiel Ausführungsqualität, Zielerreichung, Transparenz und Legitimierung.⁴ Jedes Kriterium weist eine genaue inhaltliche Beschreibung auf.

Das folgende Schaubild verdeutlicht die Gliederung in Qualitätsbereiche und Aspekte mit den entsprechenden Kriterien.

Aspekt \ Bereich	Schulleitung und Steuerung der Qualitätsprozesse	Individualfeedback und persönliche Qualitätsentwicklung	Selbstevaluation und Qualitätsentwicklung der Schule	Qualitätsentwicklung im Bereich Unterricht
Durchführung	Der Aspekt der Durchführung wird unter anderem am Kriterium der Ausführungsqualität beschrieben			
Wirkung und Wirksamkeit	Der Aspekt der Wirkung und Wirksamkeit wird unter anderem am Kriterium der Zielerreichung beschrieben.			
Organisatorische Einbindung	Der Aspekt der organisatorischen Einbindung wird unter anderem am Kriterium der Transparenz und Legitimierung beschrieben.			

3 Vgl. Landesinstitut für Schulentwicklung (2010), S. 12ff.

4 Vgl. Landwehr, N./ Steiner, P. (2007b), S. 38f.

Wie erfolgt die Zuordnung eines Merkmals zu einer der vier Entwicklungsstufen?

Beim Evaluationsbesuch befragt das Evaluationsteam alle an der Qualitätsentwicklung der Schule beteiligten Personengruppen. Dabei werden die vier Qualitätsbereiche hinsichtlich *Durchführung, Wirkung und Wirksamkeit* sowie *organisatorischer Einbindung* der entsprechenden Prozesse und Projekte betrachtet. Verwendet werden hierbei die beiden Instrumente Interview (mit Schülern und Lehrkräften in getrennten Gruppen) und Ratingkonferenz⁵ (mit Schulleitung und Qualitätssteuergruppe). Diese Daten werden zusammen mit der vorab durchgeführten Analyse der Qualitätsdokumentation der Schule (zum Beispiel Qualitätshandbuch) und des Schulhausrundgangs verarbeitet und führen zu einer Einschätzung des aktuellen Standes der schulischen Qualitätsentwicklung durch das Evaluationsteam (der sogenannten Fremdeinschätzung). Für diesen Zweck liegt ein Einschätzungsbogen vor, welchen die Schulen auch zur schulischen Selbsteinschätzung nutzen. Dieser Bogen beschreibt die 51 Kriterien auf dem Niveau der Entwicklungsstufe 3 »Die Praxis ist auf gutem Niveau entwickelt«. Das Evaluationsteam schätzt auf einer vierstufigen Skala⁷ ein, in welchem Umfang die dort gemachten Ausführungen zutreffen.

Auf Grundlage der sogenannten Fremdeinschätzung stellt das Evaluationsteam fest, welcher Entwicklungsstufe die Qualitätsentwicklung in den verschiedenen Bereichen jeweils entspricht (vgl. Grafik I 2.2 (G1)). Die Mittelwerte der Fremdeinschätzung sind in den Grafiken I 2.2 (G2) bis I 2.2 (G5) dargestellt.

Die Aussagen und Ergebnisse in den Betrachtungsfeldern »Charakteristika der Schule« und »systematische Qualitätsentwicklung« bilden die Grundlage für Empfehlungen zur weiteren Entwicklung der Schule.

5 In einer Ratingkonferenz erhalten die Teilnehmer zu Beginn einen Fragebogen mit vorgegebenen Aussagen oder Fragen, die sie zuerst individuell anhand einer vorgegebenen Skala einschätzen (sogenanntes Rating). Anschließend werden die individuellen Bewertungen zusammengetragen und das Gesamtergebnis in der Gruppe besprochen und von dieser interpretiert (sogenannte kommunikative Validierung). Weitere Nachfragen vonseiten des Evaluationsteams sind hierbei möglich.

6 Vgl. Landesinstitut für Schulentwicklung (2010).

7 Die vierstufige Skala umfasst die Aussagen »trifft (fast) nie zu«, »trifft eher weniger zu«, »trifft überwiegend zu« und »trifft (fast) immer zu«.

I 2.2.2 Ergebnisse

Die Auswertung der Evaluationsergebnisse der beruflichen Schulen bezieht sich in Anlehnung an das Verfahren der externen Evaluation nach Q2E auf die vier Bereiche *Qualitätsentwicklung im Bereich Unterricht, Schulleitung und Steuerung der Qualitätsprozesse, Individualfeedback und persönliche Qualitätsentwicklung* sowie *Selbstevaluation und Qualitätsentwicklung der Schule*, die jeweils anhand von Kriterien definiert sind. Die Qualitätsbereiche werden unter den drei Aspekten *Durchführung, Wirkung und Wirksamkeit* sowie *organisatorische Einbindung* evaluiert.¹⁸ In den Schuljahren 2007/08 bis 2009/10 wurden 40 berufliche Schulen evaluiert. Diese bilden die Datenbasis für die folgenden Auswertungen. Bei den vor 2007/08 evaluierten Schulen handelt es sich um Modellschulen, deren Ergebnisse nicht berücksichtigt wurden, da die

Fremdevaluation unter anderen Bedingungen stattfand.

Häufigkeitsverteilungen der Einordnungen

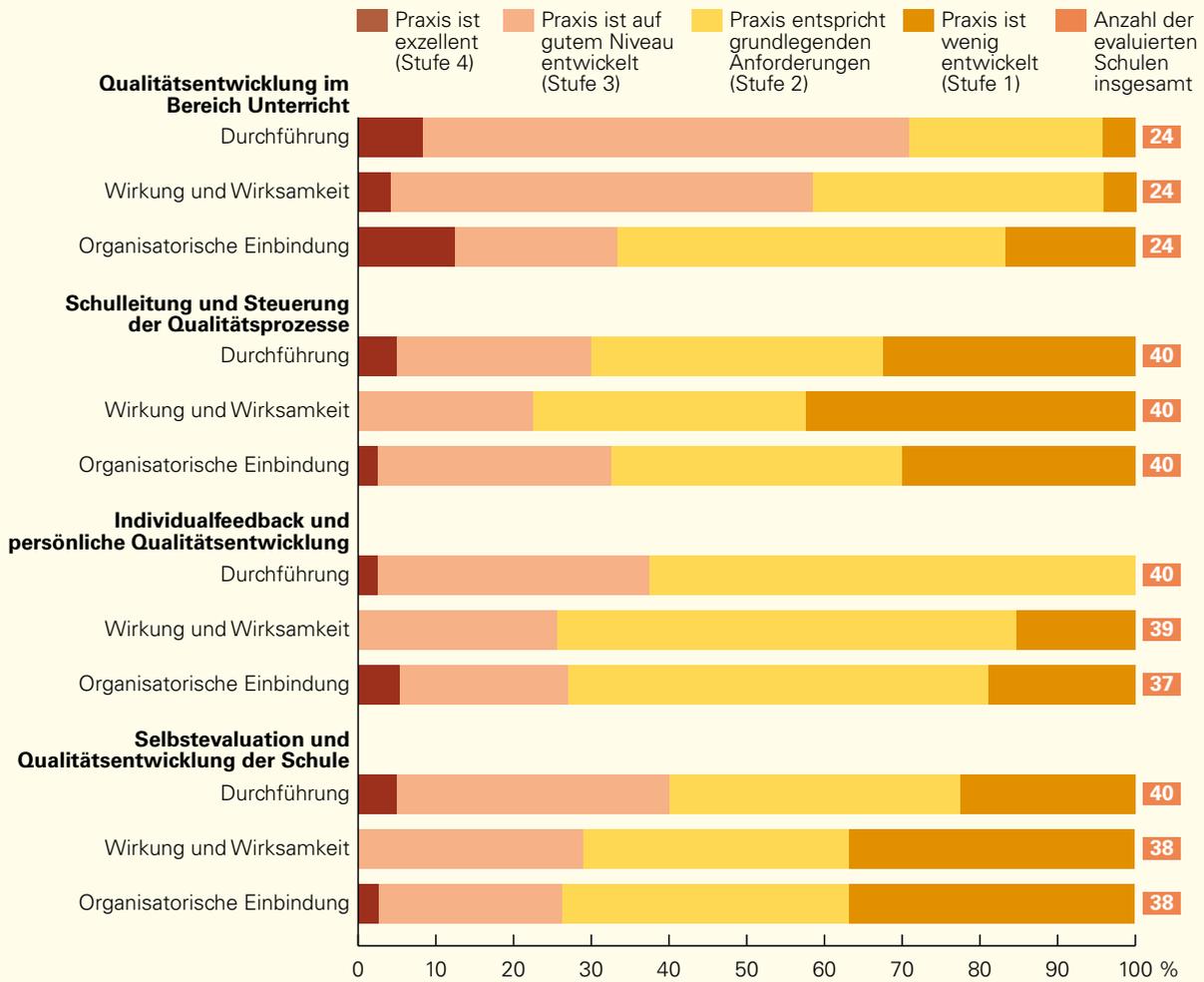
Um Unterschiede der Ergebnisse im Detail nachvollziehen zu können, werden die vergebenen Entwicklungsstufen als Häufigkeitsverteilungen in Form eines Balkendiagramms dargestellt (Grafik I 2.2 (G1) und Web-Tabelle I 2.2 (T1)). Die Bewertungsunterschiede zwischen den Qualitätsbereichen und den jeweiligen Aspekten sind dabei in ihrer prozentualen Häufigkeit ersichtlich. Im Unterschied zu dem Verfahren bei den allgemein bildenden Schulen wird der von der Fremdevaluation festgestellte Entwicklungsstand nicht in drei, sondern in vier Stufen differenziert.

Für den zum Schuljahr 2009/10 eingeführten Qualitätsbereich *Qualitätsentwicklung im Bereich Unterricht* liegen erst für 24 Schulen Einschätzungen vor. Hin-

18 Vgl. Landesinstitut für Schulentwicklung (2010).

I 2.2 (G1)

Entwicklungsstufen der beruflichen Schulen in Baden-Württemberg in den Schuljahren 2007/08 bis 2009/10 nach Qualitätsbereichen und einzelnen Aspekten



Anmerkung: Die Qualitätsentwicklung im Bereich Unterricht wurde als Qualitätsbereich erst im Schuljahr 2009/10 eingeführt, daher ist die Anzahl der Schulen geringer (n=24); auch in den anderen Bereichen wurden zum Teil nicht alle Aspekte an jeder Schule bewertet.

Datenquelle: Landesinstitut für Schulentwicklung

sichtlich der Aspekte *Durchführung* sowie *Wirkung und Wirksamkeit* zeigt sich mehrheitlich eine »auf gutem Niveau entwickelte Praxis«. Zwei Schulen erreichen bei der *Durchführung* eine »exzellente Praxis« sowie eine Schule bei der *Wirkung und Wirksamkeit* der Qualitätsentwicklung im Bereich Unterricht. Die Praxis der anderen Schulen, mit jeweils einer Ausnahme, entspricht den »grundlegenden Anforderungen«. Bezüglich der *organisatorischen Einbindung* der Qualitätsentwicklung im Bereich Unterricht wurde die Praxis der Hälfte der Schulen als »grundlegenden Anforderungen« entsprechend und bei einem Drittel der Schulen als auf »gutem Niveau entwickelt« (fünf Schulen) oder mit »exzellenter Praxis« (drei Schulen) eingeschätzt. An jeder sechsten Schule war die Praxis »wenig entwickelt«.

Für die anderen drei Qualitätsbereiche liegen Einschätzungen für 40 Schulen vor, jedoch wurden nicht bei allen Schulen Bewertungen für alle Aspekte vorgenommen, sodass sich zum Teil Differenzen in Bezug auf die Gesamtzahl ergeben.

Bezogen auf den Qualitätsbereich *Schulleitung und Steuerung der Qualitätsprozesse* zeigt sich, dass in allen drei Aspekten die Praxis in der Mehrheit mindestens als »den grundlegenden Anforderungen« entsprechend oder besser eingestuft wurde. Allerdings finden sich im Vergleich zu dem Qualitätsbereich *Qualitätsentwicklung im Bereich Unterricht* deutlich höhere Anteile an »wenig entwickelter Praxis«. Die Praxis des Aspekts *Wirkung und Wirksamkeit* wurde am häu-

figsten als »wenig entwickelt« eingeschätzt. Der Anteil der Entwicklungsstufe 1 umfasst hier fast 43 %.

Den größten Anteil mit rund 63 % stellen bei der *Durchführung des Individualfeedbacks und der persönlichen Qualitätsentwicklung* Schulen, deren Praxis den »grundlegenden Anforderungen« der Entwicklungsstufe 2 entspricht, gefolgt von 35 %, deren Praxis als »auf gutem Niveau entwickelt« eingeschätzt wurde. Eine Schule erlangt eine »exzellente Praxis«. Keine der Schulen hat, bezogen auf den Aspekt *Durchführung*, eine »wenig entwickelte Praxis«. Bei den Aspekten *Wirkung und Wirksamkeit* und *organisatorische Einbindung* sind es sechs bzw. sieben Schulen, deren Praxis zum Zeitpunkt der Evaluation »wenig entwickelt« war. Insgesamt haben jedoch auch bei diesen beiden Aspekten Schulen mit den »grundlegenden Anforderungen entsprechender Praxis« die größten Anteile (knapp 59 % bzw. 54 %). Außerdem weisen fast 26 % bzw. rund 22 % eine Praxis auf, die auf »gutem Niveau« entwickelt ist. Zu letzteren sind nochmals zwei

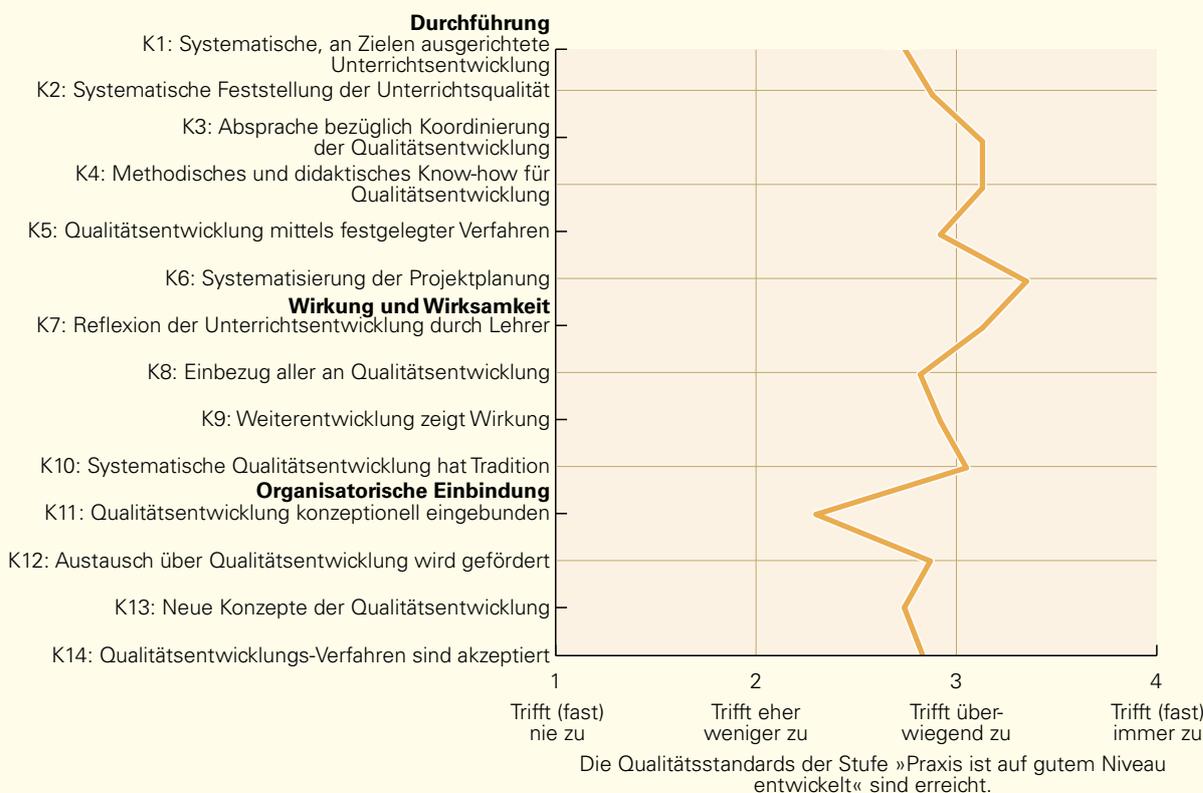
Schulen zu ergänzen, die bei der organisatorischen Einbindung des Individualfeedbacks und der persönlichen Qualitätsentwicklung eine »exzellente Praxis« vorweisen können.

