



Universität Zürich  
Institut für Bildungsevaluation

**Institut für Bildungsevaluation**  
Assoziiertes Institut  
der Universität Zürich

## **Check 5 2008**

Schlussbericht  
zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau

Florian Keller & Urs Moser  
Zürich, Juni 2008

## Inhalt

1	Einleitung .....	4
2	Das Wichtigste in Kürze .....	5
3	Fachleistungen der Schülerinnen und Schüler .....	7
3.1	Durchschnittliche Leistungen in Mathematik und Deutsch .....	7
3.2	Vergleich der durchschnittlichen Leistungen in den Teilbereichen der Mathematik .....	8
3.3	Vergleich der durchschnittlichen Leistungen in den Teilbereichen von Deutsch.....	9
3.4	Veränderungen in den Verteilungen der Schülerinnen und Schüler nach Kompetenzniveaus in der Mathematik .....	10
3.5	Veränderungen in den Verteilungen der Schülerinnen und Schüler nach Kompetenzniveau in Deutsch .....	11
3.6	Leistungsunterschiede zwischen Knaben und Mädchen in der Mathematik	12
3.7	Leistungsunterschiede zwischen Knaben und Mädchen in Deutsch.....	13
3.8	Leistungsunterschiede nach Erstsprache in der Mathematik .....	14
3.9	Leistungsunterschiede nach Erstsprache in Deutsch .....	15
3.10	Effekte der individuellen Lernvoraussetzungen auf die Leistungen der Schülerinnen und Schüler .....	16
3.11	Effekte der Klassenzusammensetzung auf die Leistungen der Schülerinnen und Schüler .....	18
4	Selbstreguliertes Lernen.....	20
4.1	Schulzufriedenheit und Prüfungsangst .....	20
4.2	Interesse und Selbstvertrauen in der Mathematik .....	21
4.3	Interesse und Selbstvertrauen in Deutsch .....	22

5	Unterrichtswahrnehmung .....	23
6	Beurteilung von Check 5 durch die Lehrpersonen.....	24
6.1	Beurteilung der Durchführung von Check 5 .....	24
6.2	Beurteilung der Testinstrumente von Check 5 .....	25
6.4	Art der getroffenen leistungsfördernden Massnahmen.....	28
6.5	Interesse an einer zweiten Durchführung von Check 5 .....	29
7	Fazit.....	30
	Anhang: Glossar der statistischen Begriffe .....	32
	Weiterführende Informationen .....	34

## **1 Einleitung**

Check 5 bietet Lehrpersonen eine freiwillige und unabhängige Standortbestimmung ihrer Schülerinnen und Schüler. Dazu werden einerseits die Leistungen der Schülerinnen und Schüler in Mathematik und Deutsch mit einem Test geprüft. Andererseits werden überfachliche Kompetenzen wie die Problemlösefähigkeit getestet und Einschätzungen zum selbstregulierten Lernen sowie zur Unterrichtswahrnehmung der Schülerinnen und Schüler erhoben.

An Check 5 können alle Lehrpersonen der 5. Klassen im Kanton Aargau teilnehmen. Wer sich für Check 5 anmeldet, verpflichtet sich, die Ergebnisse des Leistungstests und der Befragung der Schülerinnen und Schüler zu analysieren und im Unterricht fördernde und leistungsorientierte Massnahmen zu ergreifen. Check 5 bietet damit den Lehrerinnen und Lehrern eine Grundlage zur Reflexion über das eigene Handeln im Unterricht mit dem Ziel, die Unterrichtsqualität und die Förderung der Schülerinnen und Schüler weiterzuentwickeln und zu verbessern.

Check 5 wurde im Jahr 2004 vom Institut für Bildungsevaluation (IBE) im Auftrag des Departements Bildung, Kultur und Sport des Kantons Aargau entwickelt. Seither wird Check 5 jährlich Anfang September durchgeführt.

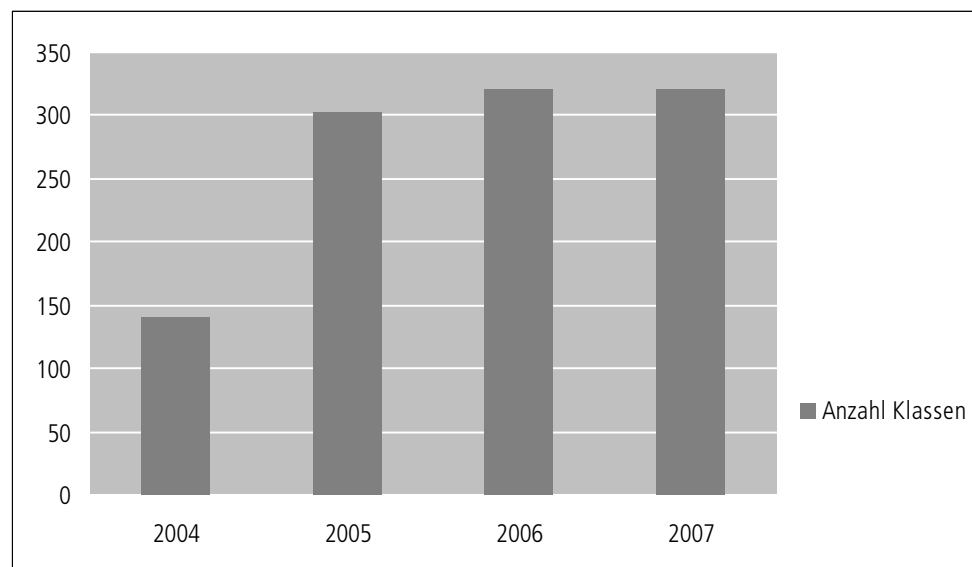
Der vorliegende Schlussbericht knüpft inhaltlich am Schlussbericht des Check 5 2007 an. Erneut werden die Ergebnisse der Schülerinnen und Schüler des Kantons Aargau anhand einiger zentraler Indikatoren verglichen. Die Ergebnisse der Fachleistungen, des selbstregulierten Lernens und der Unterrichtswahrnehmung der Schülerinnen und Schüler werden in eine vergleichbare Metrik gebracht und die Veränderungen seit 2004 als Trend dargestellt. Der Schwerpunkt des Berichts liegt somit beim Vergleich der Leistungsentwicklung der Schülerinnen und Schüler im Kanton Aargau in den letzten vier Jahren. Weitere Ergebnisse des Check 5 2007, unter anderem Analysen der unterschiedlichen Leistungen der Klassen und der Bezirke, sind im Zwischenbericht Check 5/2007 beschrieben.

Der Bericht ist in sieben Kapitel unterteilt. In Kapitel 2 wird kurz auf die Durchführung der Tests sowie auf das Vorgehen bei der Analyse der Ergebnisse eingegangen. In Kapitel 3 sind Trends für die Leistungen in den Fächern Mathematik und Deutsch und in ihren Teilbereichen dargestellt. Zudem werden der Einfluss der individuellen Lernvoraussetzungen auf die Fachleistungen sowie die Effekte der Klassenzusammensetzung auf die Leistungen in den Jahren 2004 bis 2007 analysiert. Im vierten Kapitel werden die Veränderungen im selbstregulierten Lernen dargestellt und im fünften Kapitel wird auf die Unterrichtswahrnehmung der Schülerinnen und Schüler eingegangen. In Kapitel 6 werden die Beurteilungen der Lehrpersonen zum Check 5 ausgewertet und es wird auf Veränderungen der Beurteilung zwischen den Jahren 2005 und 2007 hingewiesen. In Kapitel 7 werden in einem Fazit noch einmal die wichtigsten Resultate zusammengefasst.

## 2 Das Wichtigste in Kürze

*Teilnahme* – 2004 wurde Check 5 zum ersten Mal mit 140 ausgewählten Klassen durchgeführt. Im Herbst 2005 wurde Check 5 ein zweites Mal und ohne Beschränkung der Teilnehmerzahl angeboten. Von allen 5. Klassen im Kanton Aargau meldeten sich 302 Klassen mit 5394 Schülerinnen und Schülern für eine Teilnahme an. Damit nahmen rund 80 Prozent aller 5. Primarschulklassen des Kantons Aargau teil. Im Jahr 2006 stieg das Interesse an Check 5 noch einmal an. 320 Klassen mit 5618 Schülerinnen und Schülern oder rund 85 Prozent aller 5. Klassen im Kanton Aargau führten Check 5 freiwillig durch. Diese grosse Nachfrage hielt auch im Jahr 2007 weiter an. 2007 nahmen 321 Klassen mit 5530 Schülerinnen und Schülern teil. Dies entspricht einer Teilnahmequote von fast 90 Prozent aller 5. Primarschulklassen im Kanton Aargau (Abbildung 2.1).

Abbildung 2.1: Teilnehmende Klassen in den Jahren 2004, 2005 und 2006



*Durchführung* – Check 5 besteht aus einem Leistungstest in den Bereichen Mathematik, Deutsch und Problemlösen sowie aus einem Fragebogen zum selbstregulierten Lernen, zur Unterrichtswahrnehmung und zur sozioökonomischen Herkunft der Kinder. Tests und Befragung werden von den Lehrpersonen während der regulären Unterrichtszeit nach standardisierten Anweisungen durchgeführt. Die Durchführung beansprucht rund vier Stunden. Anschliessend werden die Tests von Mitarbeiterinnen und Mitarbeitern des Instituts für Bildungsevaluation (IBE) korrigiert, elektronisch erfasst und ausgewertet.

*Testinstrumente* – Die Tests zur Erfassung der fachlichen Leistungen werden vom IBE in Zusammenarbeit mit erfahrenen Lehrpersonen aus dem Kanton Aargau entwickelt. Jedes Jahr werden neue Testaufgaben erarbeitet und erprobt und so die Testinstrumente ständig weiterentwickelt und optimiert. Damit die Testresultate gleichwohl vergleichbar

bleiben, werden ausgewählte Aufgaben als «Link-Items» während mehreren Jahren eingesetzt.

*Kompetenzbereiche* – Die Testaufgaben prüfen die Fähigkeiten der Schülerinnen und Schüler in Mathematik und Deutsch. In den einzelnen Fächern können zudem die Fähigkeiten in Teilbereichen wie «Arithmetik», «Mathematisches Problemlösen» oder «Geometrie» beziehungsweise «Leseverständnis», «Sprachbetrachtung» oder «Wortschatz» ausgewiesen werden.

*Check-5-Skala* – Die Fähigkeiten der Schülerinnen und Schüler werden als Punktzahl auf der Check-5-Skala dargestellt. Basis für die Skala sind die Testergebnisse der Schülerinnen und Schüler im Jahr 2004. Die Check-5-Skala wurde so standardisiert, dass der Mittelwert der Schülerinnen und Schüler im Jahr 2004 in jedem Fach 500 Punkte und die Standardabweichung 100 Punkte beträgt. Diese Skalierung und Standardisierung erlaubt einerseits einen direkten Vergleich der Leistungen in den verschiedenen Testjahren unabhängig von der Stichprobe und unabhängig von der Art und der Anzahl der eingesetzten Testaufgaben. Andererseits erleichtert sie auch die Interpretation der Ergebnisse: Werte über 500 Punkte entsprechen einer höheren Kompetenz im Vergleich zur Stichprobe 2004, Werte unter 500 Punkte einer tieferen Kompetenz im Vergleich zur Stichprobe 2004.

Leistungsunterschiede – beispielsweise zwischen Mädchen und Knaben oder zwischen zwei Testjahren – sind dann relevant, wenn sie 20 Punkte oder mehr betragen. Unterschiede von 80 oder mehr Punkten sind sehr gross.

*Ergebnisrückmeldung* – Mitte November, rund sechs Wochen nach der Durchführung, wird den Lehrpersonen eine Ergebnisrückmeldung zugestellt. Die Ergebnisrückmeldung enthält die Testergebnisse sowohl für jedes Schulkind einzeln als auch für die gesamte Klasse sowie die anonymisierten Testergebnisse der anderen Klassen, die an Check 5 teilgenommen haben. So können die Lehrpersonen zum einen ihre Schülerinnen und Schüler im Verlauf des 5. Schuljahres gezielt dort fördern, wo es notwendig ist. Zum anderen können sie die Leistung ihrer Klasse mit den Leistungen der anderen Klassen vergleichen.

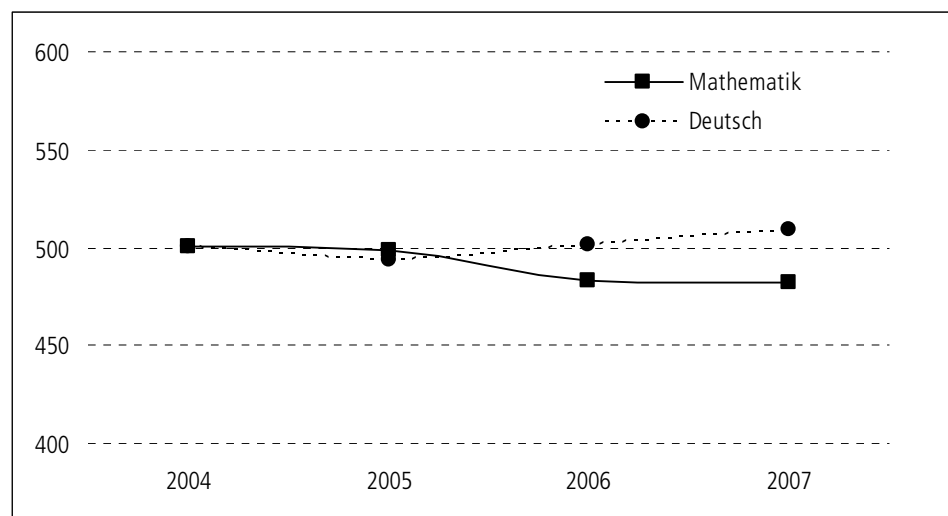
*Zwischenbericht* – Jeweils im November erstellt das IBE einen Zwischenbericht zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau. Im Zwischenbericht 2007 wurden die Ergebnisse des Check 5 2007 dargestellt und die Leistungsunterschiede zwischen den Schülerinnen und Schülern, zwischen den Klassen und zwischen den Bezirken des Kantons Aargau analysiert. Es zeigte sich, dass die Leistungsunterschiede zwischen den Schülerinnen und Schülern zu einem grossen Teil mit der sozialen Herkunft und der Kenntnis der Unterrichtssprache erklärt werden können. Auch zwischen den Klassen bestehen beträchtliche Leistungsunterschiede. Die Unterschiede zwischen den Klassen können insbesondere in den Bereichen «Geometrie», «Arithmetik», «Texte schreiben» und «Sprachbetrachtung» vor allem auf Merkmale des Unterrichts zurückgeführt werden. Zwischen den einzelnen Bezirken des Kantons Aargau hingegen konnten keine Leistungsunterschiede festgestellt werden. (Keller, F. & Moser, U. (2007). *Check 5. Zwischenbericht zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau*. Zürich: IBE.)

