



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

VERA 3:

**Vergleichsarbeiten in der Jahrgangsstufe 3
im Schuljahr 2011/2012**

Länderbericht Berlin



Katharina Thoren

Lisa Steinke

Dominique Gönder

Peter Harych

Martin Brunner

Impressum:

Herausgeber:

Institut für Schulqualität der Länder Berlin und Brandenburg e.V. (ISQ)

Otto-von-Simson-Str. 15

14195 Berlin

Tel.: 030/844 166 8 – 0

Fax.: 030/844 166 8 – 10

Mail: info@isq-bb.de

Internet: www.isq-bb.de

Autoren/Redaktion:

Katharina Thoren

Lisa Steinke

Dominique Gönder

Peter Harych

Martin Brunner

Mit fachdidaktischer Unterstützung und Beratung von:

Petra Bittins

Bärbel Hagedorn

Anita Pfeng

Inhaltsverzeichnis

Zusammenfassung.....	1
1. VERGLEICH SARBEITEN ALS KOMPETENZBASIERTES INSTRUMENT ZUR UNTERRICHTS- UND SCHULENTWICKLUNG.....	2
2. ERFASSTE SCHÜLERDATEN.....	7
3. ERGEBNISSE IN BERLIN.....	8
3.1 DEUTSCH – <i>LESEVERSTEHEN</i>	10
Lösungshäufigkeiten im Bereich <i>Leseverstehen</i>	10
Detaillierte Analyse einzelner Aufgaben aus dem Bereich <i>Leseverstehen</i>	11
Auswertung nach Punkten der KMK-Bildungsstandards (BISTA) im Bereich <i>Leseverstehen</i>	14
Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen im Bereich <i>Leseverstehen</i>	15
3.2 DEUTSCH – <i>SPRACHGEBRAUCH</i>	17
Lösungshäufigkeiten im Bereich <i>Sprachgebrauch</i>	17
Detaillierte Analyse einzelner Aufgaben aus dem Bereich <i>Sprachgebrauch</i>	18
Auswertung nach Punkten der KMK-Bildungsstandards (BISTA) im Bereich <i>Sprachgebrauch</i>	20
Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen im Bereich <i>Sprachgebrauch</i>	21
3.3 MATHEMATIK.....	24
3.3.1 <i>Muster und Strukturen</i>	24
Lösungshäufigkeiten im Bereich <i>Muster und Strukturen</i>	24
Detaillierte Analyse einzelner Aufgaben aus dem Bereich <i>Muster und Strukturen</i>	25
Auswertung nach Punkten der KMK-Bildungsstandards (BISTA) im Bereich <i>Muster und Strukturen</i>	28
Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen in <i>Muster und Strukturen</i>	30
3.3.2 <i>Größen und Messen</i>	31
Lösungshäufigkeiten im Bereich <i>Größen und Messen</i>	31
Detaillierte Analyse einzelner Aufgaben aus dem Bereich <i>Größen und Messen</i>	33
Auswertung nach Punkten der KMK-Bildungsstandards (BISTA) im Bereich <i>Größen und Messen</i>	36
Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen in <i>Größen und Messen</i>	37
4. ÜBERGREIFENDE ANALYSEN.....	40
4.1 SOZIALE HERKUNFT – VERA-ERGEBNISSE NACH KONTEXTGRUPPEN.....	40
4.2 KORRELATIONEN UND KREUZTABELLIERUNGEN DER EINZELTESTS.....	43

Zusammenfassung

Im Mai 2012 nahmen etwa 22.600 Berliner Schülerinnen und Schüler an den bundesweit einheitlichen Vergleichsarbeiten in der Jahrgangsstufe 3 (VERA 3) teil. An drei verschiedenen Testtagen wurden in Deutsch und Mathematik Aufgaben zu verschiedenen Inhaltsbereichen bearbeitet; die Ergebnisse werden im vorliegenden Bericht deskriptiv ausgewertet. Die Leistungen der Schülerinnen und Schüler werden zum einen anhand von Lösungshäufigkeiten dargestellt, zum anderen auch Kompetenzstufen zugeordnet, die angeben, zu welchem Grad die für das Ende der Jahrgangsstufe 4 definierten Bildungsstandards in Jahrgangsstufe 3 bereits erreicht wurden (siehe auch www.isq-bb.de/vera3).

Vergleichsarbeiten werden in der Öffentlichkeit häufig noch als ein Instrument des Systemmonitorings, das vor allem der Kontrolle durch externe Institutionen dient, missverstanden. Sie sind vielmehr als Instrument der internen Evaluation von Schule und Unterricht angelegt, das Akteuren vor Ort Anstöße bietet, das professionelle Handeln datengestützt weiterzuentwickeln. Dieser Bericht ist als Ergänzung der Ergebnismeldungen zu betrachten, die jede Schule vor den Sommerferien erhalten hat. Er eröffnet Schulleitungen und Lehrkräften weitere Vergleichsmöglichkeiten.