Auch im Bereich *Selbstevaluation und Qualitätsentwicklung der Schule* erreichen die Schulen in allen Aspekten mehrheitlich eine Praxis, die mindestens den »grundlegenden Anforderungen« der Entwicklungsstufe 2 entspricht. Bei den Aspekten *Wirkung und Wirksamkeit* und *organisatorische Einbindung* gibt es jedoch verhältnismäßig hohe Anteile an noch »wenig entwickelter Praxis«, die bei jeweils rund 37 % liegen.

Insgesamt erzielen die beruflichen Schulen in allen Qualitätsbereichen beim Aspekt *Durchführung* die besten Bewertungen; hier wird mit Ausnahme des Bereichs *Schulleitung und Steuerung der Qualitätsprozesse* am häufigsten eine »auf gutem Niveau entwickelte Praxis« erreicht.

I 2.2 (G2)

Einschätzung der Kriterien des Qualitätsbereichs *Qualitätsentwicklung im Bereich Unterricht* an beruflichen Schulen in Baden-Württemberg im Schuljahr 2009/10



Anmerkung: Für die grafische Darstellung wurde die Beschreibung der Kriterien stark gekürzt. Für eine ausführliche Beschreibung vgl. Landesinstitut für Schulentwicklung (2010).
 Datenquelle: Landesinstitut für Schulentwicklung

Die Anzahl der evaluierten beruflichen Schulen ist zurzeit noch relativ gering. Die Aussagen über die bislang evaluierten Schulen sind deshalb noch nicht repräsentativ für die Gesamtheit der beruflichen Schulen in Baden-Württemberg.

Einschätzung der Qualitätskriterien durch die Evaluationsteams

Tiefergehende Einblicke und eine differenziertere Betrachtung der evaluierten Schulen ermöglichen die in den Grafiken I 2.2 (G2) bis I 2.2 (G5) dargestellten Analysen der Kriterien einschätzungen.

Die Evaluatorenteams schätzen auf einer vierstufigen Skala ein, in welchem Umfang einzelne, den vier Qualitätsbereichen zugeordnete Qualitätskriterien von der Schule umgesetzt werden. Die für die Einschätzung herangezogenen Kriterien sind innerhalb der einzelnen Qualitätsbereiche den Aspekten *Durchführung*, *Wirkung und Wirksamkeit* sowie *organisatorische Einbindung* zugeordnet und beziehen sich auf die Stufe 3 »die Praxis ist auf gutem Niveau entwickelt«. Die vorgenommenen Einschätzungen der Kriterien werden für die Einstufung der Qualitätsbereiche im Evaluationsbericht herangezogen.

Der Grad des Zutreffens der einzelnen Kriterien wird mithilfe einer vierstufigen Skala eingeschätzt. Sie umfasst die vier Abstufungen »trifft (fast) nie zu«, »trifft eher weniger zu«, »trifft überwiegend zu« und »trifft (fast) immer zu«. Um Berechnungen vornehmen zu können, wurde diese Skala in die Zahlenwerte 1 (»trifft (fast) nie zu«) bis 4 (»trifft (fast) immer zu«) überführt, mit deren Hilfe anschließend Mittelwerte für jedes Kriterium ermittelt werden konnten. In diesem Zusammenhang wurden die Kriterien eines jeden Qualitätsbereichs beginnend mit K1 durchnummeriert und für die Schaubilder I 2.2 (G2) bis I 2.2 (G5) stichwortartig bezeichnet.¹⁹

Für den *Qualitätsbereich Qualitätsentwicklung im Bereich Unterricht* zeigt sich, dass sich die Kriterien im Mittel in der Region »trifft überwiegend zu« bewegen (Mittelwerte zwischen 2,7 und 3,4) und die Schulen den formulierten Qualitätsanforderungen somit überwiegend auf der Entwicklungsstufe »die Praxis ist auf gutem Niveau entwickelt« gerecht werden (Grafik I 2.2 (G2)). Eine Ausnahme bildet das Kriterium K11, welches eine konzeptionelle Einbindung der Vorgehensweise bei der Qualitätsentwicklung sowie das Vorliegen eines ausgewogenen Verhältnisses zwischen institutionellen Vorga-

ben und persönlichem Gestaltungsspielraum erfordert. Dieses, zum Aspekt der *organisatorischen Einbindung* zählende, Kriterium erhielt im Mittel die Einschätzung »trifft eher weniger zu« (Mittelwert 2,3).

Die Einschätzungen im Qualitätsbereich *Schulleitung und Steuerung der Qualitätsprozesse* stellen sich insgesamt heterogen dar (Grafik I 2.2 (G3)). Fünf Kriterien können ihren im Durchschnitt erzielten Werten nach als »eher weniger zutreffend« eingestuft werden (Mittelwerte zwischen 1,9 und 2,4) und acht Kriterien (Mittelwerte von 2,6 bis 3,4) als »überwiegend zutreffend«. Die drei Kriterien K7, K8 und K12 wurden am niedrigsten bewertet. K7 und K12 schätzen jeweils in den Aspekten *Wirkung und Wirksamkeit* sowie *organisatorische Einbindung* ein überwiegend noch nicht gelingendes Erkennen und Beseitigen von Qualitätsdefiziten ein. K8 fragt nach der Akzeptanz, welche das Qualitätskonzept an der Schule genießt. Das Kriterium K13 erhält insgesamt die beste Bewertung in diesem Qualitätsbereich und bestätigt, dass die Schulen die Qualitätsentwicklung durch gute Rahmenbedingungen unterstützen.

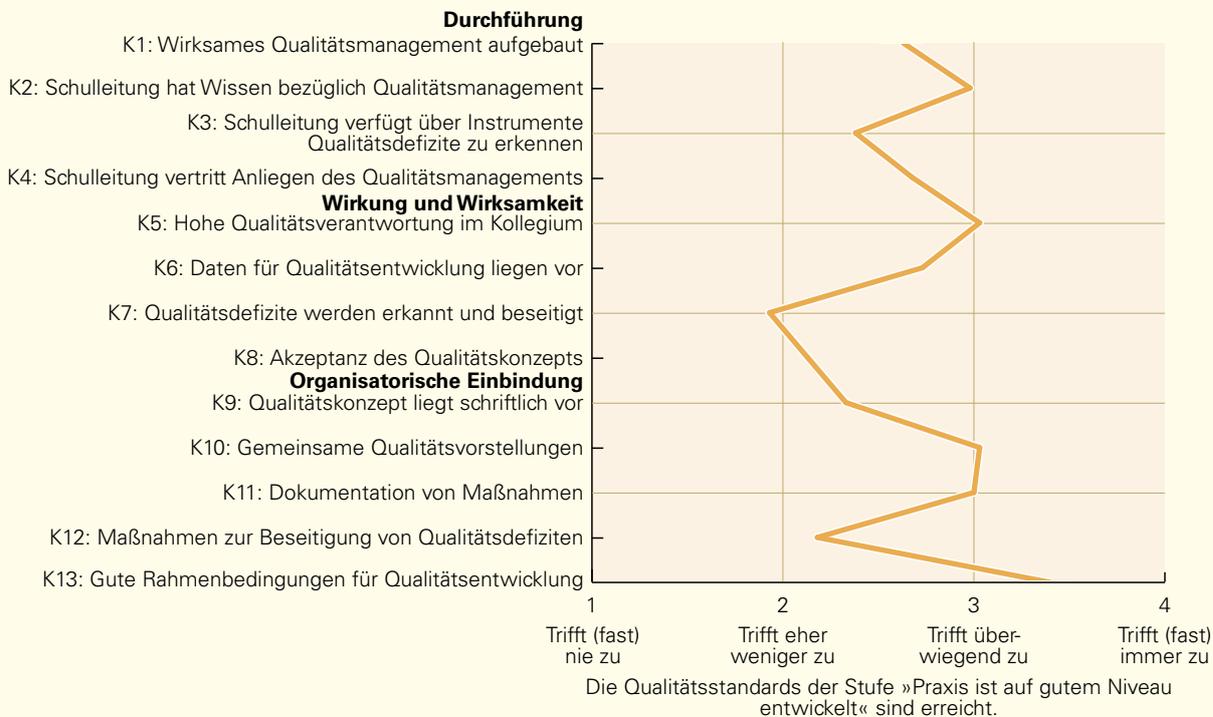
Die meisten Kriterien im Qualitätsbereich *Individualfeedback und persönliche Qualitätsentwicklung* haben eine durchschnittliche Bewertung zwischen »trifft eher weniger zu« und »trifft überwiegend zu« (Grafik I 2.2 (G4)). Zwei Kriterien werden im Durchschnitt aller Schulen am besten erfüllt: Es handelt sich um die Kriterien K9 und K11, welche zeigen, dass die Vorgaben der Schulen zur Feedbackpraxis im Kollegium akzeptiert sind und dass die von den Schulen zur Verfügung gestellten Ressourcen die Feedbackpraxis erleichtern. Kriterien mit Mittelwerten von 2,3 und 2,4 sind K4, K5 und K6, welche sich auf den Aspekt der *Wirkung und Wirksamkeit* der Feedbackpraxis beziehen. Sie nehmen in den Blick, ob das Individualfeedback Auswirkungen zeigt, die Wirksamkeit der durchgeführten Feedbackmaßnahmen reflektiert wird und ob die Auswirkungen der eigenen Feedbackpraxis realistisch wahrgenommen werden.

Die Kriterien des Qualitätsbereichs *Selbstevaluation und Qualitätsentwicklung* der Schule weisen sehr heterogene Bewertungen auf (Grafik I 2.2 (G5)). »Eher weniger« bis »(fast) nie« wurde nach Einschätzung der Evaluatorenteams ein differenziertes Repertoire an Verfahren zur Selbstevaluation aufgebaut und erprobt (K5), lag eine Sammlung von erprobten Evaluationsinstrumenten (K8) vor oder wurden Evaluations- und Entwicklungsmaßnahmen kritisch überprüft (K10). Dass sich die an Selbstevaluationsprojekten beteiligten Lehrkräfte für eine sorgfältige Umsetzung einsetzen (K4) und geeignete Rahmenbedingungen für die Evaluations- und Entwicklungsprozesse an den Schulen vorliegen (K12) wurde hingegen als überwiegend zutreffend eingestuft.

¹⁹ Für eine ausführliche Darstellung der Kriterien vgl. Landesinstitut für Schulentwicklung (2010).

I 2.2 (G3)

Einschätzung der Kriterien des Qualitätsbereichs Schulleitung und Steuerung der Qualitätsprozesse an beruflichen Schulen in Baden-Württemberg in den Schuljahren 2007/08 bis 2009/10



Anmerkung: Für die grafische Darstellung wurde die Beschreibung der Kriterien stark gekürzt. Für eine ausführliche Beschreibung vgl. Landesinstitut für Schulentwicklung (2010).

Datenquelle: Landesinstitut für Schulentwicklung

I 2.2.3 Vertiefende Analysen

Um Erkenntnisse aus der Fremdevaluation zu gewinnen, die über die Zuordnung in Entwicklungsstufen und die Darstellung von mittleren Bewertungen hinausgehen, wurden weiter reichende qualitative Analysen vorgenommen. Diese Auswertungen ermöglichen es, tiefer gehende Informationen zur Qualität schulischer Prozesse und Maßnahmen zu gewinnen. Als Datenmaterial wurden die einzelnen Evaluationsberichte der beruflichen Schulen herangezogen.