### 3 Fachleistungen der Schülerinnen und Schüler

#### 3.1 Durchschnittliche Leistungen in Mathematik und Deutsch

Abbildung 3.1 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Fächern Mathematik und Deutsch in den Jahren 2004 bis 2007. Die Mittelwerte in Mathematik sind als Quadrate, die Mittelwerte in Deutsch als Punkte dargestellt.

Abbildung 3.1: Leistungsentwicklung in den Fächern Mathematik und Deutsch



Anmerkung: Mathematik:

2004:  $n = 2537$ ;  $M = 500$ ;  $SE = 3,5$ ;  $SD = 97$   
2005:  $n = 5394$ ;  $M = 499$ ;  $SE = 2,0$ ;  $SD = 99$   
2006:  $n = 5618$ ;  $M = 483$ ;  $SE = 3,0$ ;  $SD = 102$   
2007:  $n = 5530$ ;  $M = 482$ ;  $SE = 3,9$ ;  $SD = 102$

Deutsch:

2004:  $n = 2537$ ;  $M = 500$ ;  $SE = 3,2$ ;  $SD = 96$   
2005:  $n = 5394$ ;  $M = 494$ ;  $SE = 2,7$ ;  $SD = 98$   
2006:  $n = 5618$ ;  $M = 502$ ;  $SE = 4,1$ ;  $SD = 102$   
2007:  $n = 5530$ ;  $M = 509$ ;  $SE = 3,7$ ;  $SD = 108$

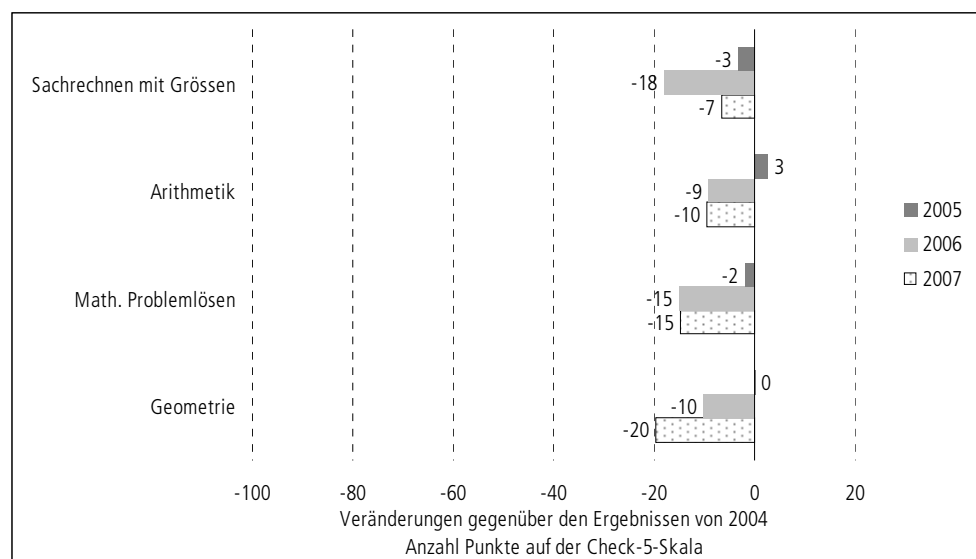
In Mathematik sinken die Leistungen seit dem Jahr 2004. Die grösste Differenz trat im Jahr 2006 auf, als die Mathematikleistungen im Vergleich zum Vorjahr um 16 Punkte gesunken waren. Im Jahr 2007 blieben die Leistungen in der Mathematik praktisch stabil (minus 1 Punkt). Die Leistungsdifferenz zwischen den Jahren 2004 und 2007 beträgt damit knapp 18 Punkte. Diese Differenz ist, obwohl statistisch signifikant, klein und der Effekt der Differenz ist nur schwach ( $d = 0.17$ ).

In Deutsch haben sich 2007 die Leistungen der Schülerinnen und Schüler weiter verbessert. Im Jahr 2007 wurden 509 Punkte erreicht. Das sind 7 Punkte mehr als 2006. Alles in allem bleiben die Veränderungen in den Deutschleistungen in den letzten drei Jahren aber vernachlässigbar klein.

### 3.2 Vergleich der durchschnittlichen Leistungen in den Teilbereichen der Mathematik

Abbildung 3.2 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Teilbereichen der Mathematik in den Jahren 2005, 2006 und 2007 als Differenz zu den Leistungen im Jahr 2004. Die Abweichungen der Stichprobe 2005 sind als dunkelgraue Balken, die Abweichungen der Stichprobe 2006 als hellgraue Balken und die Abweichungen der Stichprobe 2007 als grau gepunktete Balken dargestellt. Die angegebenen Werte entsprechen der Differenz zur Stichprobe 2004 in Punkten auf der Check-5-Skala.

Abbildung 3.2: Leistungsveränderungen in den Teilbereichen der Mathematik



Im Jahr 2005 übertreffen die Schülerinnen und Schüler im Teilbereich «Arithmetik» die Leistungen der Stichprobe 2004 um 3 Punkte. In den Bereichen «Sachrechnen mit Grössen» und «Mathematisches Problemlösen» erzielen sie leicht tiefere Ergebnisse. Im Bereich «Geometrie» zeigen sich keine Unterschiede zum Jahr 2004.

Im Jahr 2006 sind die Leistungen in allen Teilbereichen der Mathematik deutlich tiefer als die Vergleichswerte 2004 und 2005. Der geringsten Rückstand zeigt sich im Bereich «Arithmetik» (9 Punkte), der grösste Rückstand im Bereich «Sachrechnen mit Grössen» (18 Punkte).

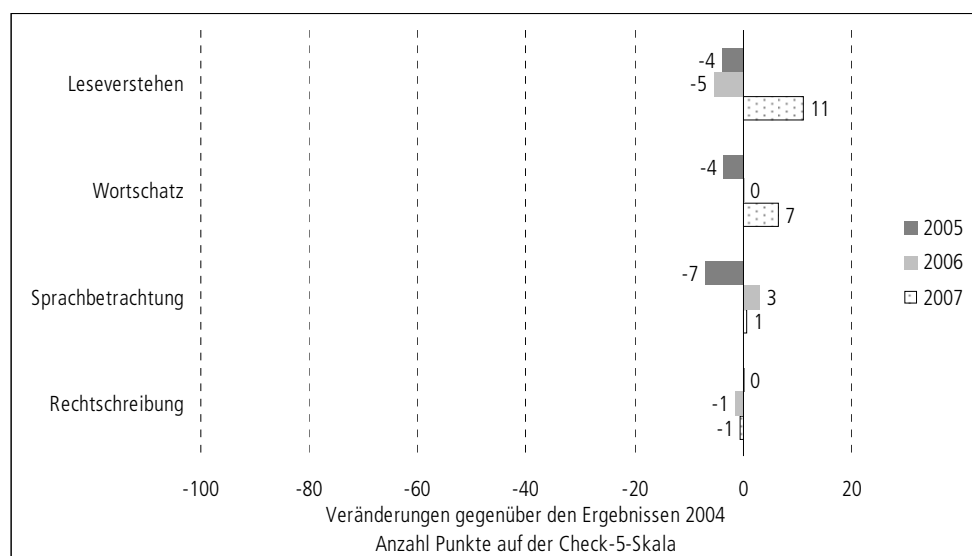
Im Jahr 2007 sind die Leistungen in allen Teilbereichen tiefer als in den Jahren 2004 und 2005. Im Vergleich zu den Leistungen im Jahr 2006 hat sich der Rückstand im Teilbereich «Sachrechnen mit Grössen» auf 7 Punkte reduziert. Im Teilbereich «Geometrie», der gemäss Lehrplan in der 4. Klasse noch nicht verbindlich ist, hat sich der Rückstand im Vergleich zum Jahr 2006 um 10 Punkte auf 20 Punkte erhöht.



### 3.3 Vergleich der durchschnittlichen Leistungen in den Teilbereichen von Deutsch

Abbildung 3.3 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Teilbereichen von Deutsch in den Jahren 2005, 2006 und 2007 als Differenz zu den Leistungen im Jahr 2004. Die Abweichungen der Stichprobe 2005 sind als dunkelgraue Balken, die Abweichungen der Stichprobe 2006 als hellgraue Balken und die Abweichungen der Stichprobe 2007 als grau gepunktete Balken dargestellt. Die angegebenen Werte entsprechen der Differenz zur Stichprobe 2004 in Punkten auf der Check-5-Skala.

Abbildung 3.3: Leistungsveränderungen in den Teilbereichen von Deutsch



Im Jahr 2005 werden in allen Teilbereichen von Deutsch mit Ausnahmen des Bereichs «Rechtschreibung» geringfügig tiefere Leistungen erbracht als im Jahr 2004. Der grösste Rückstand zeigt sich im Bereich «Sprachbetrachtung» (7 Punkte).

Im Jahr 2006 unterscheiden sich die Leistungen in den Teilbereichen von Deutsch ebenfalls kaum von den Vorjahren. Einzig im Teilbereich «Sprachbetrachtung» werden Leistungen erbracht, die 3 Punkte über den Leistungen von 2004 und 10 Punkte über den Leistungen von 2005 liegen.

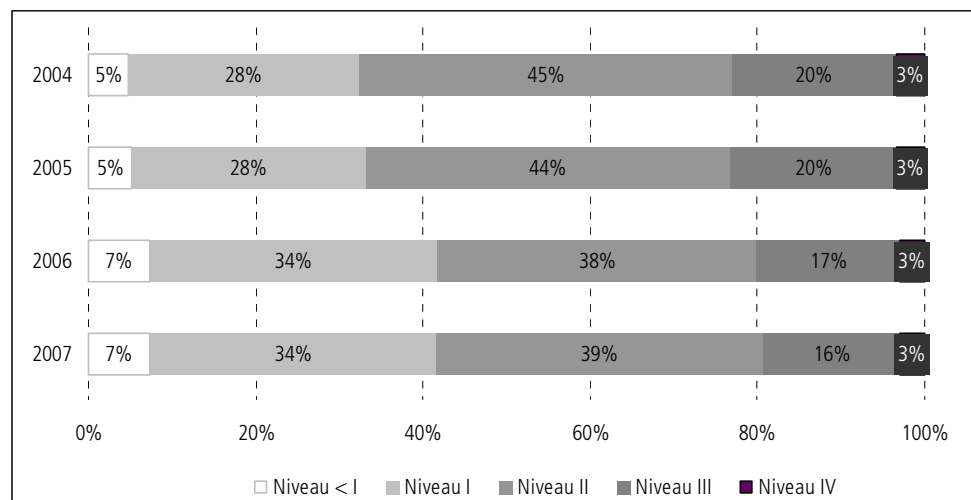
Im Jahr 2007 werden insbesondere in den Teilbereichen «Leseverstehen» und «Wortschatz» bessere Leistungen erbracht. Im «Leseverstehen», das dieses Jahr ausführlicher als in den Vorjahren getestet wurde, haben sich die Leistungen im Vergleich zum Jahr 2004 um 11 Punkte und im Vergleich zum Jahr 2006 um 16 Punkte verbessert. Alles in allem bleiben die Differenzen zwischen den einzelnen Testjahren von Check 5 aber gering.

### 3.4 Veränderungen in den Verteilungen der Schülerinnen und Schüler nach Kompetenzniveaus in der Mathematik

Abbildung 3.4 zeigt die Verteilung der Schülerinnen und Schüler in den Jahren 2004 bis 2007 nach den vier Kompetenzniveaus in der Mathematik.

Im Jahr 2007 erreichen rund 7 Prozent der Schülerinnen und Schüler das Kompetenzniveau I noch nicht. Das bedeutet, dass diese Schülerinnen und Schüler die Lernziele der 4. Klasse in der Mathematik in weiten Teilen noch nicht erfüllen. Rund 34 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau I. Sie können beispielsweise Additionen und Subtraktionen im Zahlenraum bis Zehntausend lösen oder eine einfache zweidimensionale Figur durch innere Vorstellung drehen. Rund 39 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau II. Sie können beispielsweise Multiplikationen mit zweistelligen Faktoren im Zahlenraum bis Zehntausend und Sachaufgaben, die zwei aufeinanderfolgende Rechenoperationen verlangen, lösen. Rund 16 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau III. Sie können beispielsweise Divisionen mit zweistelligen Divisoren im Zahlenraum bis Tausend lösen und dreidimensionale Figuren durch innere Vorstellung drehen. Rund 3 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau IV. Sie können beispielsweise komplexe Zahlenfolgen weiterführen und Sachaufgaben lösen, die das Erschliessen von Informationen aus Tabellen oder Abbildungen verlangen.