Im Fach Deutsch *Leseverstehen* und *Sprachgebrauch* sowie in den beiden Inhaltsbereichen *Muster und Strukturen* sowie *Größen und Messen* in Mathematik werden neben den Lösungshäufigkeiten auch erreichte Kompetenzpunkte und Kompetenzstufenverteilungen berichtet. In allen drei Bereichen gelang es einem Teil der Schülerinnen und Schüler bereits am Ende der Jahrgangsstufe 3, die in den Bildungsstandards für die Jahrgangsstufe 4 definierten Regelstandards (Kompetenzstufe III) zu erreichen oder zu übertreffen. Ein besonderes Augenmerk sollte aber auch auf diejenigen Schülerinnen und Schüler gerichtet werden, denen bei VERA 3 die niedrigste Kompetenzstufe I zugeordnet wurde.

Im Bereich Deutsch *Leseverstehen* sowie *Sprachgebrauch* schnitten Mädchen durchschnittlich besser ab als Jungen, in beiden getesteten Inhaltsbereichen in Mathematik waren Jungen etwas besser als Mädchen. Im Vergleich der Leistungen zwischen den Bezirken Berlins zeigten sich für die meisten Inhaltsbereiche deutliche Unterschiede. Die Rangreihen der Bezirke nach dem Anteil der Schülerinnen und Schüler, die Kompetenzstufe II oder höher erreichen, zeigten jedoch über alle Inhaltsbereiche hinweg eine hohe Konsistenz.

Zudem ergaben sich Unterschiede zwischen deutschsprachig aufgewachsenen Kindern sowie Schülerinnen und Schülern mit nichtdeutscher Herkunftssprache, die in allen Testteilen schlechter abschnitten. Dabei scheint nicht allein die Herkunftssprache, sondern auch der sozio-ökonomische Status der Kinder ein entscheidendes Kriterium zu sein. Um den unterschiedlichen Bedingungen in den einzelnen Schulen gerecht zu werden, werden die Ergebnisse auch in diesem Jahr im Hinblick auf die Bedeutsamkeit der Zusammensetzung der Schülerschaft untersucht, wobei neben der Herkunftssprache der Schülerinnen und Schüler die Lernmittelbefreiung als Indikator für den sozio-ökonomischen Hintergrund der Schülerschaft berücksichtigt wurde. So wurden aufgrund der Ausprägung dieser beiden Merkmale drei Kontextgruppen definiert, die jeweiligen Schulen diesen zugeordnet und deren Abschneiden bei VERA 3 verglichen. Explorative Analysen zeigen, dass sich die durchschnittlichen Leistungen der drei Kontextgruppen erheblich voneinander unterscheiden, es aber in jeder Gruppe ein Teil der Schülerinnen und Schüler schafft, die für die Jahrgangsstufe 4 formulierten Erwartungen am Ende der Jahrgangsstufe 3 zu erfüllen oder sogar zu übertreffen. Schulen können die Ergebnisse derjenigen Kontextgruppe nutzen, die eine der eigenen Schule ähnliche Schülerschaft aufweist, um ihr eigenes Abschneiden „fair“ zu vergleichen. Auch enthalten die Rückmeldungen für Schulen seit dem Durchgang 2011 Vergleiche mit Schulen ähnlicher Ausgangsvoraussetzungen.

Weitere Informationen sind auch auf www.eltern.isq-bb.de und www.iqb.hu-berlin.de/vera zu finden.

1. Vergleichsarbeiten als kompetenzbasiertes Instrument zur Unterrichts- und Schulentwicklung

In ihrer **Gesamtstrategie zur Qualitätssicherung in Schulen** von 2006 legte die Kultusministerkonferenz der Länder (KMK) vier zentrale Instrumente für das Bildungsmonitoring als wesentliche Aufgabe der Bildungspolitik (KMK, 1997) fest. Hierzu zählen die Durchführung von internationalen Schulleistungsuntersuchungen (z. B.: PISA, IGLU), die zentrale Überprüfung des Erreichens der Bildungsstandards in einem Ländervergleich, die gemeinsame Bildungsberichterstattung von Bund und Ländern sowie das Durchführen von Vergleichsarbeiten in Anbindung oder Anknüpfung an die Bildungsstandards zur landesweiten Überprüfung der Leistungsfähigkeit einzelner Schulen. Die von der KMK definierten **Bildungsstandards** basieren auf allgemeinen Bildungszielen und legen fest, welche Kompetenzen Schülerinnen und Schüler bis zu einer bestimmten Klassenstufe erworben haben sollten (KMK, 2005). Dabei liegt der Fokus auf den Kernbereichen eines Faches und den in diesem Fach erwarteten Kompetenzständen.