Analyse der Empfehlungen

Die in den 40 Evaluationsberichten der beruflichen Schulen niedergelegten 132 Empfehlungen der Evaluationsteams bilden die Grundlage für die qualitative Analyse in diesem Bereich. Im Durchschnitt werden jeder evaluierten Schule etwa drei Empfehlungen für die interne Weiterentwicklung ausgesprochen. Ein ergänzender Text begründet die Empfehlung und führt aus, welche konkreten Aktivitäten aus Sicht der Fremd-

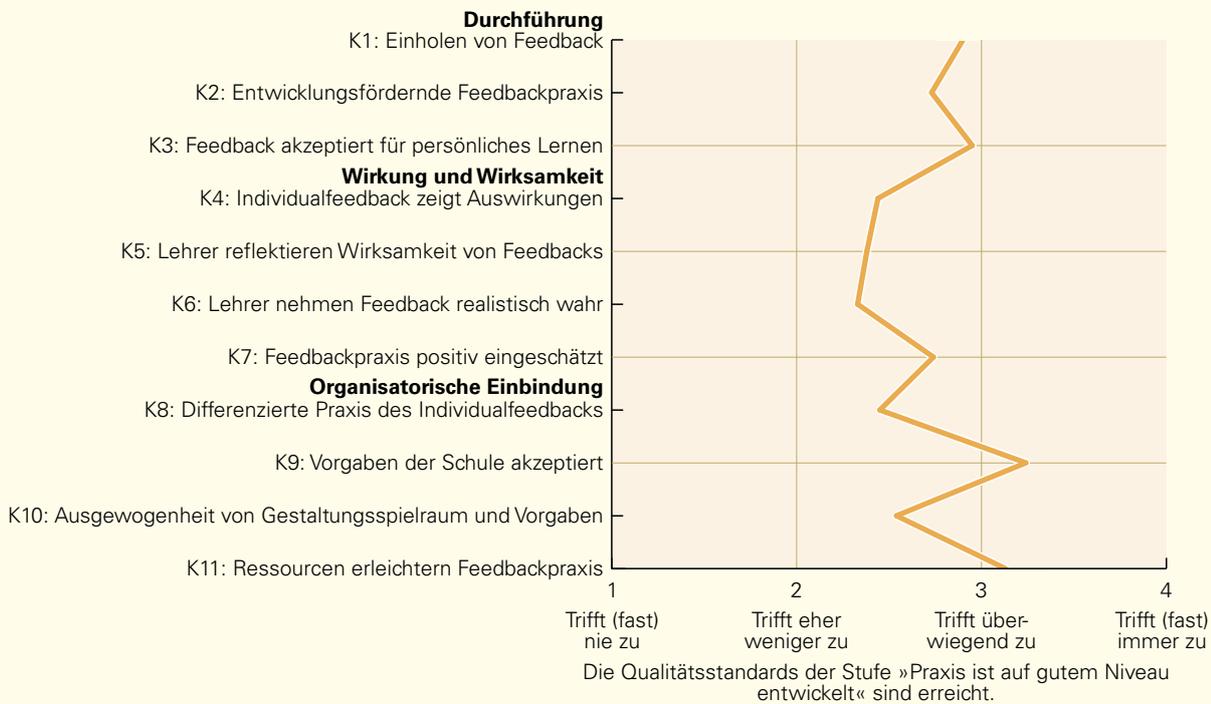
evaluation die Qualitätsentwicklung der Schule unterstützen.

Für die Auswertung im Bildungsbericht werden, wie bei der Auswertung der Evaluationsergebnisse der allgemein bildenden Schulen, die Empfehlungen systematisiert und Kategorien zugeordnet. Allerdings erfolgt hinsichtlich der beruflichen Schulen ein anderes methodisches Vorgehen bei der Kategorienbildung. Während die Empfehlungen der allgemein bildenden Schulen den entsprechenden Kriterien des Qualitätsrahmens der Fremdevaluation zugeordnet wurden, handelt es sich bei der Analyse der Empfehlungen im beruflichen Bereich um ein inhaltsanalytisches Vorgehen, bei welchem die Kategorien aus dem vorliegenden Textmaterial gebildet werden. Die Kategorien sind somit als Zusammenfassungen zu sehen; ihre Titel können als Oberbegriffe für die darunter gefassten Empfehlungen verstanden werden.

Eine Auszählung verdeutlicht, welche Empfehlungen besonders häufig genannt werden (Grafik I 2.2 (G6)). Dies lässt auf den konkreten, zum gegenwärtigen Zeit-

I 2.2 (G4)

Einschätzung der Kriterien des Qualitätsbereichs *Individualfeedback und persönliche Qualitätsentwicklung* an beruflichen Schulen in Baden-Württemberg in den Schuljahren 2007/08 bis 2009/10



Anmerkung: Für die grafische Darstellung wurde die Beschreibung der Kriterien stark gekürzt. Für eine ausführliche Beschreibung vgl. Landesinstitut für Schulentwicklung (2010).

Datenquelle: Landesinstitut für Schulentwicklung

punkt erforderlichen Entwicklungsbedarf an den evaluierten beruflichen Schulen schließen. Wie bereits bei den quantitativen Auswertungen ausgeführt, kann aufgrund der geringen Fallzahl – 40 evaluierte Schulen in den Schuljahren 2007/08 bis 2009/10 – keine Repräsentativität für die gesamten beruflichen Schulen im Land beansprucht werden.

17 % der Empfehlungen konnten in die Kategorie *Qualitätsmanagement und -entwicklung wie bestehend fortführen und (strukturell) ausbauen* eingeordnet werden. Diese Empfehlungen haben das Pflegen und Fortführen bestehender Qualitätselemente zum Gegenstand, aber auch deren Ausbau auf dem bereits eingeschlagenen Weg. Sechs der insgesamt 23 Empfehlungen dieser Kategorie nehmen hierbei speziell den strukturellen Ausbau des Qualitätsmanagements in den Blick.

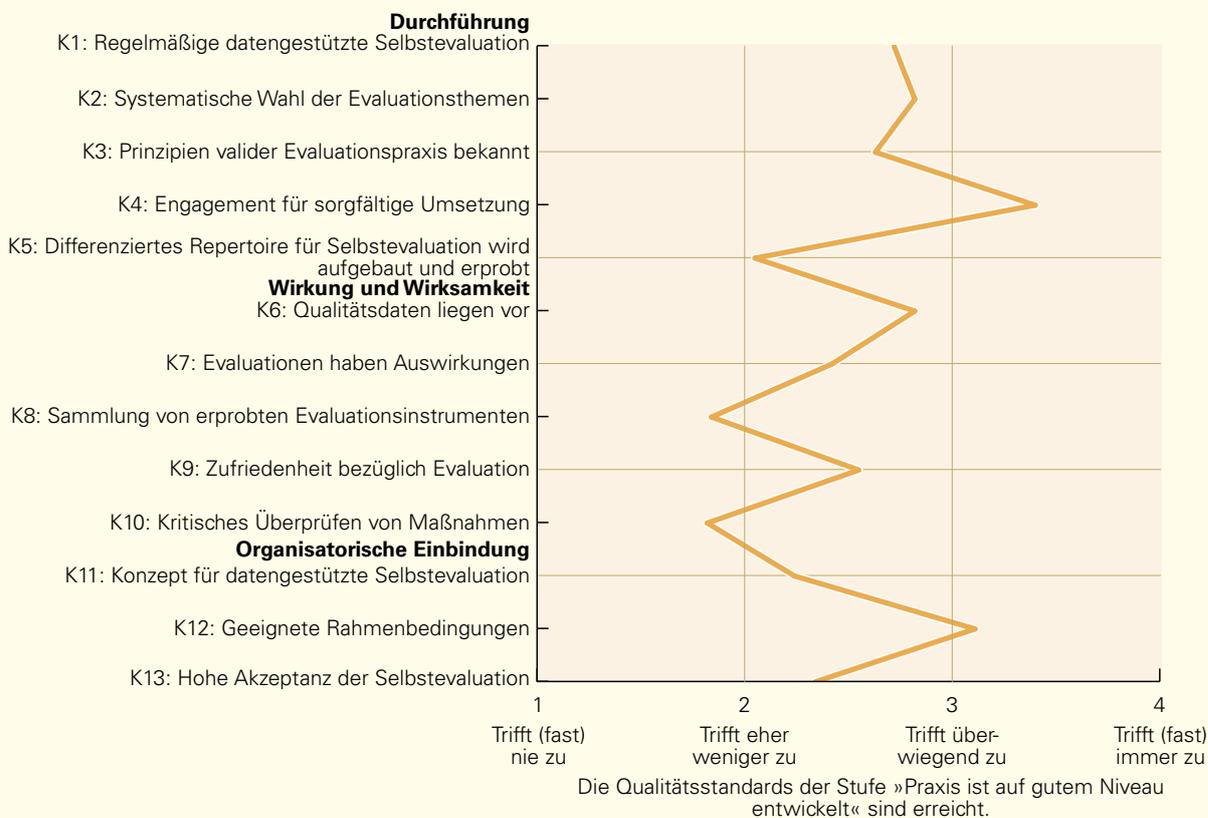
Rund jede siebte ausgesprochene Empfehlung (15 %) fällt in die Kategorie *Qualitätsarbeit konzeptuell weiterentwickeln/ ein gemeinsames Qualitätsverständnis entwickeln*. Diese Kategorie umfasst Empfehlungen, die sich auf die Arbeit am Qualitätskonzept der jewei-

ligen Schule beziehen, so zum Beispiel darauf, noch unverbundene Elemente der Qualitätsentwicklung zu einem umfassenden Konzept zusammenzuführen. Des Weiteren wurden dieser Kategorie Empfehlungen zugeordnet, die das Entwickeln eines gemeinsamen Qualitätsverständnisses betreffen, da dieses als Grundlage für die Erstellung und Weiterentwicklung eines Qualitätskonzepts angesehen werden kann.

14% der Empfehlungen lassen sich unter der Überschrift *Kommunikation ausbauen/ alle am Schulleben Beteiligten in die Qualitätsentwicklung einbeziehen* vereinen. Hierbei steht einerseits die Kommunikation, andererseits die Integration weiterer an der Schule vertretener Gruppen in den Prozess der Qualitätsentwicklung im Mittelpunkt. So weisen einige Empfehlungen darauf hin, dass Schüler noch stärker in die Qualitätsentwicklung eingebunden werden sollten. Ebenfalls 14 % der Empfehlungen sind *Empfehlungen zu der Selbstevaluation* der Schulen. Diese lassen sich inhaltlich in zwei Gruppen teilen: acht Empfehlungen beziehen sich auf den Auf- und Ausbau der Selbstevaluation an der jeweiligen Schule, während sich zehn Empfehlungen damit

I 2.2 (G5)

Einschätzung der Kriterien des Qualitätsbereichs *Selbstevaluation und Qualitätsentwicklung der Schule* an beruflichen Schulen in Baden-Württemberg in den Schuljahren 2007/08 bis 2009/10



Anmerkung: Für die grafische Darstellung wurde die Beschreibung der Kriterien stark gekürzt. Für eine ausführliche Beschreibung vgl. Landesinstitut für Schulentwicklung (2010).

Datenquelle: Landesinstitut für Schulentwicklung

Statistisches Landesamt Baden-Württemberg

309 11

befassen, die Selbstevaluation und ihre Ergebnisse (stärker) für Schul- und Unterrichtsentwicklungsprozesse zu nutzen.

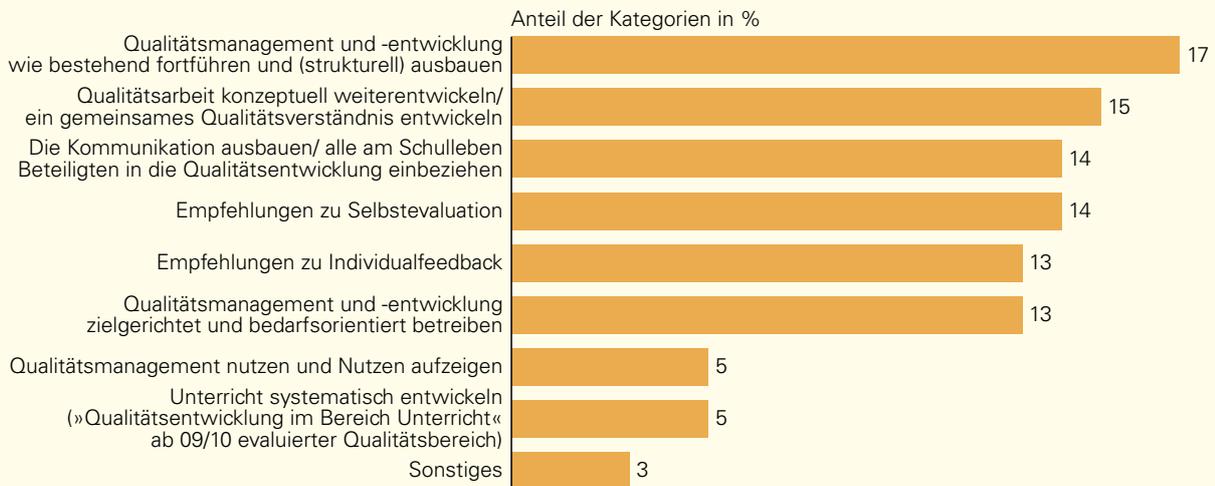
Für die Kategorie *Empfehlungen zu Individualfeedback*, welche 13 % der Empfehlungen umfasst, finden sich elf Empfehlungen, die den Auf- und Ausbau des Individualfeedbacks anregen. Sechs Empfehlungen beziehen sich auf das Nutzen des Individualfeedbacks für Entwicklungsprozesse. Weitere 13 % der Empfehlungen lassen sich unter die Kategorie *Qualitätsmanagement und -entwicklung zielgerichtet und bedarfsorientiert betreiben* fassen. Anhand dieser Empfehlungen werden die betreffenden Schulen darauf hingewiesen, dass eine Definition von Zielen für das Qualitätsmanagement bzw. die Qualitätsentwicklung und eine Ausrichtung am jeweiligen Bedarf der Schule zielführend wäre.

5 % der Empfehlungen fordern dazu auf, das Qualitätsmanagement für die schulische Arbeit zu nutzen bzw. den Nutzen des Qualitätsmanagements aufzuzeigen. Ebenfalls 5 % der Empfehlungen heben darauf ab, den Unterricht an der Schule systematisch weiterzuentwickeln. Es ist hierbei zu beachten, dass der Qualitätsbereich *Qualitätsentwicklung im Bereich Unterricht* als eigenständiger Bereich erst ab Schuljahr 2009/10 evaluiert wurde.

Die Kategorie *Sonstiges* umfasst 3 % der Empfehlungen. Bei den dieser Kategorie zugeordneten Empfehlungen handelt es sich unter anderem um den Vorschlag, Erfolge nach außen im Sinne eines Marketings zu kommunizieren und um die stärkere Würdigung der Erfolge in der Qualitätsentwicklung.

I 2.2 (G6)

**Empfehlungen der Evaluationsteams an beruflichen Schulen
in Baden-Württemberg in den Schuljahren 2007/08 bis 2009/10**



Datenquelle: Landesinstitut für Schulentwicklung

Statistisches Landesamt Baden-Württemberg

310 11

I 3 Diagnose- und Vergleichsarbeiten

Mit der Einführung des Bildungsplans 2004 legte Baden-Württemberg als erstes Bundesland im Bereich der allgemein bildenden Schulen Bildungsstandards fest. Damit verbunden ist eine systemische Umstellung von der Vorgabe dessen, was unterrichtet werden muss (Input-Steuerung), auf die Festlegung dessen, was letztlich von den Schülern gelernt werden soll (Output-Steuerung). Um auf der Ebene der Output-Steuerung den Lernstand der Schüler differenziert abbilden zu können, werden Vergleichsarbeiten durchgeführt. Vergleichsarbeiten sind standardisierte Lernstandserhebungen, die zu festgelegten Terminen den Lernstand von Klassen und auch von einzelnen Schülern in Bezug auf Bildungsstandards überprüfen. Durch die Rückmeldung der Ergebnisse erhalten Lehrkräfte objektive Informationen, welche zur Weiterentwicklung von Schule und Unterricht genutzt werden können. In Baden-Württemberg werden zwei Verfahren zur Lernstandserhebung eingesetzt: an den Grundschulen die *Diagnosearbeiten* (VERA 3, vgl. **Kapitel I 3.1**), an den weiterführenden Schulen die *Vergleichsarbeiten* (DVA, vgl. **Kapitel I 3.2**).