Abbildung 3.4: Verteilung der Schülerinnen und Schüler nach Kompetenzniveau und Jahr in der Mathematik



Im Jahr 2007 unterscheidet sich die Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in der Mathematik nicht von der Verteilung im Jahr 2006.

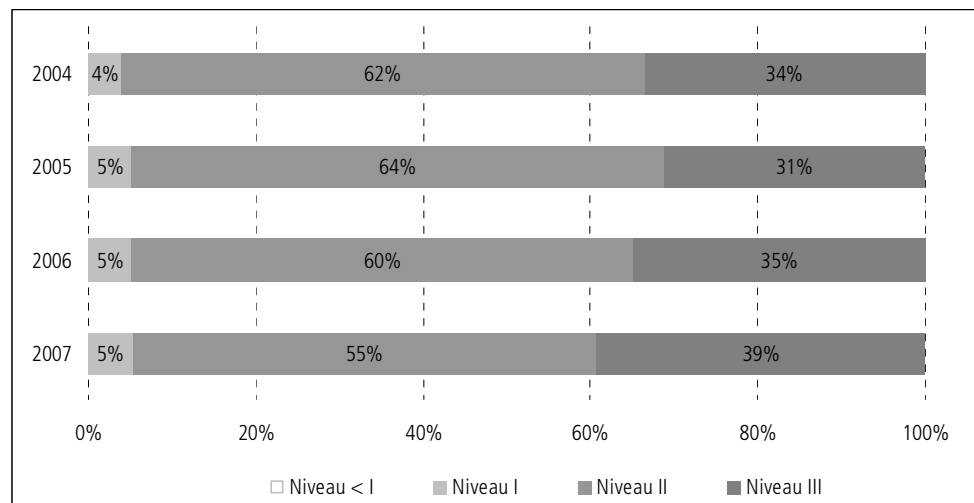
Im Vergleich zu den Verteilungen 2004 und 2005 erreichen 2007 weniger Schülerinnen und Schüler die Kompetenzniveaus II und III. Auch der Anteil Schülerinnen und Schüler, die die Lernziele der 4. Klasse noch nicht erreichen, hat sich seit dem Jahr 2004 leicht erhöht.

### 3.5 Veränderungen in den Verteilungen der Schülerinnen und Schüler nach Kompetenzniveau in Deutsch

Abbildung 3.5 zeigt die Verteilung der Schülerinnen und Schüler in den Jahren 2004 bis 2007 nach den drei Kompetenzniveaus in Deutsch.

Im Jahr 2007 erreichen sämtliche Schülerinnen und Schüler mindestens Kompetenzniveau I. Das heisst, alle Schülerinnen und Schüler erfüllen die Ziele der 4. Klasse in Deutsch. Rund 5 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau I. Sie können beispielsweise einfache Informationen in schriftlichen Texten oder in Hörtexten erfassen und können die Wortarten Nomen, Verben und Adjektive richtig bestimmen. Rund 55 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau II. Sie können beispielsweise Informationen aus Texten erfassen und mit anderen Informationen verbinden und erkennen die grammatikalischen Zeiten Präsens und Präteritum richtig. Rund 39 Prozent der Schülerinnen und Schüler erreichen Kompetenzniveau III. Sie können beispielsweise implizite Informationen in Texten erschliessen und erkennen die grammatikalische Zeit Perfekt richtig.

Abbildung 3.5: Verteilung der Schülerinnen und Schüler nach Kompetenzniveau und Jahr in Deutsch



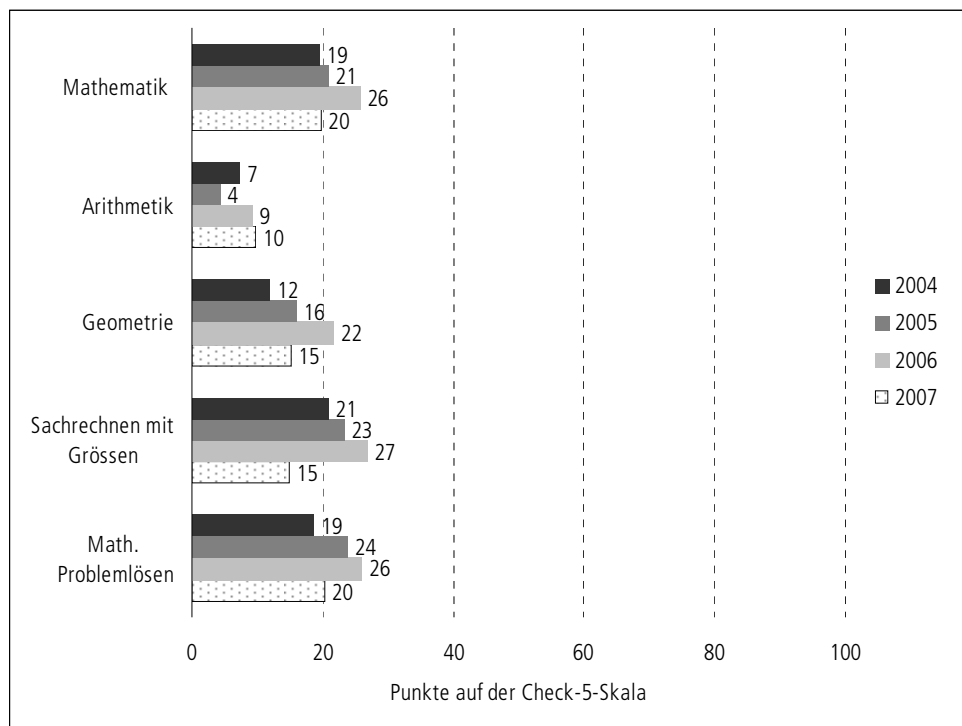
Der Anteil Schülerinnen und Schüler, die Kompetenzniveau III erreichen, ist im Jahr 2007 so gross wie noch nie (39 Prozent). Gleichzeitig reduziert sich der Anteil Kinder, die Kompetenzniveau II erreichen, im Vergleich zum Jahr 2006 um 5 Prozent, während der Anteil Schülerinnen und Schüler, die Kompetenzniveau I erreichen, seit 2004 weitgehend stabil bleibt.

### 3.6 Leistungsunterschiede zwischen Knaben und Mädchen in der Mathematik

Abbildung 3.6 zeigt den Leistungsvorsprung der Knaben in der Mathematik und in den Teilbereichen der Mathematik pro Jahr auf der Check-5-Skala.

Im Jahr 2004 betrug der Vorsprung der Knaben gegenüber den Mädchen durchschnittlich 19 Punkte (schwarzer Balken). Im Jahr 2005 stieg die Differenz zwischen Knaben und Mädchen auf 21 Punkte (dunkelgrauer Balken) und im Jahr 2006 auf 26 Punkte (hellgrauer Balken). Im Jahr 2007 betrug die Leistungsdifferenz in der Mathematik zwischen Knaben und Mädchen 20 Punkte (grau gepunkteter Balken). Das sind 6 Punkte weniger als im Jahr 2006. Das heisst, die Leistungen der Mädchen und der Knaben haben sich geringfügig angeglichen. Das heisst aber auch, dass die mathematischen Leistungen der Knaben zu Beginn der 5. Klasse im Durchschnitt immer noch 20 Punkte höher sind als die Leistungen der Mädchen. Dieser Vorsprung der Knaben ist mit einer Effektgrösse von  $d = 0.20$  als schwach zu beurteilen.

Abbildung 3.6: Leistungsvorsprung der Knaben in der Mathematik nach Jahr



Auch in den Teilbereichen der Mathematik reduzierte sich der Vorsprung der Knaben im Vergleich zum Jahr 2006. Die grösste Reduktion zeigt sich in den Teilbereichen «Geometrie» (minus 7 Punkte) und «Sachrechnen mit Grössen» (minus 12 Punkte).

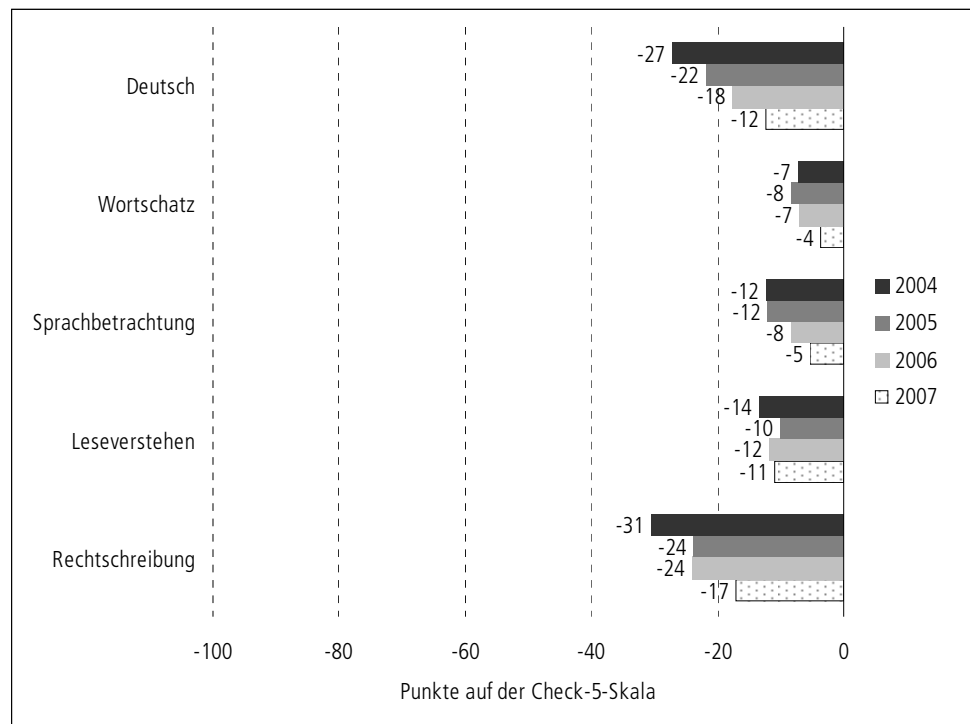
Im Jahr 2007 besteht der grösste Leistungsunterschied zwischen Mädchen und Knaben im Teilbereich «Mathematisches Problemlösen» (20 Punkte). Am wenigsten unterscheiden sich die Leistungen der Mädchen und Knaben im Teilbereich «Arithmetik» (10 Punkte).

### 3.7 Leistungsunterschiede zwischen Knaben und Mädchen in Deutsch

Abbildung 3.7 zeigt den Leistungsrückstand der Knaben in Deutsch und in den Teilbereichen von Deutsch pro Jahr auf der Check-5-Skala.

Seit dem Jahr 2004 werden die Leistungsunterschiede zwischen Mädchen und Knaben in Deutsch immer kleiner. Im Jahr 2004 betrug der Rückstand der Knaben gegenüber den Mädchen in Deutsch durchschnittlich 27 Punkte auf der Check-5-Skala. Dieser Rückstand reduzierte sich sukzessive auf 12 Punkte im Jahr 2007. Der Rückstand der Knaben ist mit einer Effektgrösse von  $d = 0.12$  als gering zu beurteilen.

Abbildung 3.7: Leistungsrückstand der Knaben in Deutsch nach Jahr



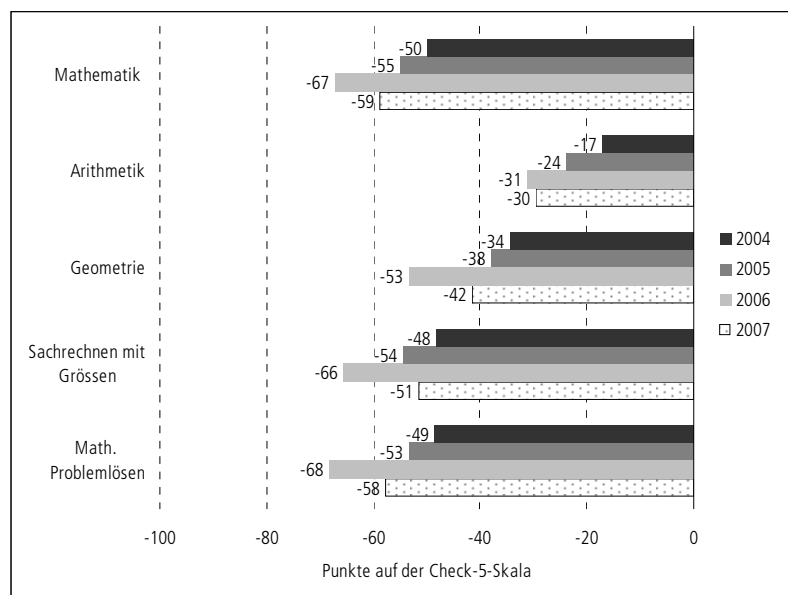
Auch in den meisten Teilbereichen von Deutsch reduzierten sich die geschlechtsspezifischen Unterschiede. So wiesen die Knaben 2004 beispielsweise im Teilbereich «Rechtschreibung» einen Rückstand von 31 Punkten auf. Bis 2007 verringerte sich dieser Rückstand um 14 auf 17 Punkte. In den Teilbereichen «Wortschatz» und «Sprachbetrachtung» unterschieden sich die Leistungen zwischen Knaben und Mädchen nicht mehr statistisch signifikant.