Die **Vergleichsarbeiten in der Jahrgangsstufe 3 (VERA 3)** haben demnach die Aufgabe, das Erreichen der **Kompetenzen**, die von einem Lernenden **am Ende der Jahrgangsstufe 4** in den Fächern Deutsch und Mathematik erwartet werden, zu überprüfen. Diese Vergleichsarbeiten werden bereits in der Jahrgangsstufe 3 (VERA 3) durchgeführt, um Eltern und Lehrkräften eine kompetente rechtzeitige Zwischenrückmeldung zu geben, inwieweit ein Kind/eine Klasse die Grundschul-Standards gegenwärtig erfüllt. Somit erhalten Schülerinnen und Schüler sowie Eltern und Lehrkräfte Informationen über spezifische Kompetenzstände und können das Lernen und Lehren gezielt darauf ausrichten. Neben der Feststellung von Förderbedarf soll die aktive Beteiligung der Lehrkräfte an der Durchführung und Auswertung zu schulinterner Kooperation und Diskussion bspw. über die Bildungsstandards, die Unterrichtsgestaltung und die eigene Beurteilungspraxis anregen.

Der Fokus der Vergleichsarbeiten liegt gegenüber den anderen oben erwähnten Bausteinen der KMK-Gesamtstrategie auf dem Aspekt der **Unterrichtsentwicklung**. Nach Beschluss der KMK soll dieser Fokus explizit durch die Länder gestärkt werden¹. Die KMK verständigte sich auf eine entsprechende Vereinbarung, die Zielbestimmungen formuliert und konkrete Hinweise zur Durchführung und Rückmeldung der Ergebnisse enthält.

- So ist festgelegt, dass **VERA nicht benotet** wird, da mit dem Test die Kompetenzen von Schülerinnen und Schülern unabhängig von unmittelbar vorgeschalteten unterrichtlichen Lernprozessen und curricularen Vorgaben überprüft werden. Es ist aber fachlich vertretbar, dass Schülerinnen und Schüler sowie deren Eltern individuelle Rückmeldungen zu VERA erhalten, die in jedem Falle pädagogisch angemessen eingeordnet werden müssen. Zulässig ist hiernach nur eine Anrechnung auf die festgelegte Zahl der Klassenarbeiten, d. h. Lehrkräfte können im zweiten Schulhalbjahr eine Klassenarbeit weniger schreiben lassen, wenn sichergestellt ist, dass insgesamt ausreichend schriftliche Leistungen der Schülerinnen und Schüler als valide und nachvollziehbare Bewertungsgrundlage vorliegen. Damit kann eine zusätzliche Arbeitsbelastung der Lehrkräfte vermieden werden und VERA wird in seiner Funktion als diagnostisches Instrument nicht sachfremd verwendet.
- So wird weiterhin auf eine **Veröffentlichung der VERA-Ergebnisse einzelner Schulen verzichtet**, da dies mit der Kernfunktion der Vergleichsarbeiten, Schul- und Unterrichtsentwicklung zu betreiben, nicht zu vereinbaren ist.

¹ Vgl. KMK (2012): Vereinbarung zur Weiterentwicklung von VERA. Beschluss der Kultusministerkonferenz vom 8. März 2012 online abrufbar unter http://www.isq-bb.de/uploads/media/Beschluss_KMK_Vera_Vereinbarung_Maerz_2012.pdf.

Die Bildungsstandards werden in Form von **Kompetenzstufenmodellen** konkretisiert.² Zu jeder Kompetenzstufe gibt es Stufenbeschreibungen (textuell) in Form von Könnens-Beschreibungen (*can-do-standards*) (vgl. Abb. 1 exemplarisch für Deutsch – *Leseverstehen*).