I 3.1 Diagnosearbeiten an Grundschulen (VERA 3)

In den Schuljahren 2005/06 bis 2007/08 wurden an den Grundschulen in Baden-Württemberg am Ende der Klassenstufen 2 und 4 landeseigene Diagnosearbeiten geschrieben. Seit dem Schuljahr 2008/09 beteiligen sich die baden-württembergischen Grundschulen verpflichtend an den bundesweit durchgeführten *Vergleichsarbeiten in der Schule* (VERA). Diese Lernstandserhebung wird im zweiten Schulhalbjahr der Klassenstufe 3 geschrieben und daher auch als VERA 3 bezeichnet. VERA orientiert sich an den von der Kultusministerkonferenz (KMK) verabschiedeten Bildungsstandards des Faches Deutsch²⁰ bzw. Mathematik²¹ für den Primarbereich. Diese sind als Regelstandards definiert und beschreiben Kompetenzen, die im Durchschnitt von allen Schülern bis zum Ende der Klasse 4 erreicht werden sollen.

Zuständig für die Konzeption der VERA-3-Arbeiten ist seit dem Schuljahr 2009/10 das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) in Berlin, welches die zuvor zuständige Universität Koblenz-Landau abgelöst hat. Baden-Württemberg entsendet zur Unter-

stützung des IQB bei der Aufgabenentwicklung landeseigene Fachdidaktik-Experten. Diese sichern den Bezug des Gesamttests zu den baden-württembergischen Bildungsplänen für die Grundschule.

Leistungsrückmeldung orientiert sich an Kompetenzstufen

Seit dem Schuljahr 2009/10 werden die Ergebnisse aus VERA 3 (bisher mit Ausnahme von Orthografie) auf fünf Kompetenzstufen im jeweiligen Fach bzw. Inhaltsbereich zurückgemeldet. Da VERA 3 die Kompetenzen der Bildungsstandards zugrunde legt, können mithilfe des Tests in den geprüften Bereichen Stärken und Schwächen der Drittklässler bezogen auf die zu erreichenden Kompetenzen am Ende der Klassenstufe 4 ermittelt werden. Für Lesen und Mathematik finden sich Kurzbeschreibungen der einzelnen Kompetenzstufen in Tabelle I 3.1 (T1). Für Mathematik sind jeweils die in Klassenstufe 4 erwarteten Anteile für die jeweiligen Stufen angegeben. Die Leistungen von Drittklässlern können diesen Erwartungen noch nicht im vollen Umfang entsprechen, dies muss bei der Interpretation der Ergebnisse berücksichtigt werden.

Eine detaillierte Beschreibung des Kompetenzstufenmodells für das Fach Mathematik kann auf der Homepage des IQB abgerufen werden.²² Das Kompetenzstufenmodell für den Inhaltsbereich Lesen im Fach Deutsch war zum Redaktionsschluss noch nicht abrufbar. Für Orthografie lag zum Zeitpunkt der Durchführung noch kein Modell vor, hier wurden die Leistungen in Leistungsgruppen zurückgemeldet. Die Grenzen der Leistungsgruppen wurden aufgrund einer gesamtdeutschen Voruntersuchung ermittelt, wobei jede Leistungsgruppe 25 % der Schüler umfasste. Die landesspezifischen Werte sind in Grafik I 3.1 (G1) dargestellt.

VERA ermöglicht einen objektiven Blick auf die Lernergebnisse baden-württembergischer Drittklässler

Die Ergebnisse der Vergleichsarbeiten sollen in erster Linie Lehrkräften einen objektiven Blick auf die Leistungen der eigenen Klasse ermöglichen. Sie bieten den Lehrkräften zuverlässige Informationen, welche

20 Vgl. Kultusministerkonferenz (Hrsg.) (2004b).

21 Vgl. Kultusministerkonferenz (Hrsg.) (2004c).

22 Vgl. Institut zur Qualitätsentwicklung im Bildungswesen (2008).

I 3.1 (T1) Allgemeine Kurzbeschreibung der Kompetenzstufenmodelle bei VERA 3 (2010): Mathematik und Lesen

Lesen Beschreibung	Kompetenzstufe	Mathematik	
		Beschreibung	Anteile in % ¹⁾
Explizit angegebene Einzelinformationen identifizieren	1	Routineprozeduren auf Grundlage einfachen begrifflichen Wissens	10
Benachbarte Informationen miteinander verknüpfen	2 Mindeststandard	Einfache Anwendungen von Grundlagenwissen	21
»Verstreute« Informationen miteinander verknüpfen und den Text ansatzweise als Ganzen erfassen	3 Regelstandard	Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext	33
Für die Herstellung von Kohärenz auf der Ebene des Textes wesentliche Aspekte erfassen	4 Regelstandard plus	Sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang	21
Auf zentrale Aspekte des Textes bezogene Aussagen selbstständig begründen	5 Maximalstandard	Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien	15

1) Anteile von Schülern in der 4. Klasse der gesamtdeutschen Vergleichsstichprobe für Mathematik. In Deutsch sind die Anteile noch nicht veröffentlicht, es ist aber zu erwarten, dass die Verteilung in etwa der in Mathematik entspricht.

Datenquelle: Institut zur Qualitätsentwicklung im Bildungswesen

Lernergebnisse ihre Schüler erreicht haben und inwieweit die Anforderungen der nationalen Bildungsstandards erfüllt wurden. So können die Vergleichsarbeiten eine wichtige Grundlage für eine systematische Schul- und Unterrichtsentwicklung bilden. Ebenso können die Ergebnisse als Grundlage für fachdidaktische Diskussionen dienen.

Um eine differenzierte Betrachtung zu ermöglichen, stehen bei VERA 3 vielfältige Informationen und insbesondere folgende Vergleichsarten zur Verfügung:

- An den Bildungsstandards orientierter Leistungsvergleich. Der Vergleich der Leistungen der einzelnen Schüler auf der Grundlage von Kompetenzstufen ermöglicht eine direkte Orientierung an den Bildungsstandards und eine inhaltliche Bewertung der Ergebnisse.
- Bezugsgruppenorientierter Vergleich. Diese Art der Darstellung ermöglicht eine vergleichende Betrachtung der Leistungen in der eigenen Klasse in Bezug auf die des eigenen Bundeslandes.

Bei jedem zehnten Drittklässler ist Deutsch nicht die dominante Alltagssprache

Es werden in der Regel jährlich wechselnd Inhaltsbereiche geprüft, wobei allerdings der Bereich Lesen in jedem Jahr Bestandteil des Tests ist. Im Jahr 2010

beinhaltete der Test für Deutsch die Bereiche Lesen und Orthografie und für Mathematik die Bereiche Daten, Häufigkeit und Wahrscheinlichkeit sowie Zahlen und Operationen.

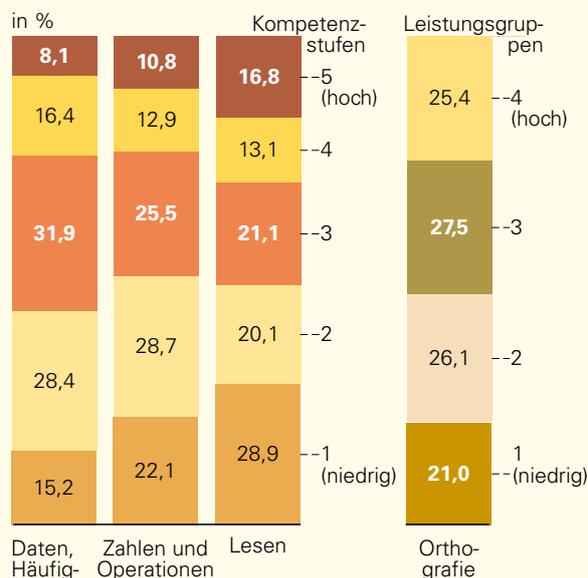
In Baden-Württemberg nahmen etwa 92 000 Grundschüler an VERA 3 teil – dies entspricht 87 % der Schüler der Jahrgangsstufe 3 im Schuljahr 2009/10. Darunter waren etwas mehr Jungen (51 %) als Mädchen. Etwas mehr als ein Zehntel aller Schüler sprechen im Alltag eine andere Sprache als Deutsch. Knapp 3 % wiederholten zum Zeitpunkt von VERA 3 die 3. Klasse. Bei rund 8 % lag eine Teilleistungsstörung (Angabe erhoben durch die Lehrkraft) in Deutsch vor (Tabelle I 3.1 (T2) im Anhang).

Jeweils ein Drittel der Schüler übertrifft bzw. unterschreitet die Regelstandards

Auffällig sind die im Inhaltsbereich Lesen relativ hohen Anteile auf den Kompetenzstufen 4 und 5 (rund 30 %), dies bedeutet, dass in Baden-Württemberg bereits in Klasse 3 knapp ein Drittel aller Schüler eine Kompetenz aufweist, die dem Regelstandard plus – bezogen auf die vierte Jahrgangsstufe – oder höher entspricht (Grafik I 3.1 (G1)). Andererseits findet sich ein fast ebenso großer Anteil von Schülern auf der anderen Seite des Kompetenzkontinuums: Nahezu 29 % der Schüler erreichen in Baden-Württemberg in Klassenstufe 3 noch nicht den Mindeststandard im Bereich Lesen.

I 3.1 (G1)

Verteilung der Schüler in VERA 3*) auf die Kompetenzstufen bzw. Leistungsgruppen in Baden-Württemberg 2010



*) Vergleichsarbeiten in der Schule in Klassenstufe 3.

Datenquelle: Institut zur Qualitätsentwicklung im Bildungswesen

Statistisches Landesamt Baden-Württemberg

301 11

Im Inhaltsbereich Zahlen und Operationen finden sich ebenfalls in etwa gleich große Anteile von Schülern, die die Mindeststandards verfehlen bzw. die Leistungen erbringen, die mindestens dem Regelstandard plus entsprechen – allerdings sind diese Anteile jeweils kleiner (rund 22 % bzw. knapp 24 %). Im Bereich Daten, Häufigkeit und Wahrscheinlichkeit ist der Anteil der Schüler auf Kompetenzstufe 1 hingegen vergleichsweise gering (gut 15 %), zugleich zeigen etwa ein Viertel der Schüler in diesem Bereich eine Kompetenz, die dem Regelstandard plus oder höher entspricht. Gut 60 % der Schülerschaft erreichen mit ihren Leistungen in diesem Bereich Minimal- bzw. Regelstandards.

Stellt man diese Ergebnisse den erwartbaren Anteilen auf den Kompetenzstufen (Tabelle I 3.1 (T1)) bezogen auf die 4. Klasse gegenüber, so zeigt sich, dass hinsichtlich der Inhaltsbereiche in Mathematik zwar die unteren Kompetenzstufen erwartungsgemäß stärker besetzt sind, aber auch im oberen Bereich (Stufen 4 und 5) bereits knapp ein Viertel aller Schüler zu finden ist. Im Bereich Lesen kommen Unterschiede in der Schülerschaft besonders zum Tragen. Der Anteil derer, die bereits jetzt den Maximalstandard erreichen, ist sogar größer als er für die Jahrgangsstufe 4 erwartbar ist. Andererseits ist – wie bereits erwähnt – der Anteil derjenigen groß, die den Mindeststandard noch nicht erreichen.

Für den Bereich der Orthografie – hier wurden die Leistungen anstelle von Kompetenzstufen in Leistungsgruppen zurückgemeldet – zeigt sich, dass im Vergleich zur Voruntersuchung, bei der jede Leistungsgruppe 25 % der Schüler umfasst, der Anteil der Schüler, die der untersten Leistungsgruppe zuzuordnen sind, für Baden-Württemberg geringer ist. Gleichzeitig sind die mittleren Leistungsgruppen etwas stärker besetzt. Dies bedeutet, dass die baden-württembergischen Schüler bereits in einem etwas größeren Umfang Kompetenzen erworben haben, als aufgrund der Voruntersuchung erwartbar gewesen wäre.

Unterschiedliche Ergebnisse von Mädchen und Jungen

Wie in zahlreichen anderen Schulleistungsstudien zeigt sich, dass Mädchen in den sprachbezogenen Bereichen höhere Anteile auf den beiden oberen Kompetenzstufen bzw. in der oberen Leistungsgruppe aufweisen. Umgekehrt befinden sich weniger Mädchen in der unteren Kompetenzstufe bzw. Leistungsgruppe. Für die Jungen sind die beschriebenen Verteilungen auf den mathematischen Bereich »Zahlen und Operationen« übertragbar (Grafik I 3.1 (G2)).

Im Inhaltsbereich Daten, Häufigkeit und Wahrscheinlichkeit ist der Unterschied am geringsten. Jungen und Mädchen sind zu fast gleichen Anteilen in den jeweiligen Kompetenzstufen vertreten.

In den Bereichen Lesen und Orthografie weist ein relativ großer Teil der Mädchen bereits zu diesem Zeitpunkt eine sehr hohe Kompetenz auf, allerdings gilt gleichzeitig im Bereich Lesen – dies jedoch für beide Geschlechter –, dass auf der unteren Kompetenzstufe der Anteil der Schüler ebenfalls relativ groß ist.