### 3.8 Leistungsunterschiede nach Erstsprache in der Mathematik

Abbildung 3.8 zeigt die Leistungsunterschiede zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache (DE) und Schülerinnen und Schülern mit Deutsch als Zweitsprache (DZ) in der Mathematik und in den Teilbereichen der Mathematik. Die Unterschiede sind dargestellt nach Testjahr auf der Check-5-Skala.

In den Jahren 2004 bis 2006 vergrösserte sich in der Mathematik der Rückstand der Schülerinnen und Schüler mit DZ gegenüber den Schülerinnen und Schülern mit DE kontinuierlich von 50 Punkte auf 67 Punkte. Im Jahr 2007 jedoch verringerte sich die Leistungsrückstand der Schülerinnen und Schüler mit DZ auf 59 Punkte (grau gepunkteter Balken). Das heisst, die mathematischen Leistungen der Kinder mit DZ liegen 2007 zu Beginn der 5. Klasse 59 Punkte tiefer als die Leistungen der Kinder mit DE. Dieser Rückstand der Kinder mit DZ ist mit einer Effektgrösse von  $d = 0.59$  als mittelstark zu beurteilen.

Abbildung 3.8: Leistungsrückstand der Kinder mit DZ in Mathematik nach Jahr



Anmerkung: Anteil Schülerinnen und Schüler mit DZ in der Stichprobe:  
 2004: 19,0% (n = 482)  
 2005: 20,1% (n = 1082)  
 2006: 21,0% (n = 1177)  
 2007: 22,1% (n = 1220)

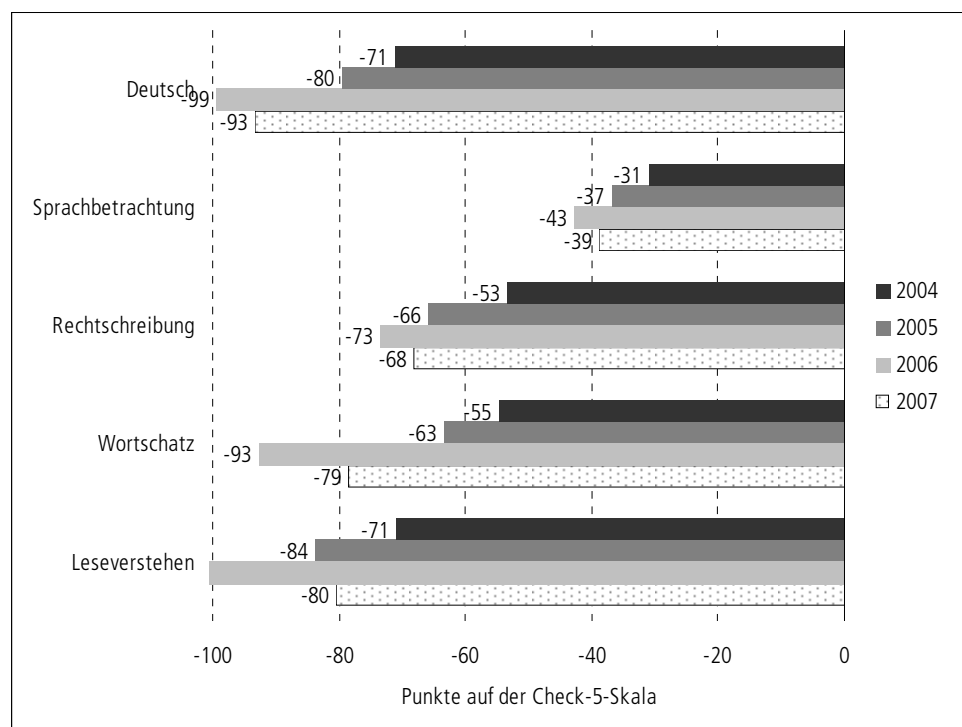
Im Jahr 2007 reduzierten sich die Leistungsunterschiede zwischen Schülerinnen und Schülern mit DZ und Schülerinnen und Schülern mit DE auch in den Teilbereichen der Mathematik. Die grössten Unterschiede zwischen Kindern mit DE und Kindern mit DZ bestehen weiterhin in den Teilbereichen «Mathematisches Problemlösen» und «Sachrechnen», der kleinste Unterschied besteht im Teilbereich «Arithmetik». Ob es sich bei dieser Reduktion der Leistungsunterschiede um einen längerfristigen Trend handelt, müssen die Check 5 der folgenden Jahre zeigen.

### 3.9 Leistungsunterschiede nach Erstsprache in Deutsch

Abbildung 3.9 zeigt die Leistungsunterschiede zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache (DE) und Schülerinnen und Schülern mit Deutsch als Zweitsprache (DZ) in Deutsch und in den Teilbereichen von Deutsch. Die Unterschiede sind dargestellt nach Testjahr auf der Check-5-Skala.

Im Jahr 2004 betrug der Rückstand der Kinder mit DZ gegenüber den Kindern mit DE in der Mathematik 71 Punkte (schwarzer Balken). Bis ins Jahr 2006 erhöhte sich der Rückstand auf 99 Punkte (hellgrauer Balken). Im Jahr 2007 hingegen reduzierte sich der Leistungsunterschied zwischen Schülerinnen und Schülern mit DE und Schülerinnen und Schülern mit DZ auf 93 Punkte. Dieser Rückstand der Kinder mit DZ ist mit einer Effektgrösse von  $d = 0.93$  aber weiterhin als stark zu beurteilen.

Abbildung 3.9: Leistungsrückstand der Kinder mit DZ in Deutsch nach Jahr



Anmerkung: Anteil Schülerinnen und Schüler mit DZ in der Stichprobe:  
 2004: 19,0% (n = 482)  
 2005: 20,1% (n = 1082)  
 2006: 21,0% (n = 1177)  
 2007: 22,1% (n = 1220)

Die grössten Unterschiede zwischen Kindern mit DE und Kindern mit DZ bestehen in den Teilbereichen «Leseverstehen» und «Wortschatz», in denen auch das semantische Sprachverständnis getestet wird (je rund 80 Punkte). Der kleinste Unterschied besteht im Teilbereich «Sprachbetrachtung», in dem vorwiegend von der Schule vermitteltes Wissen und Können geprüft werden (39 Punkte).

### 3.10 Effekte der individuellen Lernvoraussetzungen auf die Leistungen der Schülerinnen und Schüler

Mit einer Mehrebenenanalyse wurden die Effekte der individuellen Lernvoraussetzungen auf die Leistungen der Schülerinnen und Schüler untersucht. Zu den individuellen Lernvoraussetzungen werden das Geschlecht, die Erstsprache sowie die sozioökonomische Herkunft gezählt.

Tabelle 3.1 zeigt die Ergebnisse der Mehrebenenanalyse im Fach Mathematik. In der ersten Zeile steht die Konstante. Die Konstante ist die durchschnittliche Mathematikleistung. In der zweiten Zeile ist die Wirkung des Geschlechts, in der dritten Zeile die Wirkung der Erstsprache und in der vierten Zeile die Wirkung der sozioökonomischen Herkunft auf die Mathematikleistung (Koeffizient) angegeben, wenn sich die anderen Lernvoraussetzungen nicht unterscheiden.

Tabelle 3.1: Effekte der individuellen Lernvoraussetzungen auf die Mathematikleistungen

	2004	2005	2006	2007
	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>
<i>Konstante (Anzahl Punkte auf der Check-5-Skala)</i>	501.7	499.9	483.3	482.5
Knaben	20.2**	20.9**	24.4**	17.2**
Deutsch als Zweitsprache	-33.6**	-35.6**	-45.8**	-40.7**
Sozioökonomische Herkunft	10.1**	32.3**	32.0**	32.2**

Anmerkung: \*\* =  $p < 0.001$

Knaben erzielen 2007 eine um 17,2 Punkte höhere Mathematikleistung als Mädchen unabhängig von der Erstsprache und von der sozioökonomischen Herkunft.

Kinder mit Deutsch als Zweitsprache (DZ) erreichen 2007 in der Mathematik eine Leistung, die um 40,7 Punkte tiefer ist als die Leistung von Kindern mit Deutsch als Erstsprache (DE) unabhängig von Geschlecht und von der sozioökonomischen Herkunft.

Der Effekt der sozioökonomischen Herkunft beträgt im Jahr 2007 32,2 Punkte. Steigt die sozioökonomische Herkunft eines Kindes um einen Indexpunkt, so erhöht sich die Mathematikleistung des Kindes um 32,2 Punkte unabhängig von Geschlecht und Erstsprache.

Die Effekte der individuellen Lernvoraussetzungen auf die Leistungen in der Mathematik bleiben seit 2006 weitgehend konstant. Die grössten Wirkungen haben die Erstsprache und die sozioökonomische Herkunft der Kinder. Insgesamt können rund 12 Prozent der Leistungsunterschiede in der Mathematik durch die Lernvoraussetzungen der Schülerinnen und Schüler erklärt werden.



Tabelle 3.2 zeigt die Ergebnisse der Mehrebenenanalyse im Fach Deutsch. In der ersten Zeile steht die Konstante. Die Konstante ist die durchschnittliche Deutschleistung. In der zweiten Zeile ist die Wirkung des Geschlechts, in der dritten Zeile die Wirkung der Erstsprache und in der vierten Zeile die Wirkung der sozioökonomischen Herkunft auf die Deutschleistung (Koeffizient) angegeben, wenn sich die anderen Lernvoraussetzungen nicht unterscheiden.

Tabelle 3.2: Effekte der individuellen Lernvoraussetzungen auf die Deutschleistungen

	2004	2005	2006	2007
	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>
<i>Konstante (Anzahl Punkte auf der Check-5-Skala)</i>	501.0	493.6	500.2	507.8
Knaben	-26.1**	-21.7**	-19.1**	-15.6**
Deutsch als Zweitsprache	-51.1**	-57.3**	-73.8**	-70.3**
Sozioökonomische Herkunft	11.7**	34.5**	38.2**	42.0**

Anmerkung: \*\* =  $p < 0.001$

Knaben erzielen 2007 eine um 15,6 Punkte tiefere Deutschleistung als Mädchen unabhängig von der Erstsprache und von der sozioökonomischen Herkunft.

Kinder mit Deutsch als Zweitsprache (DZ) erreichen 2007 in Deutsch eine Leistung, die um 70,3 Punkte tiefer ist als die Leistung von Kindern mit Deutsch als Erstsprache (DE) unabhängig von Geschlecht und von der sozioökonomischen Herkunft der Kinder.

Der Effekt der sozioökonomischen Herkunft beträgt im Jahr 2007 42,0 Punkte. Steigt die sozioökonomische Herkunft eines Kindes um einen Indexpunkt, so erhöht sich die Deutschleistung des Kindes um 42 Punkte unabhängig von Geschlecht und Erstsprache.

Die Effekte der individuellen Lernvoraussetzungen auf die Leistungen in Deutsch verändern sich seit 2006 nur wenig. Die grössten Wirkungen haben die Erstsprache und die sozioökonomische Herkunft der Kinder. Vergleicht man die letzten vier Testjahre, so fällt einerseits auf, dass der Einfluss des Geschlechts auf die Deutschleistungen immer kleiner wird. Andererseits steigt der Einfluss der sozialen Herkunft stetig an und ist im Jahr 2007 so hoch ist wie noch nie seit 2004.

Die Lernvoraussetzungen der Schülerinnen und Schüler erklären rund 19 Prozent der Leistungsunterschiede in Deutsch. Der Einfluss der individuellen Lernvoraussetzungen auf die Leistungen ist damit in Deutsch insgesamt grösser als in der Mathematik.

### 3.11 Effekte der Klassenzusammensetzung auf die Leistungen der Schülerinnen und Schüler

Mit einer Mehrebenenanalyse wurde der Einfluss der Klassenzusammensetzung auf die Fachleistung einer Klasse untersucht bei gleichzeitiger Kontrolle der individuellen Lernvoraussetzungen Geschlecht, Erstsprache und soziale Herkunft. In die Analyse mit einbezogen wurde der Anteil Knaben in einer Klasse, der Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache (DZ), die soziale Zusammensetzung der Klasse (Mittelwert des sozioökonomischen Indexes einer Klasse) sowie die Klassengröße.

Tabelle 3.3 zeigt die Effekte der Klassenzusammensetzung auf die Mathematikleistungen einer Klasse. In der ersten Zeile steht die Konstante. Die Konstante ist die durchschnittliche Mathematikleistung. In der zweiten bis fünften Zeile sind die Wirkungen der untersuchten Merkmale auf die Mathematikleistung angegeben, wenn gleichzeitig die individuellen Lernvoraussetzungen der Kinder konstant gehalten werden (Koeffizient).

Tabelle 3.3: Effekte der Klassenzusammensetzung auf die Mathematikleistungen

	2004	2005	2006	2007
	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>
<i>Konstante (Anzahl Punkte auf der Check-5-Skala)</i>	501.8	502.0	485.5	482.5
Anteil Knaben	-0.2	0.1	0.1	0.3
Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache	0.3	0.1	0.0	0.1
Mittlere soziale Herkunft	4.3	4.7	16.5*	11.1
Klassengröße	-0.4	-0.3	0.0	-0.1

Anmerkung: \* =  $p < 0.005$

Aus Tabelle 3.3 lässt sich ablesen, dass bei Check 5 in den Jahren 2004, 2005 und 2006 weder der Anteil Knaben noch der Anteil Schülerinnen und Schüler mit DZ noch die Klassengröße einen statistisch signifikanten Effekt auf die individuellen Leistungen in einer Klasse haben. Auch die mittlere soziale Herkunft der Kinder einer Klasse hat 2007 keinen statistisch signifikanten Einfluss mehr auf die Mathematikleistungen der Schülerinnen und Schüler.