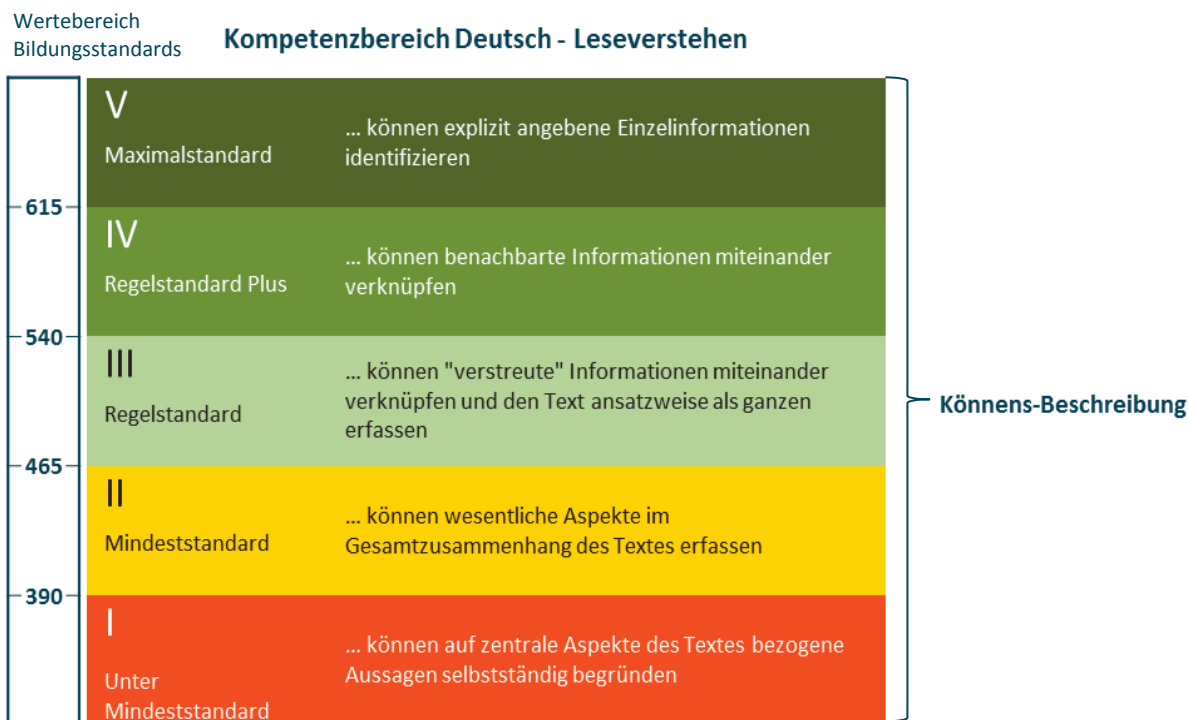


Abbildung 1: Kompetenzstufenmodell für die 4. Jahrgangsstufe im Bereich Deutsch – *Leseverstehen*

Die Kompetenzstufen im Einzelnen:

- Als durchschnittliche Erwartung an Leistungsstände gilt der **Regelstandard** (Kompetenzstufe **III**). Über die dort beschriebenen Kompetenzen sollte der Durchschnitt der Schülerinnen und Schüler am Ende der Jahrgangsstufe 4 verfügen.
- Der sogenannte **Mindeststandard** (Kompetenzstufe **II**) definiert ein Minimum an Kompetenzen, das alle Schülerinnen und Schüler bis zu einem bestimmten Bildungsabschnitt (hier zum Ende der 4. Jahrgangsstufe) erreicht haben sollten.
- Die Gruppe der Schülerinnen und Schüler **unter Mindeststandard** (Kompetenzstufe **I**) erreichen diese Mindestanforderungen nicht. Hier fehlen den Schülerinnen und Schüler basale Kenntnisse, welche für die weiterführende Schulbildung eine fundamentale Basis darstellen und für einen erfolgreichen Umgang mit darauf aufbauenden Unterrichtsinhalten unabdingbar sind. Schülerinnen und Schülern auf dieser Kompetenzstufe sollten im Fokus der Kompetenzentwicklung stehen.

Oberhalb des Regelstandards wurden zwei weitere Kompetenzstufen definiert:

- Der **Regelstandard Plus** (Kompetenzstufe **IV**) geht über die Kompetenzen des Regelstandards hinaus.

² Die jeweiligen Kompetenzstufenmodelle wurden vom IQB entwickelt und von der KMK beschlossen und können online abgerufen werden unter <http://www.iqb.hu-berlin.de/bista?reg=r> 4.

- Der **Maximalstandard** (Kompetenzstufe **V**) beschreibt Leistungserwartungen, die unter sehr guten bzw. ausgezeichneten individuellen Lernvoraussetzungen und der Bereitstellung gelingender Lerngelegenheiten innerhalb und außerhalb der Schule erreicht werden und bei Weitem die Erwartungen der Bildungsstandards übertreffen.

Die **Messung des Erreichens der Bildungsstandards** beruht auf einer Metrik (Skala), auf der man die Kompetenzstände der Schülerinnen und Schüler verorten kann. Eine solche Metrik lässt sich als ein Lineal vorstellen, welches man anlegt, um etwas zu messen. Um das gleiche Lineal bei allen Kompetenztests (Messungen) an der gleichen Stelle anzulegen, wurde die Festlegung getroffen, einen Mittelwert von 500 Punkten und eine Standardabweichung (Maß für die Streuung der Werte um den Mittelwert; legt die „Einheit“ bzw. die Messabstände des Lineals fest) von 100 Punkten anzugeben. Diese Metrik wird kurz als **BISTA-Skala** (Messskala der Bildungsstandards) bezeichnet. Sie wird in allen mit den Bildungsstandards verbundenen Tests (wie PISA, VERA, Ländervergleich) verwendet. Legt man diese BISTA-Skala an die beschriebenen Kompetenzstufen an, so lassen sich für bestimmte Kompetenzbereiche Kompetenzstufengrenzen bestimmen. Diese Kompetenzstufengrenzen wurden von der KMK und dem IQB in einem umfangreichen und langwierigen Prozess unter Mitarbeit von erfahrenen Lehrkräften, Fachdidaktikern und Wissenschaftlern festgelegt.

Die Abbildung 2 zeigt VERA-3-Beispielaufgaben und ihre Verortung auf der Skala der Bildungsstandards für den Bereich *Muster und Strukturen* des Fachs Mathematik.