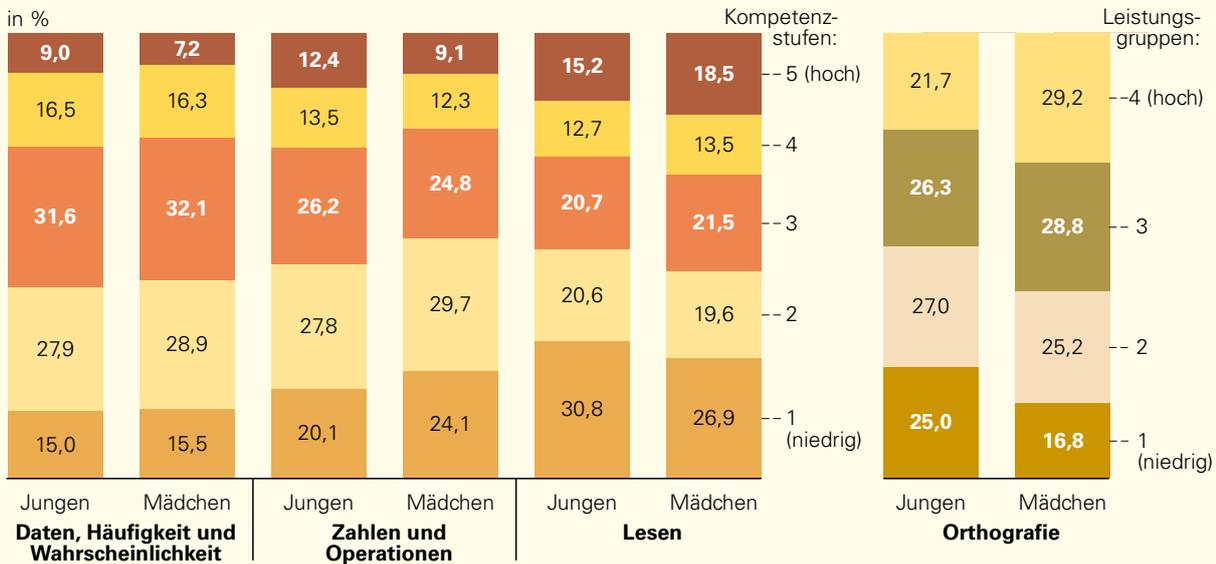
60 % der Schüler mit nicht-dominanter Sprache Deutsch verfehlen die Mindeststandards

Etwas mehr als 11 % der Schüler der Stichprobe sprechen im Alltag eine andere Sprache als Deutsch (Tabelle I 3.1 (T2) im Anhang). Für die Zuordnung der Schüler wurde die Einschätzung der Lehrkräfte zugrunde gelegt, ob für Schüler, unabhängig von Nationalität und Geburtsort, Deutsch die dominante Sprache im Alltag ist oder nicht.

In den Inhaltsbereichen des Faches Deutsch zeigen sich erwartungsgemäß die größten Unterschiede bezüglich der Sprachdominanz. Am deutlichsten zeigen sich die Differenzen im Bereich Lesen. Insgesamt knapp 61 %

I 3.1 (G2)

Verteilung der Schüler in VERA 3*) auf die Kompetenzstufen bzw. Leistungsgruppen in Baden-Württemberg 2010 nach Geschlecht



*) Vergleichsarbeiten in der Schule in Klassenstufe 3.
 Datenquelle: Institut zur Qualitätsentwicklung im Bildungswesen

der Schüler mit Deutsch als nicht dominanter Sprache erreichen nicht einmal den Mindeststandard, während andererseits knapp ein Fünftel dieser Gruppe über Kompetenzen verfügt, die dem Regelstandard entsprechen oder sogar darüber hinausgehen (Grafik I 3.1 (G3)). Hingegen zeigen etwas mehr als die Hälfte aller Schüler mit deutschem Sprachhintergrund Leistungen, die dem Regelstandard bzw. darüber hinausgehenden Kompetenzen entsprechen. Allerdings erreicht auch in dieser Gruppe etwa ein Viertel in Klassenstufe 3 noch keine Mindeststandards.

In Mathematik ist zwar derselbe Trend deutlich erkennbar, insbesondere bezogen auf die jeweils unterste Kompetenzstufe, aber weniger ausgeprägt. Beispielsweise befinden sich im Bereich Daten, Häufigkeit und Wahrscheinlichkeit etwas mehr als ein Drittel der Schüler mit nicht-deutschem Sprachhintergrund auf Kompetenzstufe 1. Andererseits gilt insgesamt für dieselben Schüler, dass in der Mathematik die Anteile auf den Stufen 3 bis 5 im Vergleich zum Lesen stärker besetzt sind.

»Fairer Vergleich« ermöglicht angemessene Einschätzung der Schülerleistungen unter Berücksichtigung des sozialen Kontextes

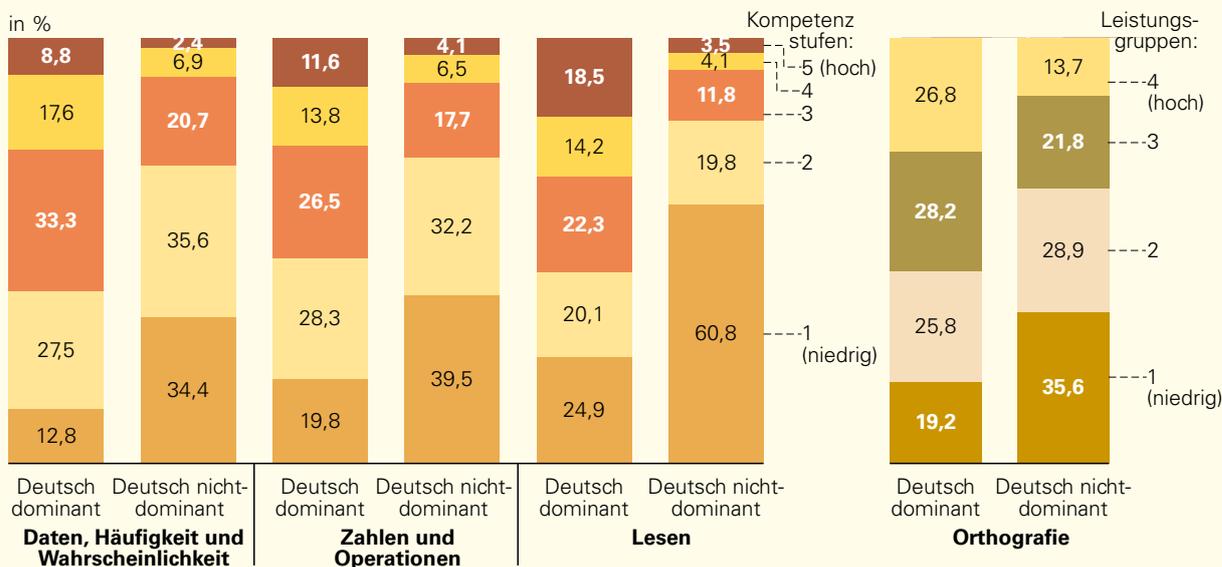
Für VERA 3 werden sogenannte Kontextgruppen für den »fairen Vergleich« gebildet, der die unterschied-

lichen sozial determinierten Lernvoraussetzungen der Schüler berücksichtigt. Diese sind definiert durch bestimmte Aspekte des sozialen Umfelds von Klassen bzw. Schulen. Lehrkräfte können so die Leistungen ihrer Klasse nicht nur mit dem Durchschnitt des Landes (Grafik I 3.1 (G1)) vergleichen, sondern sie können sie mit dem Durchschnitt einer Gruppe von Klassen vergleichen, die bezogen auf den sozialen Kontext der eigenen Klasse ähneln (Grafik I 3.1 (G4)), die also unter gleichen Rahmenbedingungen arbeiten. Hierbei werden Variablen berücksichtigt, die nicht von Lehrkräften verändert werden können, wie beispielsweise der Sprachhintergrund der Kinder, das Bildungsniveau der Eltern oder die sozialen Verhältnisse der Familien. Durch Angaben der Lehrkraft hinsichtlich der in Tabelle I 3.1 (T3) genannten Kriterien werden die Werte der eigenen Klasse denen der am besten passenden Kontextgruppe gegenübergestellt.

So kann es beispielsweise für sogenannte Brennpunktschulen sinnvoll sein, sich nicht nur mit dem Landesdurchschnitt zu vergleichen, da die betreffenden Schulen bei diesem Vergleich meist schlechter abschneiden und die Einordnung ihrer Ergebnisse möglicherweise als nicht aussagekräftig angesehen werden. Über den »fairen Vergleich« können die Ergebnisse zusätzlich vor dem Hintergrund des sozialen Umfelds interpretiert und dementsprechend gewürdigt werden.

I 3.1 (G3)

Verteilung der Schüler in VERA 3*) auf die Kompetenzstufen bzw. Leistungsgruppen in Baden-Württemberg 2010 nach Deutsch als dominante/nicht dominante Sprache



*) Vergleichsarbeiten in der Schule in Klassenstufe 3.
 Datenquelle: Institut zur Qualitätsentwicklung im Bildungswesen

Die Kontextgruppe 1 weist die für die schulische Leistung günstigste soziale Zusammensetzung auf, Kontextgruppe 3 die ungünstigste. So zeichnet sich die erste Kontextgruppe beispielsweise dadurch aus, dass nahezu kein Kind eine andere Sprache als Deutsch im Alltag spricht. Weiterhin liegen die zugeordneten

Schulen nicht in einem sozialen Brennpunkt und der Anteil von Schülern aus Familien der Grund- und Unterschicht beträgt rund ein Fünftel. Umgekehrt ist beispielsweise dieser Anteil in Kontextgruppe 3 sehr hoch (49 %) und auch andere soziale Kontextfaktoren sind hier anders gewichtet.

I 3.1 (T3)

Soziale Zusammensetzung der Kontextgruppen anhand durchschnittlicher Merkmalsausprägungen für Baden-Württemberg 2010

Merkmal	Kontextgruppe ¹⁾		
	1	2	3
	Anteil in %		
Kinder mit Deutsch als nicht dominanter Sprache	0	8	29
Klasse im sozialen Brennpunkt gelegen	nein	nein	ja
Schüler aus Familien der Grund- oder Unterschicht	20	23	49
Schüler aus Familien, deren Lebensunterhalt überwiegend durch Bezug von Sozialleistungen bestritten wird	7	6	20
Schüler aus Familien, die von Arbeitslosigkeit betroffen sind	4	5	15

1) Kontextgruppe 1: günstigster Kontext, Kontextgruppe 3: ungünstigster Kontext.
 Datenquelle: Institut zur Qualitätsentwicklung im Bildungswesen

Es ist deutlich zu erkennen, dass der soziale Kontext einer Klasse mit ihrem Leistungsvermögen zusammenhängt. So finden sich für alle Inhaltsbereiche in der günstigsten Kontextgruppe (Gruppe 1) größere Anteile von Schülern auf den Kompetenzstufen 4 und 5 als in der ungünstigsten Kontextgruppe (Gruppe 3). Umgekehrt sind die Anteile auf der niedrigsten Kompetenzstufe in der ungünstigsten Kontextgruppe wesentlich größer (Grafik I 3.1 (G4)).

Auffällig ist für alle Testinhalte, dass es insbesondere für Schüler in Klassen, die der Kontextgruppe 3 zuzuordnen sind, schwierig ist, gute Leistungen zu erbringen. Am auffälligsten erscheint der Einfluss beim Lesen: Mehr als 44% der Schüler, die der ungünstigsten Kontextgruppe zuzuordnen sind, erreichen hier nicht den Mindeststandard. Andererseits erfüllen trotz ungünstiger Kontextbedingungen gut 18% dieser Gruppe bereits in Klasse 3 den Regelstandard plus oder sogar darüber hinaus. Obwohl günstigste Kontextbedingungen vorliegen, erreichen fast 22% der Schülerschaft in Kontextgruppe 1 keine Minimalstandards, während andererseits über ein Viertel bereits Maximalstandards erreichen.

In der Orthografie zeigt ein knappes Drittel der Schüler der ungünstigsten Kontextgruppe nur Leistungen, die sich der untersten Leistungsgruppe zuordnen lassen, während andererseits nahezu 19% dieser Schülerschaft

die obere Leistungsgruppe erreicht. Für die günstigste Kontextgruppe kehren sich diese Anteile fast um.

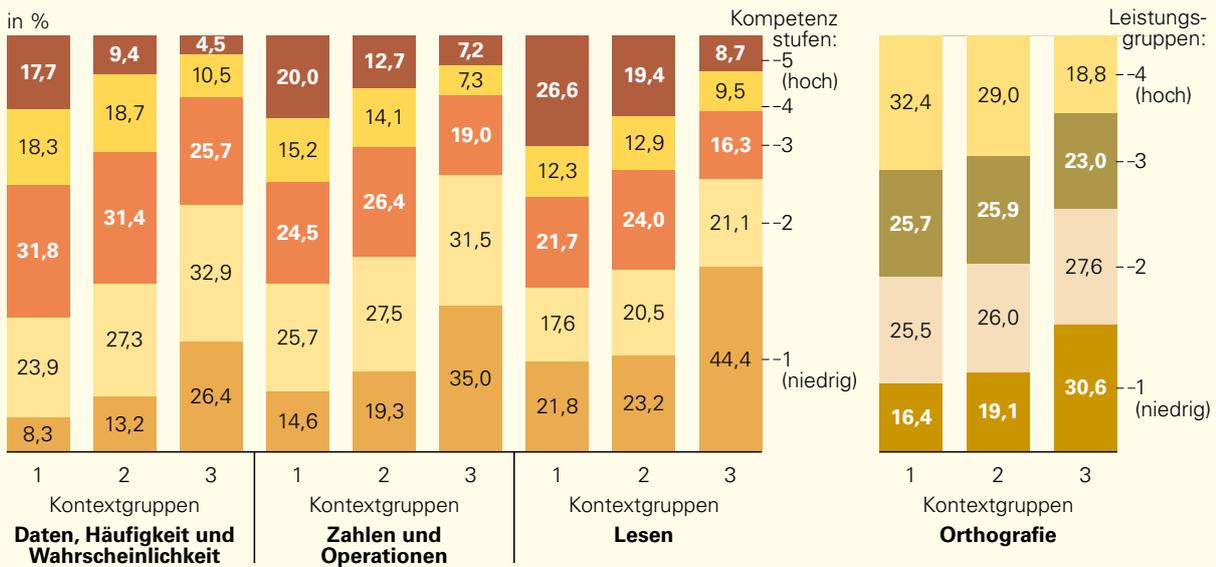
I 3.2 Vergleichsarbeiten an weiterführenden Schulen (DVA)

Die zum Schuljahr 2005/06 an öffentlichen Hauptschulen/ Werkrealschulen, Realschulen und Gymnasien verpflichtend eingeführten Vergleichsarbeiten (DVA) sind standardisierte Tests, die den Lernstand von Klassen und auch von einzelnen Schülern in Bezug auf die baden-württembergischen Bildungsstandards des vorangegangenen Bildungsabschnitts überprüfen. Die Arbeiten werden an festgelegten Terminen zu Beginn des neuen Schuljahres in den Klassenstufen 7 und 9 an den Schulen eingesetzt. Durch standardisierte Korrekturanweisungen und einen Vergleich mit dem baden-württembergischen Leistungsdurchschnitt liefern die Ergebnisrückmeldungen auf Klassen- bzw. Schulebene objektive Informationen, die von den Fachlehrkräften sowie Schulleitungen zur Weiterentwicklung von Schule und Unterricht genutzt werden können. Die Vergleichsarbeiten stellen ein verpflichtendes Instrument der Selbstevaluation dar und sind ein wichtiges Element der Qualitätsentwicklung an Schulen.