Tabelle 3.4 zeigt die Effekte der Klassenzusammensetzung auf die Deutschleistungen einer Klasse. In der ersten Zeile steht die Konstante. Die Konstante ist die durchschnittliche Deutschleistung. In der zweiten bis fünften Zeile sind die Wirkungen der untersuchten Merkmale auf die Deutschleistung angegeben, wenn gleichzeitig die individuellen Lernvoraussetzungen der Kinder konstant gehalten werden (Koeffizient).

Aus Tabelle 3.4 lässt sich ablesen, dass in den Jahren 2004 bis 2007 die soziale Klassenzusammensetzung, aber auch die Klassengrösse keine statistisch signifikanten Effekte auf die Leistungen der Schülerinnen und Schüler in einer Klasse haben.

Tabelle 3.4: Effekte der Klassenzusammensetzung auf die Deutschleistungen

	2004	2005	2006	2007
	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>	<i>Koeffizient</i>
<i>Konstante</i> (Anzahl Punkte auf der Check-5-Skala)	501.0	495.6	502.3	506.9
Anteil Knaben	0.0	-0.1	0.2	0.0
Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache	0.1	0.1	0.1	0.2
Mittlere soziale Herkunft	1.3	-3.2	9.6	-0.8
Klassengrösse	-0.1	-0.1	0.2	0.6

Anmerkung: \* =  $p < 0.005$

Diese Analysen zeigen, dass im Kanton Aargau die soziale Zusammensetzung einer Klasse für den Schulerfolg nicht relevant ist. Weder Kinder mit Deutsch als Erstsprache noch Kinder mit Deutsch als Zweitsprache haben negative Auswirkungen der Klassenzusammensetzung oder der Klassengrösse auf den Lernerfolg zu befürchten.

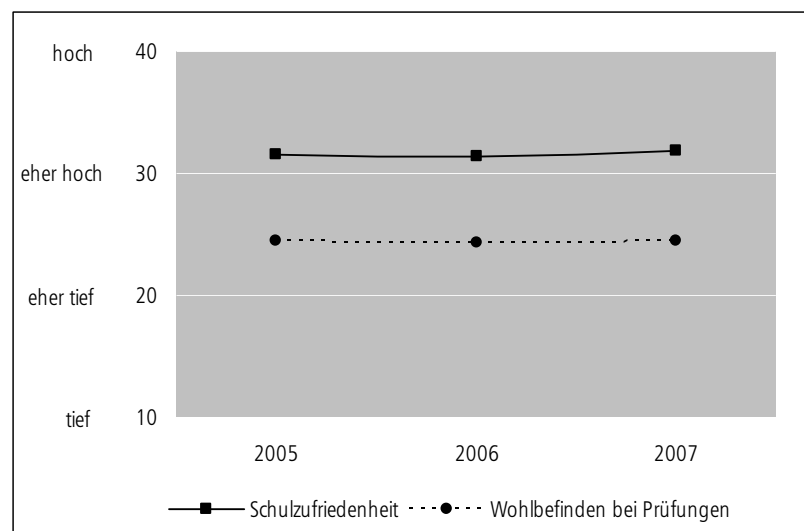
## 4 Selbstreguliertes Lernen

Als Indikatoren des selbstregulierten Lernens werden in Check 5 seit 2005 die Schulzufriedenheit, die Prüfungsangst sowie das fachspezifische Interesse und Selbstvertrauen erfasst. Den Schülerinnen und Schülern wurden zu jedem Bereich Aussagen vorgelegt wie «Ich gehe gerne zur Schule», «Ich habe bei Prüfungen immer Angst, dass ich schlechte Noten bekomme», «Mir gefällt Mathematik» oder «In Deutsch bin ich gut», denen sie mit Hilfe einer vierstufigen Skala entsprechend zustimmen konnten. Berechnet wurde für jeden Bereich die durchschnittliche Stärke der Zustimmung zu den Aussagen. Dazu wurden den vier Antwortkategorien die Zahlen 10 («stimmt überhaupt nicht»), 20 («stimmt eher nicht»), 30 («stimmt eher») und 40 («stimmt genau») zugeordnet.

### 4.1 Schulzufriedenheit und Prüfungsangst

Abbildung 4.1 zeigt die durchschnittliche «Schulzufriedenheit» und das durchschnittliche «Wohlbefinden bei Prüfungen», die positive Formulierung der Prüfungsangst. Die durchschnittliche Schulzufriedenheit ist 2007 mit 31,8 Punkten eher hoch. Das Wohlbefinden bei Prüfungen ist hingegen eher tief (24,5 Punkte). Sowohl die Schulzufriedenheit als auch das Wohlbefinden bei Prüfungen unterscheiden sich zwischen 2005 und 2007 nur zufällig.

Abbildung 4.1: Schulzufriedenheit und Prüfungsangst



Anmerkung: «Schulzufriedenheit»:  
2005: n = 4995; M = 31.5 Punkte; SE = 0.23 Punkte; SD = 6.6 Punkte  
2006: n = 5174; M = 31.4 Punkte; SE = 0.16 Punkte; SD = 6.5 Punkte  
2007: n = 5193; M = 31.8 Punkte; SE = 0.16 Punkte; SD = 6.7 Punkte

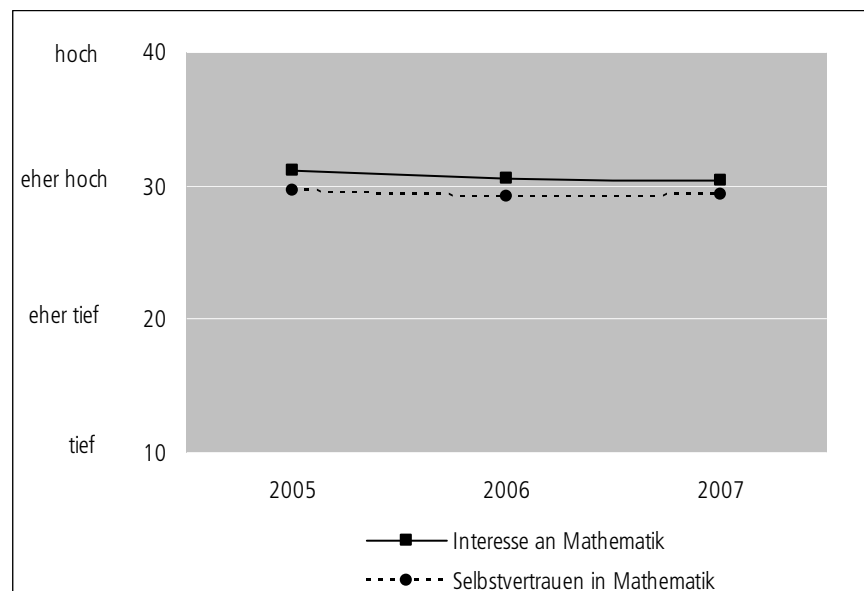
«Wohlbefinden bei Prüfungen»:  
2005: n = 4995; M = 24.5 Punkte; SE = 0.19 Punkte; SD = 7.8 Punkte  
2006: n = 5174; M = 24.4 Punkte; SE = 0.11 Punkte; SD = 7.7 Punkte  
2007: n = 5193; M = 24.5 Punkte; SE = 0.23 Punkte; SD = 7.7 Punkte

## 4.2 Interesse und Selbstvertrauen in der Mathematik

Abbildung 4.2 zeigt das durchschnittliche «Interesse an der Mathematik» und das durchschnittliche «Selbstvertrauen in der Mathematik» der Schülerinnen und Schüler. Das durchschnittliche Interesse an Mathematik ist als eher hoch zu beurteilen und beträgt 2007 30,4 Punkte. Seit 2005 ist das Interesse an Mathematik um rund 0,7 Punkte gesunken. Dieser Rückgang ist zwar statistisch signifikant, aber mit einer Effektgrösse von  $d = 0,08$  sehr schwach.

Das Selbstvertrauen in der Mathematik wird von den Schülerinnen und Schülern als eher hoch eingestuft. 2007 beträgt das durchschnittliche Selbstvertrauen in der Mathematik 29,3 Punkte. Die Einschätzung des Selbstvertrauens in der Mathematik unterscheidet sich zwischen 2005 und 2007 nur zufällig.

Abbildung 4.2: Interesse und Selbstvertrauen in der Mathematik



Anmerkung: «Interesse an Mathematik»:

2005:  $n = 4995$ ;  $M = 31.2$  Punkte;  $SE = 0.13$  Punkte;  $SD = 8.5$  Punkte  
2006:  $n = 5174$ ;  $M = 30.5$  Punkte;  $SE = 0.14$  Punkte;  $SD = 8.6$  Punkte  
2007:  $n = 5193$ ;  $M = 30.4$  Punkte;  $SE = 0.19$  Punkte;  $SD = 8.7$  Punkte

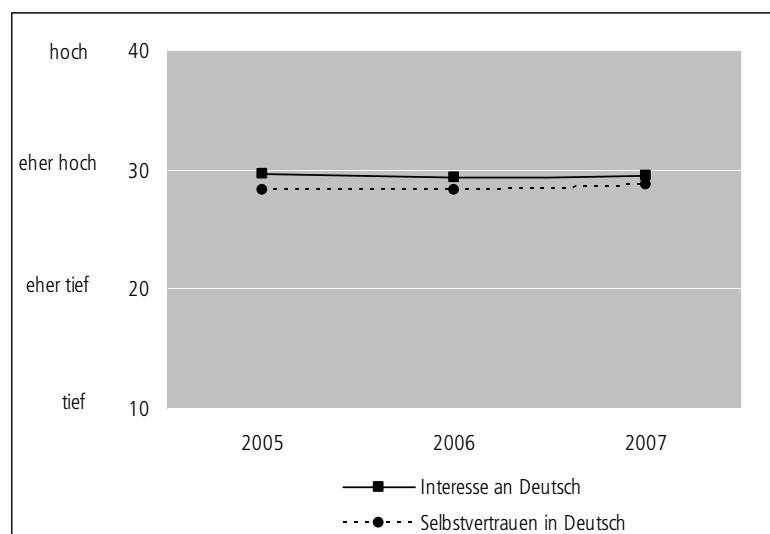
«Selbstvertrauen in Mathematik»:

2005:  $n = 4995$ ;  $M = 29.6$  Punkte;  $SE = 0.08$  Punkte;  $SD = 7.0$  Punkte  
2006:  $n = 5174$ ;  $M = 29.2$  Punkte;  $SE = 0.10$  Punkte;  $SD = 7.0$  Punkte  
2007:  $n = 5193$ ;  $M = 29.3$  Punkte;  $SE = 0.09$  Punkte;  $SD = 7.1$  Punkte

### 4.3 Interesse und Selbstvertrauen in Deutsch

Abbildung 4.3 zeigt das durchschnittliche «Interesse an Deutsch» und das durchschnittliche «Selbstvertrauen in Deutsch» der Schülerinnen und Schüler. Das durchschnittliche Interesse an Deutsch ist als eher hoch zu beurteilen und beträgt 2007 29,6 Punkte. Das Selbstvertrauen in Deutsch wird von den Schülerinnen und Schülern ebenfalls als eher hoch eingestuft (28,7 Punkte).

Abbildung 4.3: Interesse und Selbstvertrauen in Deutsch



Anmerkung: «Interesse an Deutsch»:

2005:  $n = 4995$ ;  $M = 29.6$  Punkte;  $SE = 0.14$  Punkte;  $SD = 8.4$  Punkte

2006:  $n = 5174$ ;  $M = 29.4$  Punkte;  $SE = 0.24$  Punkte;  $SD = 8.4$  Punkte

2007:  $n = 5193$ ;  $M = 29.6$  Punkte;  $SE = 0.21$  Punkte;  $SD = 8.4$  Punkte

«Selbstvertrauen in Deutsch»:

2005:  $n = 4995$ ;  $M = 28.3$  Punkte;  $SE = 0.07$  Punkte;  $SD = 7.1$  Punkte

2006:  $n = 5174$ ;  $M = 28.3$  Punkte;  $SE = 0.10$  Punkte;  $SD = 7.1$  Punkte

2007:  $n = 5193$ ;  $M = 28.7$  Punkte;  $SE = 0.10$  Punkte;  $SD = 7.3$  Punkte

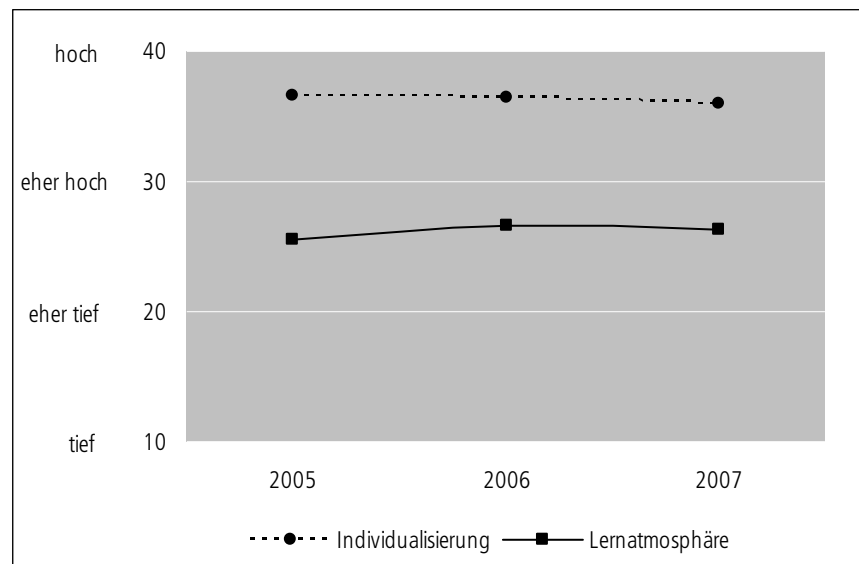
Insgesamt bleiben Interesse und Selbstvertrauen in Deutsch aber auch in der Mathematik in den Jahren 2005 bis 2007 nahezu konstant. In beiden Bereichen liegen die Einschätzungen im Fach Deutsch aber rund 1 Punkt tiefer als im Fach Mathematik. Das heisst, die Schülerinnen und Schüler beurteilen ihr Interesse und ihr Selbstvertrauen in der Mathematik höher als im Deutsch. Zudem ist die Streuung der Einschätzungen vergleichsweise hoch. Die Standardabweichungen der Einschätzungen der Schülerinnen und Schüler betragen im Interesse rund 8,5 Punkte und im Selbstvertrauen rund 7 Punkte. Das heisst, dass sich Interesse und Selbstvertrauen zwischen den Schülerinnen und Schülern erheblich unterscheiden.