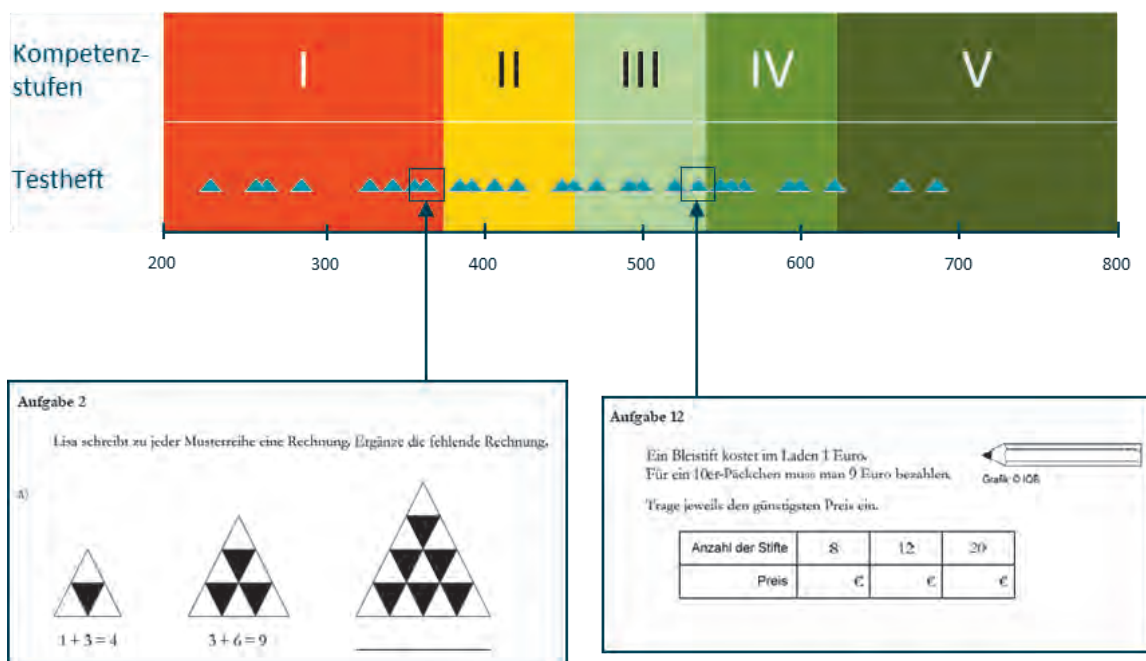


Abbildung 2: Testheftkonzeption und Aufgabenschwierigkeit (Beispiel Mathematiktest). Abgebildet sind die jeweiligen Aufgaben des Testheftes (Dreiecke) und deren Verteilung gemäß ihrer mittleren Schwierigkeit auf der Skala der Bildungsstandards (BISTA)

Es steht jeweils eine **Testheftversion** pro Fach zur Verfügung. Jeder dieser **VERA-3-Aufgaben** kann aufgrund ihrer bekannten Schwierigkeit ein bestimmter **Punktwert auf der Skala der Bildungsstandards** zugeordnet werden. Aus der Bearbeitung der Aufgaben können dann Rückschlüsse über die bildungsstandardbezogenen Kompetenzen der Schülerinnen und Schüler getroffen werden. Hinzu kommt, dass die Skala der Bildungsstandards fachbezogen in Kompetenzstufen eingeteilt wird, sodass auch eine Zuordnung zu diesen Kompetenzstufen vorgenommen werden kann.

Bei der **Testheftzusammenstellung** werden Aufgaben so ausgewählt, dass sie verschiedene Schwierigkeiten abbilden, um auch die Leistungen (sehr) schwacher und (sehr) starker Schülerinnen und Schüler differenziert erfassen zu können. Für den jeweiligen Einsatz verschiedener Testheftversionen bedeutet dies, dass die erwartete mittlere Lösungshäufigkeit pro Testheft idealerweise bei ca. 50-65 % liegen sollte.

Die Vergleichsarbeiten in der Jahrgangsstufe 3 in den Fächern **Deutsch** und **Mathematik** fanden 2012 bundesweit einheitlich am 8. (*Leseverstehen*), 10. (*Sprachgebrauch*) und 15. (Mathematik; beide Domänen) Mai statt. Verantwortlich für die Aufgabenentwicklung und Erarbeitung der didaktischen Materialien ist seit dem VERA-3-Durchgang 2009/2010 das **Institut zur Qualitätsentwicklung im Bildungswesen (IQB)** an der Humboldt-Universität zu Berlin (für weitere Informationen siehe <http://www.iqb.hu-berlin.de/vera>).

Alle Schülerinnen und Schüler öffentlicher Grundschulen in Berlin bearbeiteten im Fach Deutsch obligatorisch Aufgaben zu den in den KMK-Bildungsstandards formulierten Inhaltsbereichen *Leseverstehen* und *Sprachgebrauch*. Der Mathematiktest umfasste Aufgaben zu den beiden Inhaltsbereichen *Muster und Strukturen*³ und *Größen und Messen*. Schulen in nichtöffentlicher Trägerschaft hatten die Möglichkeit, freiwillig an den Vergleichsarbeiten teilzunehmen.