Die Vergleichsarbeiten werden am Landesinstitut für Schulentwicklung konzipiert. Im Jahr 2010 wurden zwei

I 3.1 (G4)

Verteilung der Schüler in VERA 3*) auf die Kompetenzstufen bzw. Leistungsgruppen in Baden-Württemberg 2010 nach Kontextgruppen



*) Vergleichsarbeiten in der Schule in Klassenstufe 3.
 Datenquelle: Institut zur Qualitätsentwicklung im Bildungswesen

I 3.2 (T1) Vergleichsarbeiten DVA an öffentlichen allgemein bildenden Schulen in Baden-Württemberg 2010

Schulart	Klassenstufe ¹⁾	Vergleichsarbeiten in den Fächern
Werkrealschule/ Hauptschule	7	Deutsch, Mathematik
	9	Deutsch, Mathematik, Pflichtfremdsprache Englisch oder Französisch
Realschule	7	Deutsch, Mathematik
	9	Deutsch, Mathematik, Pflichtfremdsprache Englisch oder Französisch
Gymnasium	7	Deutsch, Mathematik
	9	Deutsch, Mathematik, eine der Fremdsprachen Englisch, Französisch oder Latein ²⁾

1) Überprüft wird jeweils der Bildungsstandard des vorausgegangenen Bildungsabschnitts. – 2) Die Gesamtlehrerkonferenz entscheidet für die jeweilige Klasse, welche Fremdsprache getestet wird.

Tests für die Werkrealschule/ Hauptschule, sechs Arbeiten für die Realschule und sieben weitere Arbeiten für das Gymnasium entwickelt (Tabelle I 3.2 (T1)).

Alle Aufgaben für die Vergleichsarbeiten werden in einer Pilotierung (Vorabtestung vor dem eigentlichen Einsatz) erprobt und repräsentieren zwei oder drei Schwerpunktbereiche der Bildungsstandards des jeweiligen Faches. Für diese Vorabtestung werden in jedem Jahr für jedes Regierungspräsidium anteilig Stichprobenschulen zufällig gezogen, von denen die jeweiligen Ergebnisse anonymisiert an das Landesinstitut für Schulentwicklung zurückgemeldet werden.

Auf Grundlage der Pilotierung werden geeignete Aufgaben für den endgültigen Test ausgewählt und die Gesamtschwierigkeit festgelegt. Hierzu werden die Arbeiten so zusammengestellt, dass die mittlere Lösungswahrscheinlichkeit für den Gesamttest im Idealfall 50 % beträgt und die Aufgaben über den gesamten Schwierigkeitsbereich möglichst gleichmäßig gestreut sind. Somit ist für einen in Bezug auf die Leistung durchschnittlichen Schüler zu erwarten, dass er etwa 50 % der Aufgaben des jeweiligen Tests löst. Andererseits sind genügend viele Aufgaben vorhanden, um die Testleistungen von schwächeren bzw. leistungsstärkeren Schülern differenziert betrachten zu können. Bei der Aufgabenauswahl werden neben den Aufgabenschwierigkeiten (vermittelt über den prozentualen Anteil an Schülern, die die jeweiligen Aufgaben lösen) auch weitere teststatistische Kriterien, zeitliche Vorgaben und vor allem fachdidaktische Aspekte berücksichtigt.

Auf Klassen- bzw. Schulebene dienen die Ergebnisse der Vergleichsarbeiten in erster Linie den Schulen selbst,

damit diese für eine gezielte Schul- und Unterrichtsentwicklung genutzt werden können.²³ Zudem können die so gewonnenen Leistungsdaten auf Landesebene insgesamt betrachtet werden.

Exemplarische Auswertung erlaubt Rückschlüsse auf Abhängigkeit der Leistungen von verschiedenen Kontextmerkmalen

Eine detaillierte Ergebnisdarstellung aller 15 Vergleichsarbeiten ist im Rahmen der Bildungsberichterstattung nicht möglich. Exemplarisch werden die Ergebnisse der 9. Jahrgangsstufe Realschule und Gymnasium vorgestellt, da hier zusätzlich zu Deutsch und Mathematik auch Resultate aus den modernen Fremdsprachen betrachtet werden können. Bezüglich der Interpretation der berichteten Ergebnisse aus beiden Schularten muss beachtet werden, dass sich die Testergebnisse von Real- und Gymnasialschülern nicht in Bezug auf die Höhe der jeweiligen Testleistungen miteinander vergleichen lassen. Da die Vergleichsarbeiten die Schülerleistungen in Bezug auf die schulartspezifischen Bildungsstandards überprüfen und die Lösungshäufigkeiten speziell für die Testauf-

²³ Da in diesem Bericht der Fokus auf den landesweiten Daten liegt, kann an dieser Stelle nicht auf den in hohem Maße wichtigen Aspekt der Nutzung der Daten in den Klassen/ Schulen eingegangen werden. Auf dem Landesbildungsserver unter <http://www.schule-bw.de/entwicklung/dva/vadva/konzeption-dva/> sind unter anderem die Ziele der Vergleichsarbeiten und Hinweise zum Umgang mit den Ergebnissen in der Klasse/ Schule dargestellt.

gaben der Realschüler bzw. Gymnasiasten bestimmt werden, sind sowohl die Inhalte als auch die Ergebnisse der Vergleichsarbeiten verschiedener Schularten nicht miteinander vergleichbar. Die gemeinsame Betrachtung nimmt ausschließlich bestimmte Trends und Fragestellungen in den Blick, beispielsweise die Abhängigkeit der Leistungen von bestimmten Kontextmerkmalen wie Geschlecht, Migrationshintergrund, Beherrschung der deutschen Sprache und Noten des vorangegangenen Jahreszeugnisses. Diese Informationen wurden als Angaben bzw. Einschätzungen durch die Lehrkraft erhoben.

Die an der Pilotierungsphase beteiligten Schulen übersenden dem Landesinstitut für Schulentwicklung auch ihre DVA-Ergebnisse aus dem Pflichteinsatz in anonymisierter Form für weitere Auswertungen. Aus den Ergebnismeldungen der Realschulen liegen für das Fach Deutsch Daten von 2 295 Schülern und für Mathematik Daten von 2 279 Schülern vor. Für Englisch können Daten von 2 308 Realschülern ausgewertet werden, während für das Fach Französisch lediglich Daten von fünf Realschulen und insgesamt 99 Schülern zurückgemeldet wurden. Diese geringen Zahlen sind auf die geringe Gesamtanzahl von Realschulen in Baden-Württemberg (etwa zehn Schulen) mit der Pflichtfremdsprache Französisch zurückzuführen. Aus diesem Grund wird nachfolgend auf eine Darstellung der Ergebnisse für die Vergleichsarbeiten im Fach Französisch an den Realschulen verzichtet.

Für die Vergleichsarbeiten am Gymnasium liegen für das Fach Deutsch Daten von 3 140 Schülern und für Mathematik Daten von 3 094 Schülern vor. In Englisch sind dies etwa ein Drittel weniger (1 919), für Französisch etwa zwei Drittel weniger (1 024), da jedes Gymnasium für die einzelnen Klassen festlegt, in welcher Fremdsprache eine Vergleichsarbeit geschrieben wird.

Die Geschlechterverteilung ist bei beiden Schularten in etwa ausgewogen. An den Realschulen beträgt der Anteil der Mädchen pro Fach jeweils 50 %. An den Gymnasien sind tendenziell etwas mehr Mädchen zu finden, die Angaben schwanken je nach Test zwischen 50 % und 54 %. Der Anteil von Schülern mit Migrationshintergrund ist an Realschulen insgesamt etwas höher als an Gymnasien. So beträgt der Anteil von Realschülern mit Migrationshintergrund über alle Fächer hinweg 19 % bzw. 20 %. Am Gymnasium liegt dieser Anteil in Mathematik, Deutsch und Englisch hingegen zwischen 13 % und 16 %, in Französisch mit 10 % noch etwas darunter. Der Anteil von Schülern, welche nach Einschätzung der Lehrkräfte die deutsche Sprache »sehr gut« beherrschen, schwankt bei beiden Schularten je nach Fach. So beherrschen laut Einschätzung der Deutsch-Lehrkräfte und der Englisch-Lehrkräfte 17 % bzw. 16 % aller Realschüler die deutsche Sprache »sehr gut«. Demgegenüber trauen die Mathematik-Lehrkräfte 22 % der Realschüler eine »sehr gute« Sprachbeherrschung zu. Am Gymnasium wird die Be-

herrschung der deutschen Sprache durch die Deutsch- und Französisch-Lehrkräfte am niedrigsten bewertet. Etwa 15 % bzw. 13 % aller Gymnasiasten beherrschen nach deren Einschätzung die deutsche Sprache »sehr gut«, während Englisch- und Mathematik-Lehrkräfte 20 % bzw. 21 % ihrer Schüler eine »sehr gute« Sprachbeherrschung zuschreiben. Möglicherweise ist diese fächerbezogene Abweichung zumindest in Teilen auf die Tatsache zurückzuführen, dass Lehrkräfte in sprachbezogenen Fächern die Beurteilung in Bezug auf die Sprachbeherrschung etwas strenger vornehmen als andere Lehrkräfte, beispielsweise des Fachs Mathematik.

Mittlere Lösungshäufigkeiten variieren leicht um den angestrebten Wert von 50 %

Bei der Analyse der mittleren Lösungshäufigkeiten für die hier betrachteten Vergleichsarbeiten zeigt sich, dass sowohl Real- als auch Gymnasialschüler teilweise etwas schlechtere Leistungen erbringen, als aufgrund der Pilotierungsergebnisse zu erwarten war. An der Realschule beträgt die mittlere Lösungshäufigkeit für Deutsch und Englisch jeweils 49 %, für Mathematik 48 %. Die mittlere Lösungshäufigkeit am Gymnasium liegt für Deutsch bei 45 %, für Mathematik bei 47 % und für Französisch bei 48 %, während die Leistungen im Fach Englisch im Mittel bei den zu erwartenden 50 % liegen. Dieses insgesamt etwas schlechtere Abschneiden ist vermutlich auf die Umstellung des Erhebungsmodus zurückzuführen. Erstmals in diesem Durchgang fand die entsprechende Pilotierung unmittelbar vor den Sommerferien statt und die Vergleichsarbeiten selbst unmittelbar nach den Sommerferien des Folgejahres, im Durchgang zuvor wurden beide Erhebungen jeweils nach den Sommerferien durchgeführt.

Breite Streuung der Testergebnisse

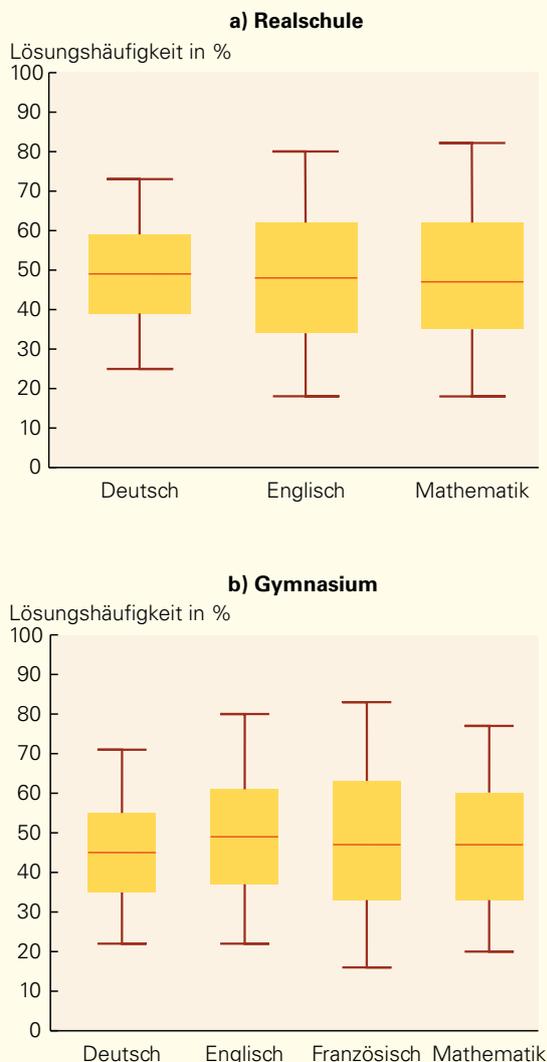
In Grafik I 3.2 (G1) sind die Ergebnisse der einzelnen Tests für Realschulen und Gymnasien entsprechend ihrer prozentualen Verteilung dargestellt. In der Mitte des jeweiligen Boxplots ist der Median²⁴ eingezeichnet. Dies

24 Der Median wird bei der Beschreibung verteilungsbezogener Merkmale verwendet, da er relativ robust gegenüber Extremwerten ist. Im Falle der DVA-Testwerte stimmen Median und Mittelwert zum Teil genau überein. An der Realschule besteht in Deutsch kein Unterschied zwischen Median und Mittelwert (49 %), in Mathematik und Englisch beträgt die Abweichung jeweils 1 Prozentpunkt (Mathematik: Median 48 %, Mittelwert 49 %, Englisch: Median 47 %, Mittelwert 48 %). Für das Gymnasium sind in Mathematik und Deutsch beide Werte identisch (Deutsch: 45 %, Mathematik: 47 %). In Englisch lag der Median bei 49 % (Mittelwert 50 %), in Französisch lag der Median bei 47 % (Mittelwert 48 %).

I 3.2 (G1)

DVA-Testergebnisse an der Realschule und am Gymnasium in Baden-Württemberg 2010

Bildungsstandards Klasse 8*), Fächer Deutsch, Englisch, Französisch**) und Mathematik



*) Getestet Anfang Klasse 9. - **) Nur Gymnasium.

Anmerkung: Die Höhe der jeweiligen Testleistungen ist zwischen beiden Schularten nicht vergleichbar

Datenquelle: Landesinstitut für Schulentwicklung.

Statistisches Landesamt Baden-Württemberg

300 11

bedeutet, dass die Hälfte der Schüler eine Leistung zeigt, die maximal dem Wert des Medians entspricht, die andere Hälfte erreicht Werte, die darüber liegen. Weiterhin markiert die »Box« jeweils die Grenze zum unteren bzw. zum oberen Leistungsviertel, wobei innerhalb der Box 50 % der Schüler mit ihren Leistungswerten repräsentiert sind.