## 5 Unterrichtswahrnehmung

Die Unterrichtswahrnehmung der Schülerinnen und Schüler wird in den beiden Bereichen «Individualisierung des Unterrichts» und «Lernatmosphäre» erhoben. Den Schülerinnen und Schülern wurden Aussagen vorgelegt wie «Unsere Lehrperson bemüht sich, dass alle mitkommen» oder «Die Lehrperson bemüht sich sehr, uns alles gut zu erklären» beziehungsweise «Im Unterricht herrscht oft Lärm» oder «Häufig ist die Klasse so unruhig, dass man sich nicht konzentrieren kann». Berechnet wurde für jeden Bereich die durchschnittliche Stärke der Zustimmung zu den Aussagen. Dazu wurden den vier Antwortkategorien die Zahlen 10 («stimmt überhaupt nicht»), 20 («stimmt eher nicht»), 30 («stimmt eher») und 40 («stimmt genau») zugeordnet.

Abbildung 5.1 zeigt die durchschnittliche Wahrnehmung der Unterrichtsmerkmale «Individualisierung des Unterrichts» und «Lernatmosphäre». Die Individualisierung des Unterrichts wird 2007 von den Schülerinnen und Schülern mit 36 Punkten als sehr hoch eingeschätzt. Die Lernatmosphäre wird mit 26,3 Punkten hingegen deutlich weniger positiv wahrgenommen. Die vergleichsweise hohe Standardabweichung von 7,1 Punkten deutet zudem darauf hin, dass die Lernatmosphäre von den Schülerinnen und Schülern sehr unterschiedlich wahrgenommen wird.

Abbildung 5.1: Individualisierung des Unterrichts und Lernatmosphäre



Anmerkung: «Individualisierung des Unterrichts»:  
2005: n = 4995; M = 36.6 Punkte; SE = 0.08 Punkte; SD = 4.4 Punkte  
2006: n = 5174; M = 36.5 Punkte; SE = 0.07 Punkte; SD = 4.2 Punkte  
2007: n = 5193; M = 36.0 Punkte; SE = 0.08 Punkte; SD = 4.7 Punkte

«Lernatmosphäre»:  
2005: n = 4995; M = 25.6 Punkte; SE = 0.12 Punkte; SD = 7.3 Punkte  
2006: n = 5174; M = 26.6 Punkte; SE = 0.14 Punkte; SD = 6.2 Punkte  
2007: n = 5193; M = 26.3 Punkte; SE = 0.21 Punkte; SD = 7.1 Punkte

## 6 Beurteilung von Check 5 durch die Lehrpersonen

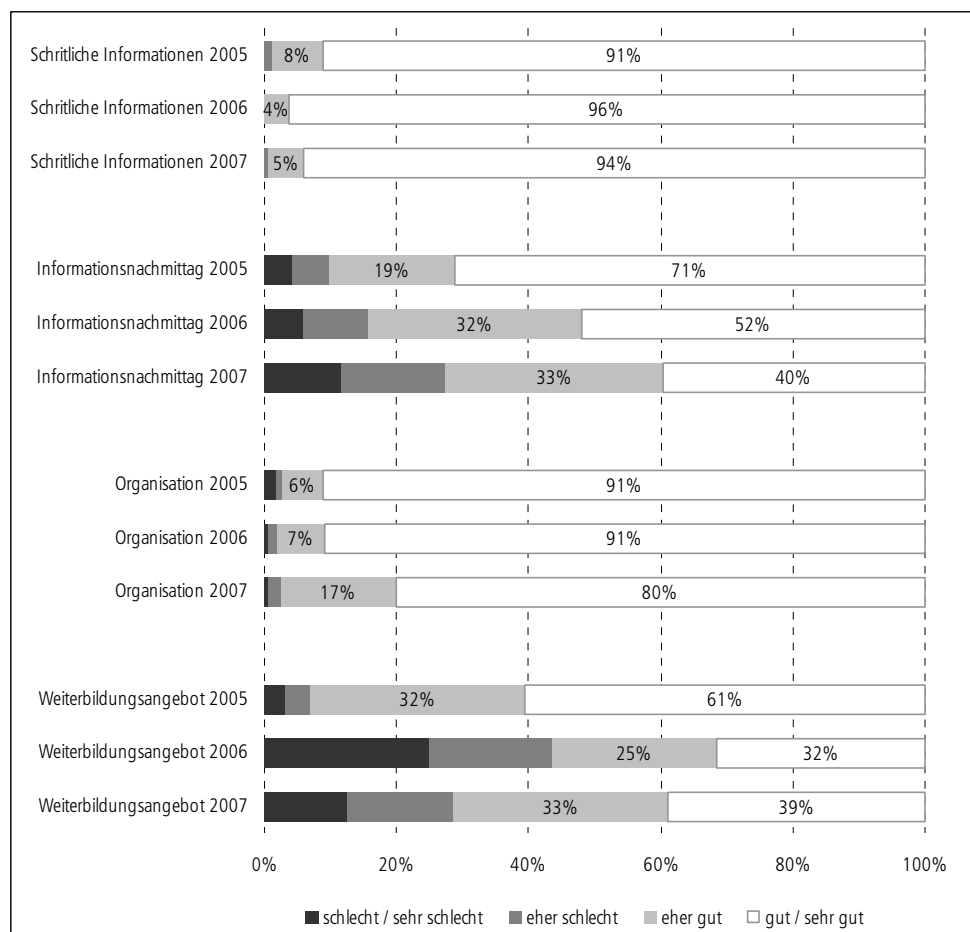
Um Check 5 laufend zu evaluieren, wird seit 2005 ein standardisierter Fragebogen eingesetzt. In diesem Fragebogen können Lehrpersonen anhand von verschiedenen Aussagen Nutzen und Qualität von Check 5 beurteilen. Zudem können sie Bemerkungen und Verbesserungsvorschläge zu Check 5 im Fragebogen schriftlich anbringen. Im vorliegenden Bericht sind die Anmerkungen von Lehrpersonen kursiv dargestellt.

2005 nahmen 252 Lehrpersonen (82 Prozent) und im Jahr 2006 267 Lehrerinnen und Lehrer (83 Prozent) an der Befragung teil. 2007 ist der Rücklauf mit 242 Lehrpersonen (75 Prozent) etwas geringer.

### 6.1 Beurteilung der Durchführung von Check 5

Abbildung 6.1 zeigt die Einschätzungen der Lehrpersonen zur Durchführung von Check 5. Weiss und hellgrau ist der Anteil an positiven Aussagen («eher gut» bzw. «gut» und «sehr gut») einfärbt, dunkelgrau und schwarz der Anteil an negativen Einschätzungen («eher schlecht» bzw. «schlecht» und «sehr schlecht»).

Abbildung 6.1: Beurteilung der Durchführung von Check 5 durch die Lehrpersonen





Wie Abbildung 6.1 zeigt, beurteilen die Lehrerinnen und Lehrer die Durchführung des Check 5 weitgehend als sehr positiv. Die schriftlichen Informationen, die im Rahmen von Check 5 abgegeben werden, werden 2007 von 99 Prozent der Lehrpersonen als «sehr gut», «gut» oder «eher gut» beurteilt. Dasselbe gilt auch für die Organisation von Check 5. Seit 2005 wird die Organisation von nahezu 100 Prozent der Lehrerschaft als «sehr gut», «gut» oder «eher gut» beurteilt.

Der Informationsnachmittag hingegen stösst zunehmend auf Ablehnung. 2005 beurteilten 71 Prozent der Lehrerinnen und Lehrer den obligatorischen Informationsnachmittag als «gut» oder «sehr gut». 2007 sind es noch 40 Prozent der Lehrerinnen und Lehrer, während rund ein Viertel der Lehrpersonen den Informationsnachmittag als «sehr schlecht», «schlecht» oder «eher schlecht» beurteilt. Von vielen Lehrpersonen, die bereits mehrmals am Check 5 teilgenommen haben, wird die Informationsveranstaltung als lästige und überflüssige Pflichterfüllung erlebt. Sie schlagen daher vor, die Teilnahme am Informationsnachmittag zumindest für *«alte Hasen»* als freiwillig zu erklären.

Auch das Weiterbildungsangebot wird immer noch relativ stark kritisiert, auch wenn sich der Anteil der negativen Beurteilungen, vermutlich aufgrund der stetigen Entwicklung und Verbesserung des obligatorischen Weiterbildungsangebots, seit 2006 von 44 Prozent auf 29 Prozent reduziert hat. Kritisiert wird auch beim Weiterbildungsangebot in erster Linie die alljährliche Teilnahmepflicht. Viele Lehrpersonen, die bereits mehrmals am Check 5 teilgenommen haben, empfinden die individuelle Weiterbildung als unnötig und langweilig. Inhaltlich werden vor allem konkrete Ideen und Vorschläge zur Verbesserung des Unterrichts oder zur Förderung einzelner Schülerinnen und Schüler vermisst. Auf der anderen Seite beurteilt jedoch rund ein Drittel der Lehrerinnen und Lehrer die Weiterbildungen als «gut» oder «sehr gut». Ein Lehrperson beispielsweise schreibt: *«Wir haben den Check 5 im Team reflektiert und super gute Erfahrungen gemacht damit!»*

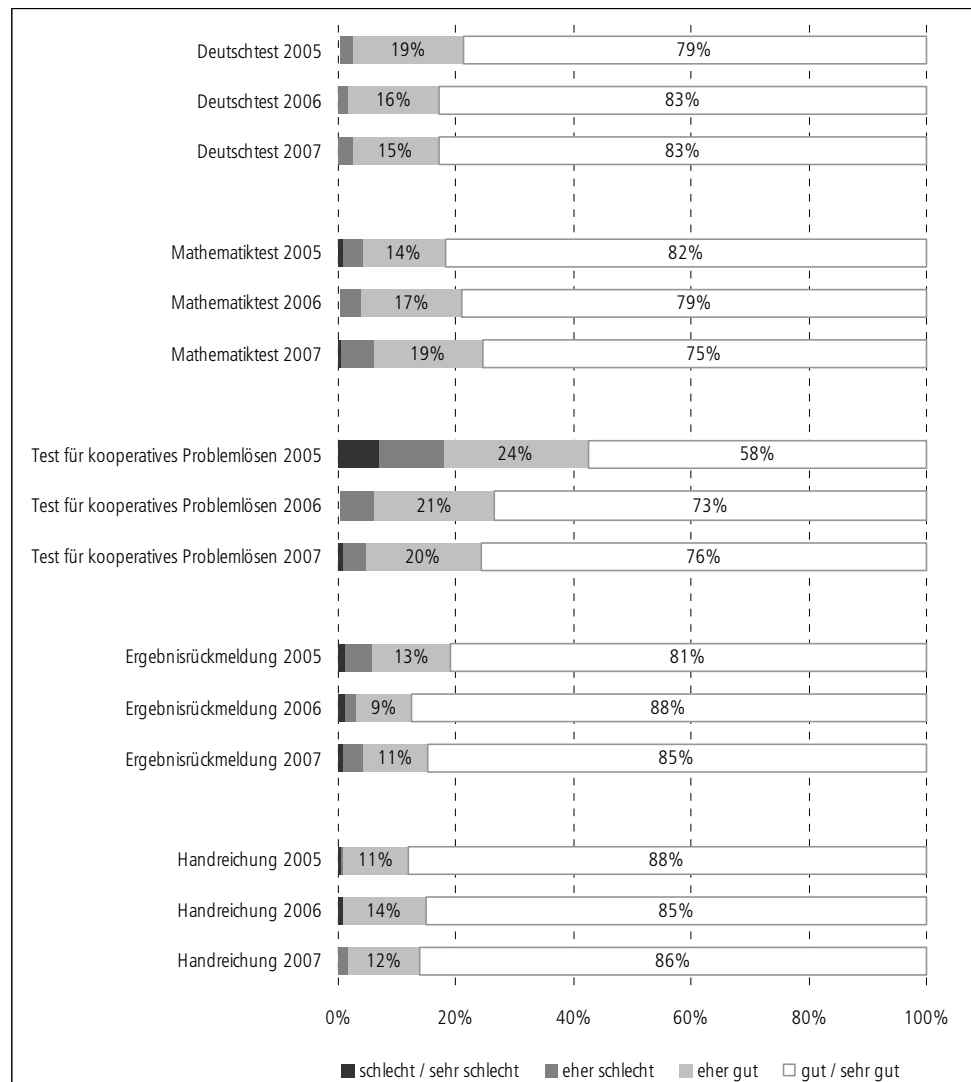
## 6.2 Beurteilung der Testinstrumente von Check 5

Abbildung 6.2 zeigt die Beurteilungen der Testinstrumente von Check 5 durch die Lehrpersonen. Weiss und hellgrau ist der Anteil an positiven Aussagen («eher gut» bzw. «gut» und «sehr gut») eingefärbt, dunkelgrau und schwarz der Anteil an negativen Einschätzungen («eher schlecht» bzw. «schlecht» und «sehr schlecht»).