Die Lehrkräfte der teilnehmenden Schulen in Berlin übernahmen die **Durchführung** und **Auswertung** der Vergleichsarbeiten auf Grundlage standardisierter Vorgaben. Unterstützt wurden sie hierbei vom **Institut für Schulqualität der Länder Berlin und Brandenburg (ISQ)**, einem An-Institut der Freien Universität Berlin (FU). Das ISQ versendete die Test- und Auswertungsmaterialien an die Schulen und bot mit vorbereitenden **Informationsveranstaltungen** (siehe Kapitel 1) und einer **Telefon-Hotline** konkrete Beratung bei Fragen der Testdurchführung und Dateneingabe an. Die Schulen übermittelten ihre Ergebnisse der Vergleichsarbeiten im passwortgeschützten **ISQ-Internetportal**.

Der vorliegende Bericht über die Ergebnisse der Schülerinnen und Schüler des Landes Berlin geht auf die Vergleichsarbeiten der Jahrgangsstufe 3 im Schuljahr 2011/12 ein und beschränkt sich dabei auf die zur Teilnahme verpflichteten Schülerinnen und Schüler aus öffentlichen Grundschulen⁴. Die Rückmeldung der Ergebnisse an die Schulen erfolgt auf Basis der Kompetenzmodelle der Kultusministerkonferenz (KMK), welche sich an den Bildungsstandards für die Primarstufe orientieren (KMK, 2004).

Informationsveranstaltungen

Die Länder Berlin und Brandenburg beauftragten das ISQ mit der Administration der Tests und der Unterstützung der Schulen bei der Durchführung der Vergleichsarbeiten. Im März und April 2012 veranstaltete das ISQ in beiden Ländern insgesamt drei Informationsveranstaltungen. Im ersten Teil dieser Veranstaltungen referierten Mitarbeiterinnen und Mitarbeiter des ISQ über die den Tests zugrunde liegenden Bildungsstandards und Kompetenzstufenmodelle, den Ablauf von VERA 3 und stellten das ISQ-Portal zur Dateneingabe vor. Nach diesen kurz gehaltenen Vorträgen wurden im zweiten Teil der Veranstaltung 90-minütige Workshops von Referentinnen und Referenten des LISUM angeboten:

³ Im Berliner Rahmenlehrplan wird dieses Themengebiet nicht als eigener Bereich geführt, sondern ist Bestandteil aller Themengebiete der Mathematik.

⁴ Auf die Darstellung der Ergebnisse aus Schulen in freier Trägerschaft wird im Rahmen dieses Ergebnisberichts verzichtet. Schüler/-innen aus diesen Schulen in freier Trägerschaft waren nicht teilnahmeverpflichtet und ihre Daten wurden nur zum Teil und freiwillig zur Verfügung gestellt, sodass keine belastbaren Aussagen über das Abschneiden dieser Schulen getroffen werden können.

1) Aufgabenkultur im Fach Deutsch (LISUM)

Den inhaltlichen Schwerpunkt bildete hier die Analyse der Ergebnisse aus den Vergleichsarbeiten. Hierzu wurde ein diagnostischer Blick auf exemplarische Schülerantworten aus dem Test des Vorjahres geworfen und diskutiert, wie Fehler als „Fenster in Kinderköpfen“ gesehen und für die Weiterarbeit im Unterricht genutzt werden können.

2) Aufgabenkultur im Fach Mathematik (LISUM)

In diesem Workshop wurden beispielhaft konkrete Aufgaben und Inhalte dargestellt, die einen kompetenzorientierten Mathematikunterricht kennzeichnen, außerdem ein Überblick über die Bildungsstandards gegeben. Schwerpunkte dieses Workshops bildeten die getesteten Inhaltsbereiche *Muster und Strukturen* sowie *Größen und Messen*.

Der Besuch einer ISQ-Informationsveranstaltung wurde als Teilnahme an einer Fortbildungsveranstaltung anerkannt.

2. Erfasste Schülerdaten

Insgesamt konnten für Berlin aus allen öffentlichen Grundschulen und teilnehmenden Schulen in freier Trägerschaft die Daten von 21.974 Tests in Deutsch *Leseverstehen*, 21.992 Tests in Deutsch *Sprachgebrauch* und 21.996 Mathematiktests erfasst werden (vgl. Tabelle 1). Die unterschiedliche Anzahl der erfassten Testleistungen erklärt sich aus den unterschiedlichen Testzeitpunkten an zwei Testtagen im Jahr 2012. Der Sprachtest in Deutsch war für Berliner Schüler/-innen dieses Jahr verpflichtend.

Tabelle 1: Erfasste Schülerdaten differenziert nach Klassen, Schulen⁵

Teilnahme	Schulen	Klassen	Schüler/-innen <i>Leseverstehen</i>	Schüler/-innen <i>Sprachgebrauch</i>	Schüler/-innen Mathematik
Öffentliche Schulen	373	1.414	20.979	20.995	20.984
Schulen in freier Trägerschaft	29	62	995	997	982
Summe	402	1.476	21.974	21.992	21.966

An den Schulen in freier Trägerschaft zeichnet sich auch in diesem Jahr eine hohe Akzeptanz des VERA-Tests als freiwillig genutztes diagnostisches Verfahren ab. In Berlin nutzten dieses Jahr 29 Schulen in freier Trägerschaft mit insgesamt rund 1.000 Schülerinnen und Schülern das zentral administrierte Test- und Auswertungsverfahren und damit landesweite Vergleichsmaßstäbe (vgl. Tabelle 1).