Nach oben und unten schließen sich an die Box Anschlussstriche, sogenannte »Whiskers«, an. Deren Enden markieren in dieser Darstellung unten den

Leistungswert, den die 5 % der schwächsten Schüler im Test maximal erreichen und oben den Leistungswert, den die 5 % der besten Schüler noch übertreffen. Am Beispiel des Deutsch-Tests an der Realschule soll dies näher erläutert werden. Die Hälfte der Schüler löst mehr als 39 % und maximal 59 % der Aufgaben. Ein weiteres Viertel der Schüler löst mehr als 59 % der Aufgaben und repräsentiert damit das obere Leistungsviertel, während entsprechend das untere Leistungsviertel der Schüler eine maximale prozentuale Lösungshäufigkeit von 39 % erreicht. Weiterhin lösen die leistungsstärksten 5 % der Schüler über 73 % der Aufgaben, während im entsprechenden unteren Abschnitt der Leistungsverteilung maximal 25 % der Aufgaben gelöst werden.

Im Vergleich zu den anderen Tests zeigt sich in Deutsch sowohl an Realschulen als auch an Gymnasien, dass viele Schüler einen Wert erreichen, der nahe am Median liegt. In allen anderen Fächern weisen die Testwerte hingegen eine größere Streuung auf. So lösen 90 % der Realschüler in Deutsch über 25 % und maximal 73 % der Aufgaben, während die 5 % leistungsschwächsten Schüler in Englisch und Mathematik jeweils nur maximal 18 % der Aufgaben lösen, die 5 % leistungsstärksten Schüler jedoch Lösungshäufigkeiten von mehr als 80 % (Englisch) bzw. 82 % (Mathematik) erreichen (Grafik I 3.2 (G1)).

Auch am Gymnasium weisen die Testleistungen der Schüler in den Fächern Englisch und Mathematik eine breitere Streuung auf als in Deutsch. So erbringen die 5 % leistungsschwächsten Schüler in Deutsch noch eine Testleistung von maximal 22 %, die leistungsstärksten 5 % der Gymnasialschüler lösen mehr als 71 % der Aufgaben. In Englisch und in Mathematik lösen 90 % der Schüler hingegen über 22 % und maximal 80 % (Englisch) bzw. über 20 % und maximal 77 % (Mathematik) der Aufgaben. Insbesondere im Fach Französisch zeigt sich eine breite Streuung der Lösungshäufigkeiten, vor allem oberhalb des Medians; die Leistungen innerhalb der Schülerschaft erscheinen hier also recht heterogen. Die besten 5 % der Schüler lösen mehr als 83 % der Französisch-Aufgaben, während andererseits maximal 16 % der Aufgaben von den schwächsten 5 % bewältigt werden. Umgekehrt bedeutet dies, dass 90 % der Schüler Lösungshäufigkeiten zeigen, die zwischen diesen Werten liegen (Grafik I 3.2 (G1)).

Unterschiedliche Testleistungen von Mädchen und Jungen in den Sprachen und in Mathematik

Wie in vielen weiteren Schülerleistungsstudien lässt sich auch bei den DVA an der Realschule und am Gym-

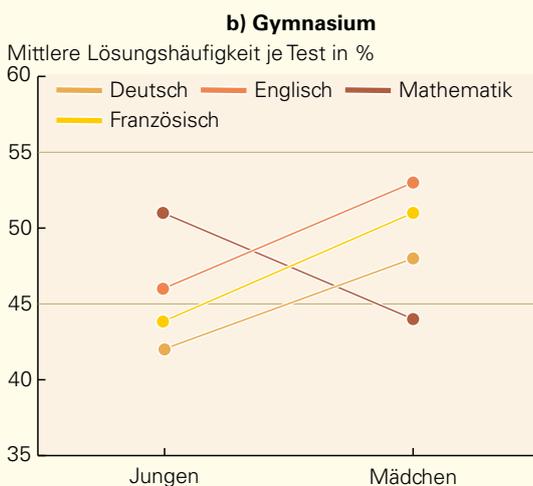
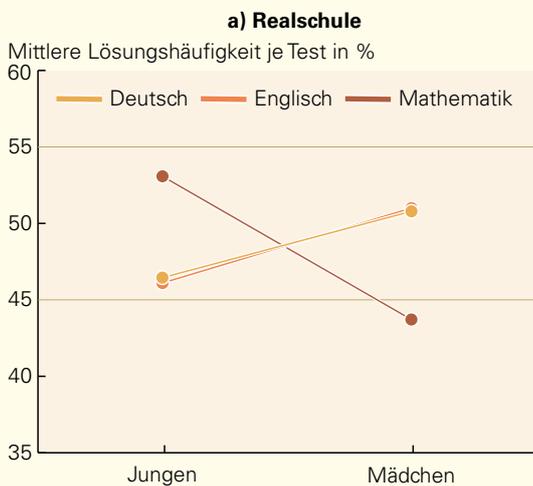
nasium ein geschlechtsspezifischer Effekt feststellen. Mädchen zeigen in den sprachbezogenen Fächern höhere Leistungen, während Jungen in Mathematik besser abschneiden (Grafik I 3.2 (G2)).

Bei Realschülern schwankt die Größe der Geschlechtsunterschiede je nach Test. Der Leistungsvorsprung der Jungen in Mathematik beträgt 9 Prozentpunkte, während Mädchen an der Realschule in Englisch und Deutsch im Schnitt 5 bzw. 4 Prozentpunkte höher liegen.

I 3.2 (G2)

Mittlere DVA-Lösungshäufigkeit der Schüler an der Realschule und am Gymnasium in Baden-Württemberg 2010 nach Geschlecht

Bildungsstandards Klasse 8*), Fächer Deutsch, Englisch, Französisch**) und Mathematik



*) Getestet Anfang Klasse 9. - **) Nur Gymnasium.
 Anmerkung: Die Höhe der jeweiligen Testleistungen ist zwischen beiden Schularten nicht vergleichbar
 Datenquelle: Landesinstitut für Schulentwicklung.

Die Leistungsunterschiede am Gymnasium betragen für alle vier Fächer konstant 6 bis 7 Prozentpunkte. In Deutsch lösen Mädchen im Durchschnitt 48 % der Aufgaben, Jungen 42 %. Entsprechende Unterschiede können ebenfalls bei den Fremdsprachen festgestellt werden. In Mathematik hingegen beträgt die mittlere Lösungshäufigkeit für Jungen 51 % und für Mädchen 44 %.

Migrationshintergrund hat nur zum Teil Auswirkungen auf die Testleistungen

Auch der Migrationshintergrund spielt bei den erzielten Leistungen eine Rolle, jedoch nicht in jedem Fach gleichermaßen. So beträgt im Fach Deutsch an der Realschule der Unterschied in den durchschnittlichen Testleistungen zwischen Schülern mit bzw. ohne Migrationshintergrund 5 Prozentpunkte (Grafik I 3.2 (G3)). In Englisch ist hingegen keine Abweichung vorhanden.

Am Gymnasium zeigen sich in Deutsch und in Französisch Leistungsunterschiede von 6 bzw. 4 Prozentpunkten in den mittleren Lösungshäufigkeiten zwischen Schülern mit und ohne Migrationshintergrund, während sich in Englisch ebenfalls fast kein Leistungsunterschied feststellen lässt (1 Prozentpunkt).

Auffällig ist, dass sowohl an der Realschule als auch am Gymnasium gerade für die Mathematikleistungen der Unterschied am größten ist. Realschüler ohne Migrationshintergrund erreichen eine im Durchschnitt um 10 Prozentpunkte bessere Testleistung als solche, die einen Migrationshintergrund aufweisen. An Gymnasien unterscheiden sich die durchschnittlichen Mathematikleistungen je nach Migrationshintergrund um 8 Prozentpunkte. Eine inhaltliche Sichtung des Mathematiktests zeigt, dass an Realschulen die drei Bereiche Messen, Modellieren und funktionale Zusammenhänge im Test 2010 als Schwerpunktbereiche repräsentiert waren. Inhalt der Vergleichsarbeit an Gymnasien waren die Bereiche Daten und Zufall sowie Vernetzung und Modellieren. Aufgaben mit diesen Kompetenzen bedürfen häufig einer textlichen Einbettung, wobei die Komplexität solcher Erläuterungen zum Teil hoch ist, da in kurzen Texten alle für die Lösung relevanten Informationen enthalten sein müssen. Dieser Umstand hat vermutlich die schon im Fach Deutsch festgestellten Unterschiede für das Fach Mathematik verstärkt.

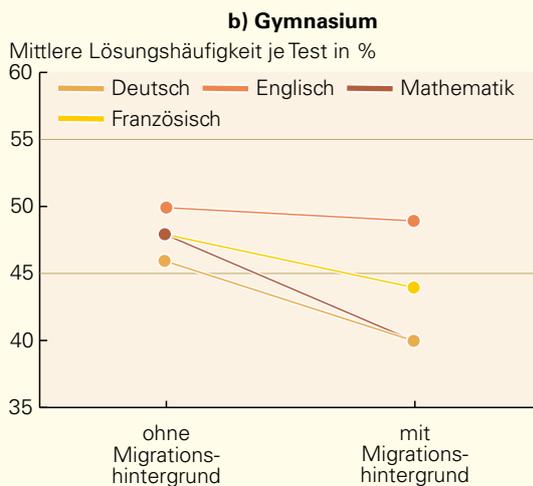
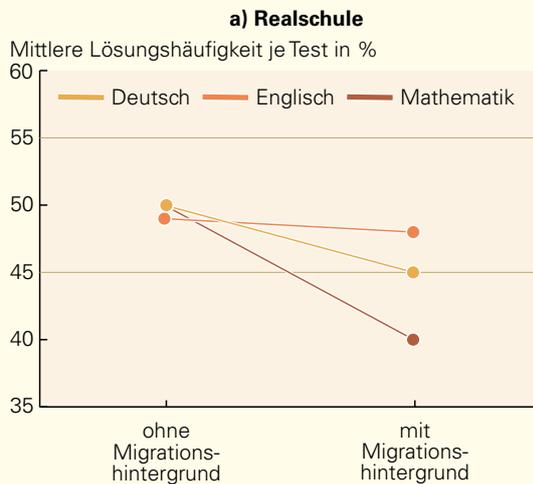
Kontextmerkmale Geschlecht und Migrationshintergrund wirken sich auf die Leistungen in den einzelnen Fächern unterschiedlich aus

In Grafik I 3.2 (G4) sind die Ergebnisse nach den beiden Kontextmerkmalen Geschlecht und Migrationshinter-

I 3.2 (G3)

Mittlere DVA-Lösungshäufigkeit der Schüler an der Realschule und am Gymnasium in Baden-Württemberg 2010 nach Migrationshintergrund

Bildungsstandards Klasse 8*), Fächer Deutsch, Englisch, Französisch**) und Mathematik



*) Getestet Anfang Klasse 9. - **) Nur Gymnasium.
Anmerkung: Die Höhe der jeweiligen Testleistungen ist zwischen beiden Schularten nicht vergleichbar
Datenquelle: Landesinstitut für Schulentwicklung.

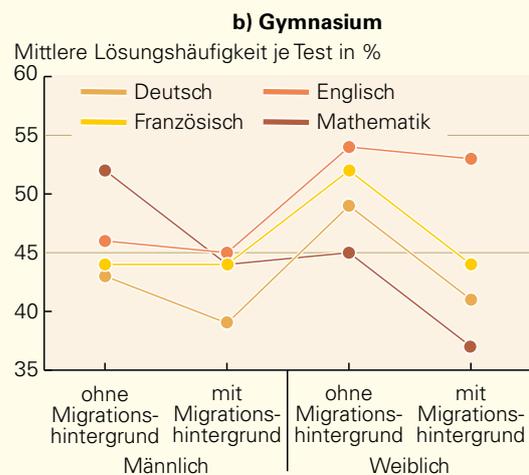
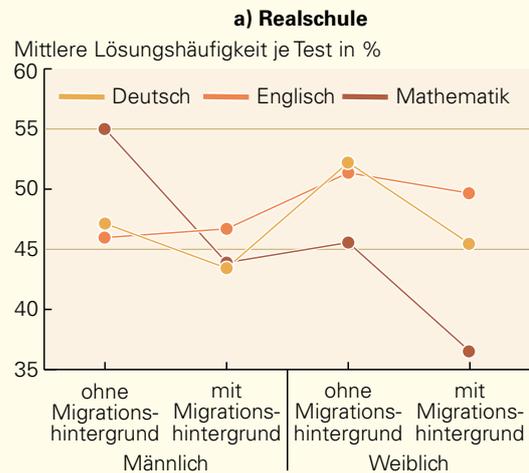
grund differenziert dargestellt. Insgesamt zeigen sich für jeden Test spezifische, jedoch für beide Schularten vergleichbare Zusammenhänge. Im Fach Deutsch weisen sowohl an Realschulen als auch an Gymnasien Mädchen ohne Migrationshintergrund bessere Leistungen auf als die übrigen Gruppen. Die Leistungen von Jungen mit Migrationshintergrund fallen dabei im Vergleich zu den anderen Gruppen noch etwas weiter ab, so erreichen sie an der Realschule im Vergleich zu Mädchen ohne Migrationshintergrund im Durchschnitt ein um 9 Prozentpunkte schwächeres Ergebnis; am

Gymnasiumweichendie durchschnittlichen Leistungen dieser beiden Gruppen um 10 Prozentpunkte voneinander ab. Im Fach Englisch hingegen zeigt sich bei beiden Schularten auch bei der gemeinsamen Betrachtung beider Kontextvariablen lediglich ein Geschlechtseffekt. Bezogen auf die Vergleichsarbeit in Französisch am Gymnasium zeigt sich eine ähnliche Verteilung wie bei der Vergleichsarbeit im Fach Deutsch: Mädchen ohne Migrationshintergrund erreichen auch hier im Vergleich zu den übrigen Gruppen eine bessere Leistung. Allerdings ist für Jungen kein Einfluss des Migra-

I 3.2 (G4)

Mittlere DVA-Lösungshäufigkeit der Schüler an der Realschule und am Gymnasium in Baden-Württemberg 2010 nach Geschlecht und Migrationshintergrund

Bildungsstandards Klasse 8*), Fächer Deutsch, Englisch, Französisch**) und Mathematik



*) Getestet Anfang Klasse 9. - **) Nur Gymnasium.
Anmerkung: Die Höhe der jeweiligen Testleistungen ist zwischen beiden Schularten nicht vergleichbar
Datenquelle: Landesinstitut für Schulentwicklung.

tionshintergrunds auf die Französischleistung feststellbar.