Allgemein werden die Testinstrumente, die beim Check 5 eingesetzt werden, von den Lehrerinnen und Lehrern sehr positiv beurteilt. Annähernd 100 Prozent der Lehrerschaft beurteilen die Tests als «sehr gut», «gut» oder «eher gut». Davon zeugen auch schriftliche Kommentare von Lehrpersonen wie *«Ich finde Check 5 eine sehr gute Sache!»* oder *«Allgemein guter Test!»*. Dies gilt insbesondere für den Deutsch- und den Mathematiktest. 2007 beurteilen 83 Prozent der Lehrpersonen den Deutschtest und 75 Prozent den Mathematiktest als «sehr gut» oder «gut». Kritisiert wird, dass der diesjährige Deutschtest zu stark den Bereich «Textverständnis» in den Mittelpunkt stelle und dass der Mathematiktest zu viele Geometrieaufgaben beinhalte, die zudem gemäss Aargauer Mathematiklehrmittel freiwillig seien.

Der Test für das kooperative Problemlösen wird seit seiner Neukonzeption 2006 von rund drei Vierteln der Lehrpersonen als «sehr gut» oder «gut» beurteilt.

Abbildung 6.2: Beurteilung der Testinstrumente von Check 5 durch die Lehrpersonen

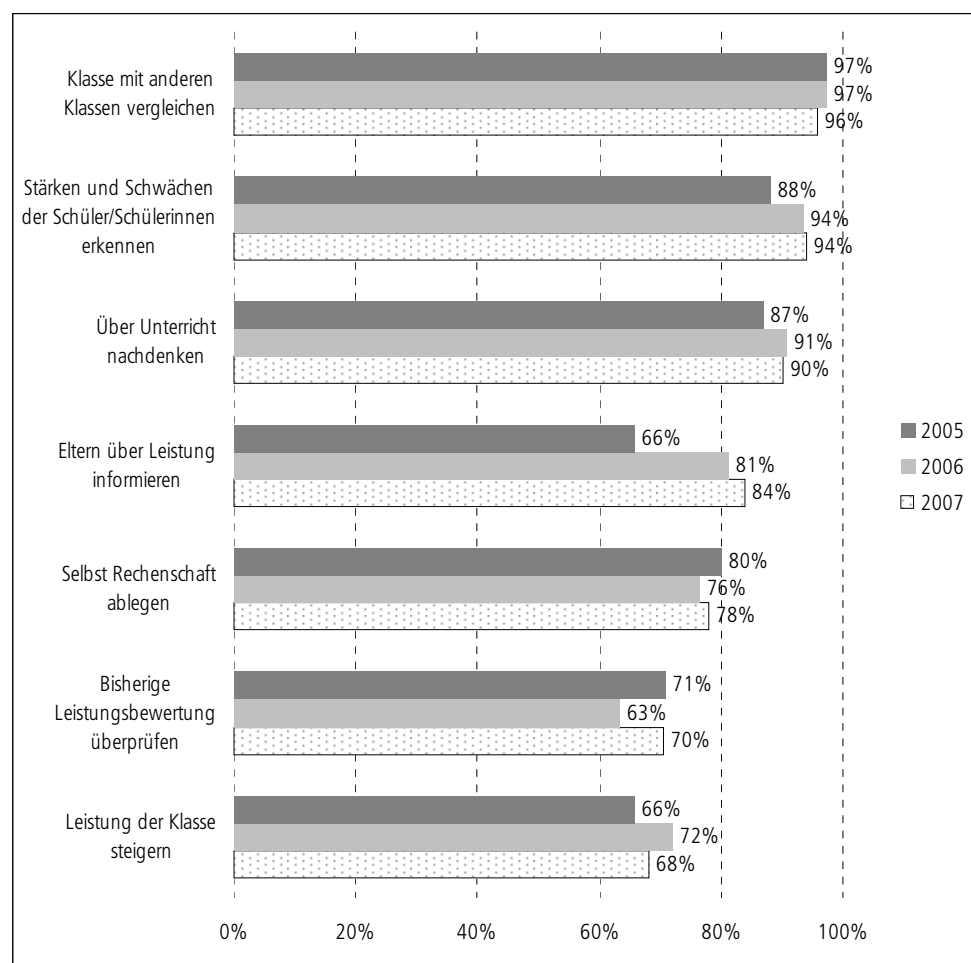


Konstant sehr positiv werden zudem auch die Ergebnisrückmeldung und die Handreichung beurteilt. Das zeigen beispielsweise auch folgende Aussagen von Lehrpersonen zu den Ergebnisrückmeldungen 2007: «Die Ergebnisrückmeldung ist sehr aussagekräftig und sehr wertvoll» oder «Mir hilft die gewählte Art gut weiter!». Einige Lehrpersonen vermissen die fehlende oder unzureichende Möglichkeit zur Individualdiagnose. Sie hätten daher gerne Einsicht in die Tests ihrer Schülerinnen und Schüler. Gewünscht wird von einigen Lehrpersonen zudem eine digitale Version der Ergebnisrückmeldung sowie ein leeres Testheft als Belegexemplar, um die gestellten Aufgaben anzusehen und mit den Schülerinnen und Schülern oder auch mit den Eltern zu besprechen. Andere Lehrerinnen und Lehrer sehen aber gerade darin eine Gefahr für die Testobjektivität und vermuten, dass sich gewisse Lehrpersonen nicht an die Spielregeln halten und mit ihren Klassen auf die Prüfungen üben.

### 6.3 Nutzung der Ergebnissrückmeldung

Abbildung 6.3 zeigt, wozu die Ergebnissrückmeldung von den Lehrpersonen genutzt wird. Am meisten wird die Ergebnissrückmeldung zum Leistungsvergleich mit anderen Klassen genutzt (2007: 96 Prozent). Über 70 Prozent der Lehrerinnen und Lehrer nutzten die Ergebnissrückmeldung aber auch, um «über den eigenen Unterricht nachdenken» (2007: 90 Prozent), um «sich selbst Rechenschaft ablegen» (2007: 78 Prozent) oder um die «eigene Leistungsbewertung zu überprüfen» (2007: 70 Prozent), alles explizite Ziele von Check 5.

Abbildung 6.3: Nutzung der Ergebnissrückmeldung



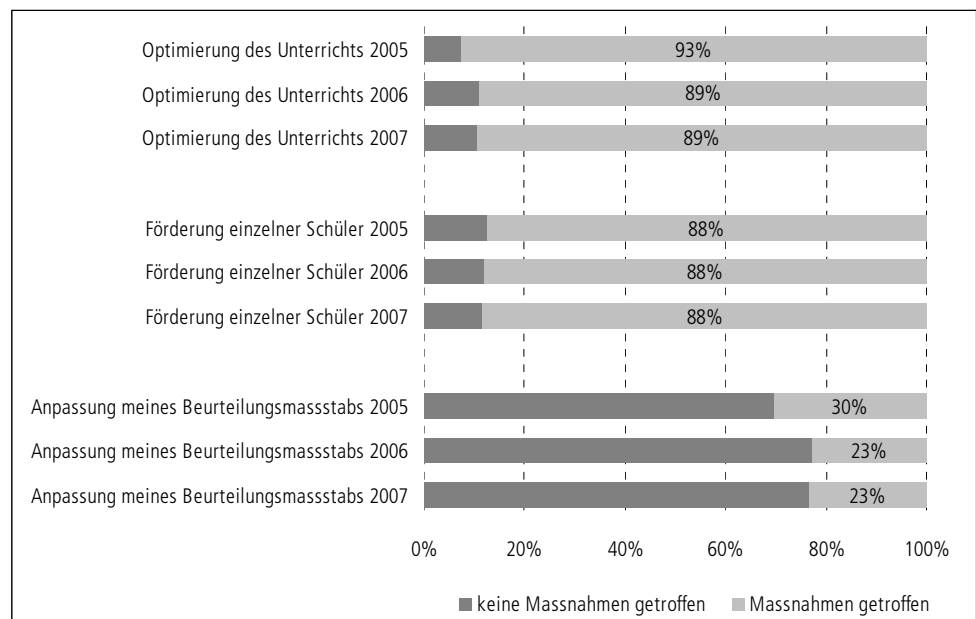
Bei der Nutzung der Ergebnissrückmeldung zeigen sich bei den meisten Nutzungsmöglichkeiten nur geringe Unterschiede zwischen den Durchführungen 2005, 2006 und 2007. Auffällig ist allenfalls, wie viele Lehrpersonen die Ergebnissrückmeldung nutzen, um individuelle Stärken und Schwächen ihrer Schülerinnen und Schüler zu erkennen (2007: 94 Prozent), sowie die steigende Bedeutung, die der Ergebnissrückmeldung bei der Information der Eltern über die Leistung ihrer Kinder zukommt (2007: 84 Prozent). Einige Lehrpersonen wünschen sich daher Ergebnissrückmeldungen, die für Eltern aussagekräftiger und einfacher verständlich sind.

#### 6.4 Art der getroffenen leistungsfördernden Massnahmen

Aufgrund der Ergebnismeldung haben 2007 93 Prozent der Lehrerinnen und Lehrer leistungssteigernde oder -fördernde Massnahmen ergriffen.

Abbildung 6.4 zeigt, welche Massnahmen die Lehrpersonen getroffen haben.

Abbildung 6.4: Getroffene Massnahmen



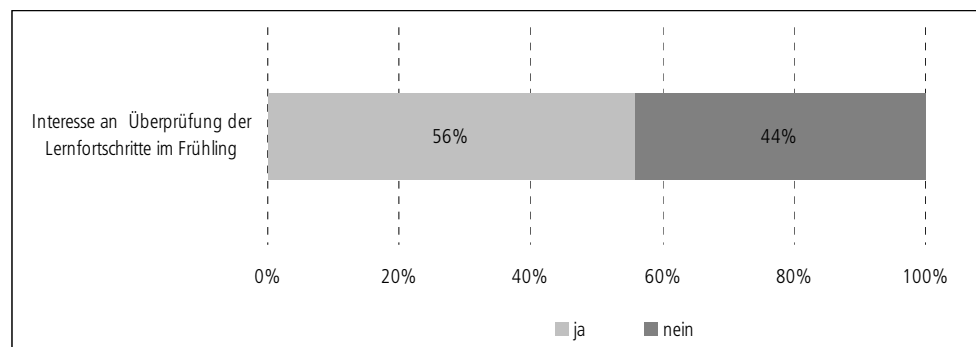
Im Jahr 2007 ergriffen 89 Prozent der Lehrpersonen Massnahmen zur Optimierung ihres Unterrichts, 88 Prozent Massnahmen zur Förderung einzelner Schülerinnen und Schüler und 23 Prozent ergriffen Massnahmen zur Anpassung ihres Beurteilungsmassstabs. Die Art der von den Lehrpersonen getroffenen Massnahmen unterscheidet sich seit 2005 kaum.

21 Prozent der Lehrpersonen geben an, die getroffenen Massnahmen im Frühjahr 2007 bereits voll und ganz umgesetzt zu haben, 79 Prozent haben die Massnahmen bereits teilweise umgesetzt.

### 6.5 *Interesse an einer zweiten Durchführung von Check 5*

Abbildung 6.5 zeigt das Interesse der Lehrpersonen an einer zweiten, verkürzten Durchführung des Check 5 gegen Ende der 5. Klasse. Anhand dieser zweiten Durchführung wäre es möglich, die Effekte der getroffenen Massnahmen auf die Leistung von einzelnen Schülerinnen und Schülern oder auch der ganzen Klasse zu überprüfen.

Abbildung 6.5: Interesse an einer Überprüfung der Lernfortschritte im Frühling



Etwas mehr als die Hälfte der Lehrerinnen und Lehrer (56 Prozent) würde eine Überprüfung der Lernfortschritte im Frühling begrüßen. 44 Prozent der Lehrpersonen stehen dieser Idee hingegen ablehnend gegenüber.

## 7 Fazit

Im Schuljahr 2007/08 haben 321 Klassen mit 5530 Schülerinnen und Schülern an Check 5 teilgenommen. Das sind rund 90 Prozent aller 5. Klassen im Kanton Aargau und damit so viele wie noch nie.

Die Leistungen in der Mathematik und in Deutsch haben sich im Kanton Aargau in den letzten vier Jahren nur sehr wenig verändert. Trotzdem unterscheidet sich die Leistungsentwicklung in der Mathematik und in Deutsch. In Deutsch werden die Leistungen seit 2005 tendenziell besser, in der Mathematik ist seit 2005 ein langsamer Abwärtstrend feststellbar. Noch sind die Unterschiede zwischen den einzelnen Testjahren aber schwach und sollten nicht überinterpretiert werden.

Die Leistungsunterschiede zwischen Mädchen und Knaben haben sich im Vergleich zum Vorjahr sowohl in der Mathematik als auch in Deutsch reduziert. Kleine Unterschiede bleiben aber bestehen. Die Knaben haben weiterhin einen kleinen Leistungsvorsprung in der Mathematik, besonders ausgeprägt im Teilbereich «Mathematisches Problemlösen». Die Mädchen haben einen kleinen Leistungsvorsprung in Deutsch, insbesondere im Teilbereich «Rechtschreibung». In Deutsch ist der Vorsprung der Mädchen jedoch nur noch schwach und so gering wie noch nie seit 2004.