Tabelle 2: Zusammensetzung der Teilnehmerinnen und Teilnehmer an den öffentlichen Schulen Berlins bei VERA 3 im Schuljahr 2011/12

	Anzahl	Anteil
Eingegebene Stammdaten	22.635	
Zur Teilnahme verpflichtet	21.953	100,0 %
davon weiblich	10.774	49,1 %
davon nichtdeutsche Herkunftssprache (ndH)	6.802	31,0 %
davon Teilleistungsstörung Deutsch	1.179	5,4 %
davon Teilleistungsstörung Mathematik	490	2,2 %
<i>Leseverstehen</i> - gewertete Schüler/-innen	20.738	94,5 %
<i>Sprachgebrauch</i> - gewertete Schüler/-innen	20.754	94,5 %
Mathematik - gewertete Schüler/-innen	20.724	94,4 %

⁵ Für die im Datensatz erfassten Schüler/-innen schwanken die gültigen Antworten in Deutsch und Mathematik, wodurch sich die den Auswertungen zugrunde liegenden Fallzahlen unterscheiden können.

Tabelle 2 zeigt die Zusammensetzung der Schülerschaft an den öffentlichen Schulen Berlins. Knapp die Hälfte der Schüler/-innen waren Mädchen, rund 31 % sprechen nach Angabe der Lehrkraft zu Hause hauptsächlich eine andere Sprache als Deutsch, was in den Rückmeldungen als „Deutsch nicht dominant“, in diesem Bericht aber im Folgenden als nichtdeutsche Herkunftssprache (ndH) bezeichnet wird. Nach Angabe der Lehrkräfte weisen rund 5,4 % der Schülerinnen und Schüler eine Teilleistungsstörung im Fach Deutsch auf und ca. 2,2 % der Kinder eine Teilleistungsstörung im Fach Mathematik.

3. Ergebnisse in Berlin

Die vorliegende Darstellung der Ergebnisse gliedert sich wie folgt: Zunächst werden nacheinander die Ergebnisse für die Inhaltsbereiche *Leseverstehen* und *Sprachgebrauch* im Fach Deutsch dargelegt. Anschließend folgt die Darstellung der Ergebnisse in den zwei Inhaltsbereichen *Muster und Strukturen* und *Größen und Messen* im Fach Mathematik.

Dabei umfassen die Ergebnisdarstellungen für beiden Deutsch-Inhaltsbereiche sowie die beiden Mathematik-Inhaltsbereiche jeweils drei Teile:

- **Lösungshäufigkeiten** geben an, von welchem Anteil der Schülerinnen und Schüler jede Aufgabe gelöst werden konnte. Bei der Testheftzusammenstellung wird sichergestellt, dass Aufgaben verschiedener Schwierigkeiten berücksichtigt werden, wobei die durchschnittlich zu erwartende Lösungshäufigkeit über alle Aufgaben eines Testheftes hinweg bei ca. 50-65 % liegt.
- Für die Berechnung der **Kompetenzpunkte** auf der Bildungsstandard-Metrik (im Folgenden **BISTA-Punkte**) bildet das probabilistische Testverfahren der Item-Response-Theorie (IRT) die methodische Grundlage. In Anlehnung an die internationale Bildungsforschung wurde eine normative Festlegung der Kompetenzskala vorgenommen. Der Mittelwert der Kompetenzskala beträgt $M=500$ Punkte, die Standardabweichung $SD=100$ Punkte.
- Zur Festlegung der **Kompetenzstufen** wird die Kompetenzskala in fünf aufeinander aufbauende Stufen untergliedert, für die sich typische Aufgaben und notwendige Fähigkeiten beschreiben lassen.

Differenziert dargestellt werden im Folgenden für die beiden getesteten Deutsch- und Mathematik-Inhaltsbereiche jeweils Unterschiede zwischen den Berliner Bezirken, aber auch in Bezug auf das Geschlecht und die Herkunftssprache der Schüler/-innen. Im Anschluss an die deskriptiven Ergebnisse werden die Verteilungen auf den Kompetenzstufen berichtet, wiederum differenziert nach Geschlecht, Herkunftssprache und Bezirk.

Weder beim Vergleich der Ergebnisse auf Ebene der Bezirke noch beim Vergleich anderer Werte werden statistische Signifikanzen berichtet. Stattdessen wird, sofern angebracht, auf praktisch bedeutsame Unterschiede hingewiesen, da nur diese bei flächendeckenden Erhebungen sinnvoll interpretiert werden können.

Zusätzlich zu den Testunterlagen erhielten alle Brandenburger Schulen die **didaktischen Handreichungen** in Deutsch und Mathematik in gedruckter Form.