Im Fach Mathematik lässt sich bei beiden Schularten der größte Unterschied feststellen: So haben an der Realschule Jungen ohne Migrationshintergrund einen Leistungsvorsprung von durchschnittlich 18 Prozentpunkten im Vergleich zu Mädchen mit Migrationshintergrund. Am Gymnasium unterscheiden sich die durchschnittlichen Mathematikleistungen von Jungen ohne Migrationshintergrund und Mädchen mit Migrationshintergrund um 15 Prozentpunkte. Mädchen ohne Migrationshintergrund und Jungen mit Migrationshintergrund erbringen hingegen sowohl an Realschulen als auch an Gymnasien im jeweiligen Mathematiktest fast dieselben Leistungen.

Starker Effekt der Sprachbeherrschung

Die Einschätzung der Lehrkraft, inwiefern ein Schüler eine »sehr gute«, »normale« oder »schwache« Beherrschung der deutschen Sprache aufweist und die tatsächliche Leistung der Betreffenden in den Tests weisen sowohl an Realschulen als auch an Gymnasien unabhängig vom Fach einen starken Zusammenhang auf (Grafik I 3.2 (G5)).

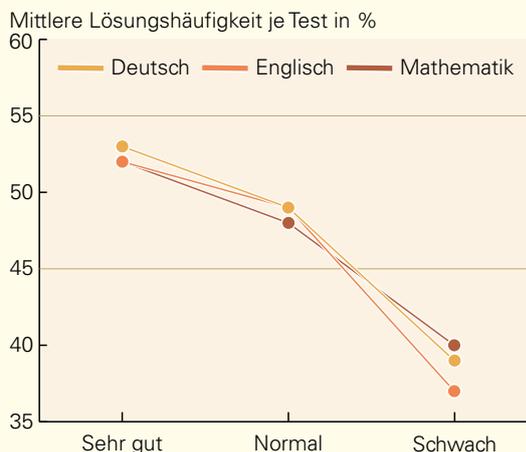
In den sprachbezogenen Fächern zeigen sich große Leistungsunterschiede zwischen Schülern, deren Sprachbeherrschung von den Lehrkräften als »sehr gut« bewertet wird und Schülern, denen eine »schwache« Sprachbeherrschung zugeschrieben wird. Die Lösungsquote der Realschüler mit einer als »sehr gut« eingeschätzten Sprachbeherrschung liegt im Fach Deutsch durchschnittlich 14 Prozentpunkte über der von Realschülern mit »schwacher« Sprachbeherrschung. Im Fach Englisch an der Realschule unterscheidet sich eine »sehr gute« Sprachbeherrschung im Vergleich zu einer »schwachen« Sprachbeherrschung um durchschnittlich 15 Prozentpunkte. Am Gymnasium sind die Unterschiede im Fach Französisch am größten. Eine sehr gute Beherrschung der deutschen Sprache geht mit einer mittleren Testleistung von 58 % gelösten Aufgaben einher, während Gymnasialschüler, denen eine schwache Sprachbeherrschung zugeschrieben wird, im Mittel nur 36 % der Aufgaben lösen. Für Englisch und Deutsch zeigen sich am Gymnasium ebenfalls große Unterschiede zwischen den einzelnen Gruppen: Eine »sehr gute« Sprachbeherrschung geht hier mit einer durchschnittlichen Testleistung von 57 % (Englisch) bzw. 53 % (Deutsch) einher, während Gymnasialschüler mit »schwacher« Sprachbeherrschung durchschnittlich lediglich 38 % (Englisch) bzw. 36 % (Deutsch) der Aufgaben lösen.

I 3.2 (G5)

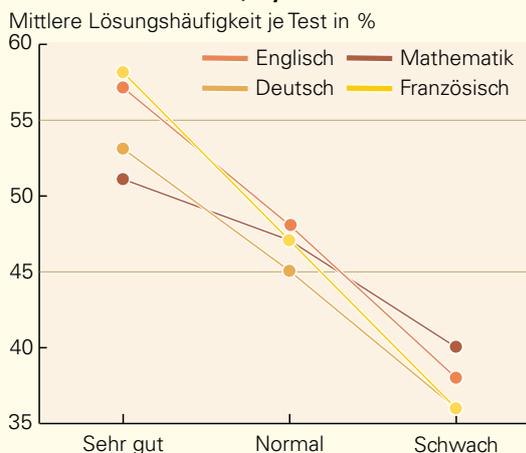
Mittlere DVA-Lösungshäufigkeit der Schüler an der Realschule und am Gymnasium in Baden-Württemberg 2010 nach Beherrschung der deutschen Sprache

Bildungsstandards Klasse 8*), Fächer Deutsch, Englisch, Französisch**) und Mathematik

a) Realschule



b) Gymnasium



*) Getestet Anfang Klasse 9. – **) Nur Gymnasium.

Anmerkung: Die Höhe der jeweiligen Testleistungen ist zwischen beiden Schularten nicht vergleichbar

Datenquelle: Landesinstitut für Schulentwicklung.

Auch für die Mathematikleistung in den Vergleichsarbeiten beider Schularten hängt eine bessere Beherrschung der deutschen Sprache mit einer höheren Leistung zusammen, allerdings in etwas geringerem Ausmaß. Die Lösungsquote sowohl der Realschüler als auch der Gymnasialschüler mit »sehr guter« Sprachbeherrschung liegt um 11 Prozentpunkte über der Lösungsquote derjenigen Schüler, denen eine unterdurchschnittliche Beherrschung der deutschen Sprache von den Mathematiklehrkräften zugeschrieben wird.

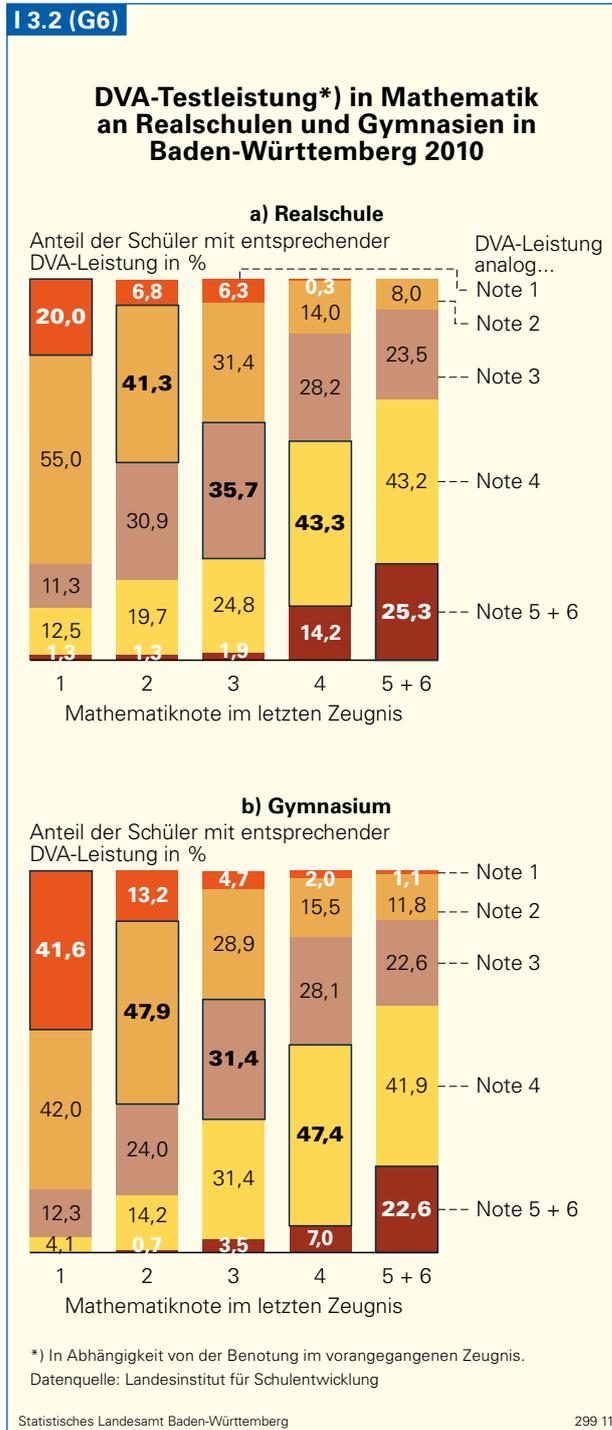
Mittlere Übereinstimmung von Testleistungen und Schulnoten

Exemplarisch soll am Beispiel der Mathematiktests für Realschule und Gymnasium erläutert werden, wie sich der Zusammenhang der Testleistung mit den Schulnoten im vorangegangenen Jahreszeugnis der 8. Klassenstufe darstellt. In Tabelle 13.2 (T2) im Anhang finden sich Informationen zur Notenverteilung in den hier analysierten repräsentativen Stichproben von Real- und Gymnasialschülern.

Für die Darstellungen in Grafik 1 3.2 (G6) wurden die Testleistungen der Real- und Gymnasialschüler zunächst in sechs Gruppen eingeteilt und zwar in der Form, dass die Schüleranteile bezogen auf die Notenverteilung auf die Leistungsverteilung des Tests übertragen wurden. Beispielsweise erhielten laut Auskunft der Lehrkräfte 3,5 % der Realschüler und 7,1 % der Gymnasialschüler im letzten Jahreszeugnis die Note 1 (Tabelle 1 3.2 (T2) im Anhang). Diese beiden Gruppen stellen also an Realschulen und Gymnasien jeweils den Anteil der leistungsstärksten Schüler bezogen auf die Noten dar. Dieser Anteil wurde auf die jeweiligen Testleistungen der Real- und Gymnasialschüler übertragen, sodass sich bezogen auf die Testleistung in der Kategorie »DVA-Leistung analog zu Note 1« bei der Realschule die 3,5 % leistungsstärksten Schüler und entsprechend beim Gymnasium die 7,1 % leistungsstärksten Schüler befinden. Analog wurde bei den weiteren Gruppen verfahren.

Für jede Schulnote ist in Grafik 1 3.2 (G6) dargestellt, wie sich die Testleistungen aller Realschüler und aller Gymnasialschüler, die diese Note erhalten haben, verteilen. Da die Note 6 in den Klassen der Stichprobe der Realschüler insgesamt nur dreimal und in der Stichprobe der Gymnasialschüler nur zweimal vergeben wurde, wurden die Noten 5 und 6 jeweils zusammengefasst. Anteile unterhalb der hervorgehobenen Säulenabschnitte der Grafik 1 3.2 (G6) beschreiben jeweils Leistungen, die schlechter sind, als aufgrund der letzten Zeugnisnote und der Notenverteilung insgesamt zu erwarten war. Anteile über diesen Säulenabschnitten kennzeichnen solche Leistungen, die diesbezüglich besser sind.

20 % der Realschüler, die im vorangehenden Jahreszeugnis die Note 1 erhalten hatten, erzielten eine vergleichbar gute Leistung im Mathematiktest (Grafik 1 3.2 (G6)). Bei den Gymnasialschülern fällt diese Übereinstimmung höher aus. So erreichten knapp 42 % der Gymnasialschüler, die im vorausgegangenen Jahr die Jahreszeugnisnote »sehr gut« im Fach Mathematik erhalten hatten, auch bei der Vergleichsarbeit in Mathematik eine entsprechende Leistung. Dieser unterschiedliche Grad an Übereinstimmung lässt sich vermutlich auf die Tatsache zurückführen, dass die Note 1 in der Stichprobe der Realschüler insgesamt nur selten



vergeben wurde (bei 3,5 % aller Realschüler), während der prozentuale Anteil der Gymnasialschüler, die die Note 1 erhalten hatten, mehr als doppelt so hoch ist (7,1 % der Gymnasialschüler). Daraus folgt, dass von den Realschülern, welche die Note 1 erhalten hatten, ein Fünftel gleichermaßen sehr gute Testleistungen zeigten. Auf die gesamte Stichprobe von Realschülern bezogen sind dies knapp 1 % der Schüler.²⁵ Entspre-

25 Ein Fünftel des Schüleranteils von 3,5 % mit Schulnote 1.

chend haben rund 42 % der Gymnasialschüler mit Note 1 in der Vergleichsarbeit sehr gut abgeschnitten. Dies entspricht einem Anteil von insgesamt knapp 3 % der Gymnasialschüler.²⁶

Setzt man alle jeweiligen Anteile aus Tabelle I 3.2 (T2) und Grafik I 3.2 (G6) in Bezug, so ist erkennbar, dass sich insgesamt bei der Stichprobe der Realschüler etwa 38 % der Bewertungen von Note und Testleistung in hohem Maße entsprechen. Bei der Stichprobe der Gymnasialschüler ergibt sich eine Übereinstimmung von insgesamt 40 %. Die Höhe der Übereinstimmung zwischen Testleistung und Note ist auch inhaltlich plausibel. Während in Notenleistungen noch weitere Beurteilungsaspekte einfließen, beispielsweise mündliche Leistung und individueller Leistungsfortschritt, ermöglicht die Vergleichsarbeit eine Momentaufnahme der schriftlichen Leistung bezogen auf Schwerpunktbereiche des Fachs. Hinzu kommt, dass Lehrkräfte sich bei der Notengebung am Leistungsstand der jeweiligen Klasse bzw. Schule orientieren, während sich die Kategorisierung der Leistungsdaten in Grafik I 3.2 (G6) auf die gesamte Schülerschaft bezieht. Eine

26 41 % des Schüleranteils von 7,1 % mit Schulnote 1.

gewisse Korrespondenz zwischen Noten und Testleistung kann und sollte folglich erwartbar sein, ein besonders hohes Ausmaß an Übereinstimmung jedoch nicht.

Beachtenswert sind an dieser Stelle allerdings die Abweichungen insbesondere in der Kategorie der Schüler mit den Jahreszeugnisnoten 5 und 6. Betrachtet man die Leistungsverteilungen für diese Kategorie, so zeigt sich bei den Testwerten der Realschüler, dass etwa 75 % der Schüler, die Note 5 oder 6 erhalten hatten, Leistungen zeigten, die besser zu bewerten wären, wenn man die Notenverteilung auf die Testleistung bezieht; 8 % erzielten eine Leistung, die demnach eher der Note 2 entspräche. Verstärkt findet sich dies am Gymnasium. So erbrachten etwa 13 % der Gymnasialschüler, die im vorangegangenen Jahreszeugnis die Note 5 oder 6 erhalten hatten, Leistungen, die nach der Notenverteilung mit der Note 1 und 2 zu bewerten wären. Auf Systemebene können solche Abweichungen der Testleistung von den erzielten Fachnoten im vorausgegangenen Jahreszeugnis unterschiedliche Ursachen haben. Auf Ebene des einzelnen Schülers liefern solche Auffälligkeiten jedoch wertvolle Hinweise, die von den betreffenden Lehrkräften genutzt werden können.