Auch die Leistungsunterschiede zwischen den Schülerinnen und Schülern mit Deutsch als Erstsprache (DE) und den Schülerinnen und Schülern mit Deutsch als Zweitsprache (DZ) haben sich im Vergleich zum Jahr 2006 geringfügig verringert. Inwieweit diese Reduktion der Leistungsunterschiede einer Trendumkehr gleichkommt, lässt sich aber erst anhand der Testresultate der nächsten Jahre abschliessend beantworten.

Sehr grosse Leistungsunterschiede zwischen Kindern mit DE und Kindern mit DZ bestehen weiterhin in den Teilbereichen «Mathematisches Problemlösen» sowie «Leseverstehen» und «Wortschatz», in denen semantisches Sprachverständnis vorausgesetzt und getestet wird. Die geringsten Leistungsunterschiede bestehen in den Teilbereichen «Arithmetik» und «Sprachbetrachtung», in denen weitgehend formalisiert, von der Schule vermittelte Fertigkeiten und Routinen geprüft werden.

Insgesamt bleibt der Einfluss der Erstsprache und der sozioökonomischen Herkunft auf die Leistungen der Schülerinnen und Schüler vor allem in Deutsch gross. Das heisst, ein grosser Teil der individuellen Leistungsunterschiede kann, unabhängig von der sozialen Zusammensetzung der Klasse und vom Unterrichtsstil der Lehrperson, mit den unterschiedlichen Lernvoraussetzungen der Schülerinnen und Schüler erklärt werden.

Wie der Zwischenbericht zum Check5 2007 gezeigt hat, sind die Unterschiede aber nicht nur zwischen den Schülerinnen und Schülern, sondern auch zwischen den einzelnen Klassen beträchtlich. Grosse Leistungsunterschiede zwischen den Klassen bestehen in den Teilbereichen «Arithmetik», «Geometrie», «Sprachbetrachtung» und «Texte schreiben», in denen hauptsächlich von der Schule vermitteltes Wissen geprüft wird. In diesen Teilbereichen können letztlich rund 20 Prozent der Leistungsunterschiede auf Merkmale des Unterrichts zurückgeführt werden.

Die Schulzufriedenheit wird von den Schülerinnen und Schülern als «eher hoch» beurteilt. Auch das Interesse sowie das Selbstvertrauen in der Mathematik und in Deutsch werden als «eher hoch» eingeschätzt. Vergleicht man die diesjährigen Resultate mit den Befragungen 2005 und 2006, so ist ein leichter Rückgang des Interesses in der Mathematik festzustellen, während das Selbstvertrauen in Deutsch geringfügig gestiegen ist. Noch immer werden aber Interesse und Selbstvertrauen in der Mathematik höher eingeschätzt als in Deutsch.

Die Lehrpersonen beurteilen Check 5 auch im Schuljahr 2007/08 sehr positiv. Sowohl die Organisation von Check 5 als auch die verschiedenen Testinstrumente werden als gut beurteilt. Auch die Weiterbildung wird etwas positiver beurteilt als noch im Jahr 2006. Es ist zu erwarten, dass sich die Einstellungen der Lehrerinnen und Lehrer gegenüber dem Weiterbildungsangebot in Zukunft dank der neu geschaffenen Möglichkeit, die individuelle Weiterbildung durch eine gemeinsame Interpretation und Diskussion der Testergebnisse im Schulteam zu ersetzen, weiter verbessern wird.

Zunehmend kritisiert wird hingegen der Informationsnachmittag. Für viele Lehrpersonen, die bereits mehrmals an Check 5 teilgenommen haben, ist der Informationsnachmittag nicht mehr so informativ, wie der Name vermuten lässt, sondern vielmehr eine langweilige und lästige Pflichtaufgabe.

Die Ergebnismrückmeldung mit den Testergebnissen der Schülerinnen und Schüler sowie den durchschnittlichen Leistungen der Klassen werden von den Lehrpersonen als sehr gut und hilfreich beurteilt. Am häufigsten wird die Ergebnismrückmeldung von den Lehrpersonen genutzt, um sich mit den anderen Klassen im Kanton zu vergleichen und um Stärken und Schwächen ihrer Schülerinnen und Schüler zu erkennen. Die Ergebnismrückmeldung wird aber von einem grossen Teil der Lehrerinnen und Lehrer auch zum Anlass genommen, um über den eigenen Unterricht nachzudenken oder sich selbst Rechenschaft abzulegen. Als Ergebnis dieser Reflexion müssen die Lehrpersonen im Rahmen von Check 5 qualitäts- und leistungssteigernde Massnahmen schriftlich formulieren und umsetzen. Dabei werden hauptsächlich Massnahmen zur Optimierung des Unterrichts und zur Förderung einzelner Schülerinnen und Schüler getroffen. Umgesetzt werden die Massnahmen bis zum Zeitpunkt der Befragung im Frühling vom grössten Teil der Lehrpersonen jedoch nur teilweise. Vermutlich stösst deshalb die Idee einer zweiten, verkürzten Durchführung des Check 5 zur Überprüfung der Lernfortschritte der Schülerinnen und Schüler nur auf ein verhaltenes Interesse. Etwa die Hälfte der Lehrerinnen und Lehrer würden einen zweiten Check 5 gegen Ende des 5. Schuljahres begrüssen.

## **Anhang: Glossar der statistischen Begriffe**

*Effektgrösse* – Zur Interpretation von statistisch signifikanten Unterschieden wird üblicherweise die Effektgrösse «d» berechnet, indem die Differenz der Mittelwerte durch die Standardabweichungen dividiert wird. Unterschiede, die aufgrund von verschiedenen Skalen zustande gekommen sind, werden so standardisiert und vergleichbar. Eine Effektgrösse von  $d = 0.2$  weist auf einen schwachen Effekt hin, eine Effektgrösse von  $d = 0.5$  auf einen mittleren Effekt und eine Effektgrösse von  $d = 0.8$  auf einen starken Effekt. Auf der standardisierten Check-5-Skala können dementsprechend Differenzen von 20 Punkten als klein, von 50 Punkten als mittelstark (oder deutlich) und von 80 Punkten als sehr gross bezeichnet werden.

*Kompetenzniveaus* – In Mathematik und Deutsch wurden Testaufgaben mit vergleichbarer inhaltlicher Ausrichtung und ähnlichem Schwierigkeitsgrad zu Kompetenzniveaus zusammengefasst. Schülerinnen und Schüler, die ein bestimmtes Kompetenzniveau erreichen, können die Aufgaben dieses Niveaus mit einer hinreichenden Wahrscheinlichkeit von mindestens 50 Prozent lösen. Dies erleichtert die inhaltliche Beschreibung der Fähigkeiten der Schülerinnen und Schüler.

*Link-Items* – Link-Items sind ausgewählte Testaufgaben, die in Check 5 während mehreren Jahren identisch eingesetzt wurden. Neue Testaufgaben sowie die Testergebnisse der Schülerinnen und Schüler der Jahre 2005 und 2006 wurden anhand der Schwierigkeitsparameter der Link-Items auf der ursprünglichen Skala 2004 positioniert. Damit ist es möglich, die Fähigkeiten der Schülerinnen und Schüler trotz teilweise unterschiedlicher Testinstrumente zu vergleichen. Durch die angepassten Modellrechnungen können allerdings einzelne Werte teilweise leicht von bereits publizierten Werten abweichen.

*Mittelwerte* – Der Mittelwert entspricht jeweils dem arithmetischen Mittel aller Einzelwerte. Der Mittelwert wird im Bericht mit «M» abgekürzt.

*Signifikantes Ergebnis* – Ein Ergebnis (Unterschied oder Zusammenhang) ist statistisch signifikant, wenn es durch ein statistisches Testverfahren überprüft und für gültig befunden wurde. Es kann mit einer bekannten, im Voraus festgelegten Irrtumswahrscheinlichkeit  $\alpha$  (üblicherweise  $\alpha = 0.05$ ) von der Stichprobe auf die Population geschlossen werden. In diesem Bericht erfüllen signifikante Ergebnisse die Bedingung  $p < 0.05$ . Die Signifikanz eines Ergebnisses sagt aber nichts aus über dessen Relevanz. Um signifikante Ergebnisse zu beurteilen, wird daher die Effektgrösse berechnet.

*Sozioökonomische Herkunft* – Für die sozioökonomische Herkunft wurde ein Index aus den Merkmalen «Ausbildung der Eltern», «Anzahl Bücher zu Hause» und «Anzahl Zimmer pro Person zu Hause» gebildet. Der Index zur sozioökonomischen Herkunft der Kinder hat den Mittelwert 0 und die Standardabweichung 1.



*Standardabweichung* – Die Standardabweichung ist ein quantitatives Mass für die Streuung der Einzelwerte um den Mittelwert. Entspricht die Verteilung der Einzelwerte einer Normalverteilung, dann besitzt die Standardabweichung die Eigenschaft, dass rund zwei Drittel (68 Prozent) der Einzelwerte zwischen dem Mittelwert  $\pm$  eine Standardabweichung  $[M \pm SD]$  liegen. Wird der Bereich um je eine Standardabweichung erweitert – Mittelwert  $\pm$  zwei Standardabweichungen  $[M \pm 2 SD]$  –, dann befinden sich darin rund 95 Prozent der Einzelwerte. Für die Leistungsdaten auf der standardisierten Check-5-Skala ( $SD = 100$  Punkte) heisst das, dass die Ergebnisse von rund zwei Dritteln der Schülerinnen und Schüler zwischen 400 und 600 Punkten liegen und rund 95 Prozent der Ergebnisse zwischen 300 und 700 Punkten liegen. Die Standardabweichung wird im Bericht mit «SD» abgekürzt.

*Stichprobe* – Die Stichprobe ist die Anzahl Schülerinnen und Schüler, die am Check 5 teilgenommen haben. Im Jahr 2007 haben 5530 Kinder teilgenommen. Die Teilnahme erfolgte aufgrund der freiwilligen Anmeldung einer Klasse durch die Lehrperson. Die Anzahl Schülerinnen und Schüler bzw. die Stichprobengrösse wird im Bericht mit «n» abgekürzt.

*Stichprobenfehler* – Bei einem Teil der Ergebnisse ist neben dem Mittelwert auch der Stichprobenfehler angegeben. Der Stichprobenfehler ist ein Mass für den Fehler eines empirischen Wertes (z.B. Mittelwert der Leistung) aufgrund der Tatsache, dass diese nur auf einer Stichprobe beruht. Je nach Genauigkeit der Stichprobe streut der empirische Wert in einem grösseren oder kleineren Bereich um den wahren Wert. Je kleiner der Stichprobenfehler, desto genauer ist die Schätzung eines Merkmals in der Population. Im Bericht wird der Stichprobenfehler mit «SE» abgekürzt.

*Trends* – Anhand der vorliegenden Resultate kann man die Ergebnisse in den Check-5-Tests der Jahre 2004 bis 2007 miteinander vergleichen. Bei der Interpretation von Trends müssen allerdings gewisse Einschränkungen gemacht werden:

Erstens liegen erst die Resultate von vier Testjahren vor. Um von einem interpretierbaren, längerfristigen Trend zu sprechen, müssen die Resultate von mindestens drei Testjahren auf eine gemeinsame Entwicklung verweisen. Anhand eines einmaligen Anstiegs oder einer einmaligen Reduktion der Leistungen kann noch nicht auf einen Trend geschlossen werden.

Zweitens treten zwangsläufig Stichproben- und Messfehler auf, wenn Tests durch eine beschränkte Zahl an gemeinsamer Aufgaben (Link-Items) im Zeitverlauf miteinander verknüpft werden. Veränderungen oder Unterschiede von weniger als 20 Punkten können daher nicht sinnvoll interpretiert werden, da sie zum grössten Teil auf Unterschiede in der Stichprobe (Auswahl und Merkmale der teilnehmenden Klassen, individuelle Merkmale der Schülerinnen und Schüler) und auf Messfehler zurückgeführt werden können.

## Weiterführende Informationen

Weiterführende Informationen zu Check 5 finden sich in folgenden Dokumenten:

- Keller, F. & Moser, U. (2007). Check 5. Zwischenbericht zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau.
- Keller, F. & Moser, U. (2007). Check 5. Schlussbericht 2007 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau.
- Keller, F. & Moser, U. (2006). Erste Ergebnisse Check 5. Zwischenbericht zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau.
- Keller, F. & Moser, U. (2006). Check 5. Schlussbericht 2006 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau.
- Tresch, S. & Moser, U. (2005). Check 5. Schlussbericht 2005 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau.
- Tresch, S. & Moser, U. (2005). Check 5. Schlussbericht 2005 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau.
- Moser, U. & Tresch, S. (2005). Leistungen messen und beurteilen. Handreichung zum Umgang mit Ergebnissen von Leistungstests. Herausgegeben vom Departements Bildung, Kultur und Sport des Kantons Aargau.

Alle diese Dokumente sowie weitere Informationen zum Ablauf von Check 5 und zur obligatorischen Weiterbildung stehen im Internet zur Verfügung:

- [www.ibe.uzh.ch/projekte/check5.html](http://www.ibe.uzh.ch/projekte/check5.html)
- [www.ag.ch/leistungstest](http://www.ag.ch/leistungstest)

Erste Erkenntnisse und Erfahrungen mit Check 5 sind zudem in einem Buch festgehalten:

- Tresch, S. (2007). *Potenzial Leistungstest*. Bern: h.e.p.-Verlag.