Diese umfassten im Durchgang 2011/12 jeweils drei Module:

- Modul A enthielt fachübergreifende Erläuterungen zu VERA 3,
- Modul B umfasste Erläuterungen zum jeweiligen Kompetenzbereich,
- Modul C beinhaltete Kommentare und Hinweise zu jeder der Aufgaben.

Im vorliegenden Bericht werden für jeden der getesteten Inhaltsbereiche einige interessante Aufgaben fachdidaktisch erläutert und – sofern sinnvoll – um Hinweise für die Unterrichtsgestaltung und Weiterarbeit mit der eigenen Klasse ergänzt. Kommentierungen aller Aufgaben finden sich in den didaktischen Handreichungen, die bei der Auswertung der Ergebnisse hinzugezogen werden sollten.

3.1 Deutsch – *Leseverstehen*

Der Test zum Inhaltsbereich *Leseverstehen* fand am 08.05.2012 statt und bestand aus zwei Texten: Einem kontinuierlichen literarischen Text (fiktional) mit zehn anschließend zu beantwortenden Aufgaben und einem mittellangen Sach- und Gebrauchstext in expositorischer Form mit hoher Informationsdichte, zu dem 13 Fragen beantwortet werden mussten. In diesem Testbereich konnten maximal 23 Punkte erreicht werden. Die Schülerinnen und Schüler hatten für die Bearbeitung des Lesetests 40 Minuten Zeit.

Lösungshäufigkeiten im Bereich *Leseverstehen*

Tabelle 3: Itemgenaue Lösungshäufigkeiten im Bereich *Leseverstehen*
(alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

	Aufgabe	Berlin (n= 20.738)	Jungen (n= 10.557)	Mädchen (n= 10.181)	Deutsch (n= 14.425)	ndH (n= 6.313)
1.1	Schafherde	95%	94%	95%	96%	91%
1.2	Grund für Verkleidung	86%	85%	87%	89%	79%
1.3	Zeit für Verkleidung	72%	71%	74%	78%	61%
1.4	Gefühl am Ende	77%	74%	80%	81%	68%
1.5	Pelz verloren	53%	52%	55%	59%	39%
1.6	richtige Reihenfolge	59%	59%	60%	67%	42%
1.7	Regenmantel & Wurst	76%	75%	77%	82%	64%
1.8	Wolf nicht gefährlich	35%	33%	36%	39%	25%
1.9	Lippen lecken	72%	71%	73%	77%	58%
1.10	Schafschur	49%	48%	50%	58%	28%
2.1	Zettel abzugeben bei	79%	77%	82%	82%	74%
2.2	Bekanntgabe Ausflugsziel	53%	51%	55%	59%	40%
2.3	schminken lassen	77%	73%	81%	81%	69%
2.4	Schatzsuche	80%	79%	81%	85%	69%
2.5	Spielgerät Angebot 4	76%	74%	78%	81%	65%
2.6	Zoo - Was stimmt?	59%	57%	60%	64%	45%
2.7	Marvin 1.Klasse	60%	57%	62%	66%	46%
2.8	Kinder dürfen ...	71%	69%	72%	76%	58%
2.9	Angebote Viertklässler	37%	36%	37%	42%	25%
2.10	Angebote für Lisa	19%	17%	21%	22%	13%
2.11	Wo befindet sich der Zoo?	71%	69%	73%	76%	59%
2.12	ungerecht / begründe	25%	23%	28%	30%	15%
2.13	Flughafen-Angebot	66%	67%	66%	71%	57%
	Gesamt	63%	61%	64%	68%	52%

Etwa 63 % der insgesamt 23 Aufgaben wurden landesweit richtig gelöst, wobei Mädchen mit 64 % im Durchschnitt eine drei Prozent höhere Lösungshäufigkeit aufwiesen als Jungen mit 61 %. Kinder, die zu Hause nur oder hauptsächlich Deutsch sprechen („Deutsch“), lösten durchschnittlich 68 % der Aufgaben im *Leseverstehen* korrekt. Schülerinnen und Schüler, die zu Hause überwiegend eine andere

Sprache als Deutsch sprechen („ndH“), lösten im Durchschnitt 52 % der Aufgaben korrekt. Sie machten bei den Vergleichsarbeiten im Inhaltsbereich *Leseverstehen* in der teilnehmenden Berliner Schülerschaft 2012 einen Anteil von rund 30 Prozent aus.

Abbildung 3 verdeutlicht die **Verteilung der erreichten Gesamtpunktzahlen** (von 23 möglichen Punkten) im *Leseverstehen* über alle Schülerinnen und Schüler Berlins hinweg. Rund 25 % der Kinder erlangten demnach mindestens 18 Punkte. Weniger als die Hälfte der Gesamtpunktzahl erreichten hingegen knapp 30 Prozent der teilnehmenden Schülerinnen und Schüler. Durchschnittlich wurden 15 der 23 Aufgaben korrekt gelöst, was einer mittleren Lösungshäufigkeit von 66 % entspricht.

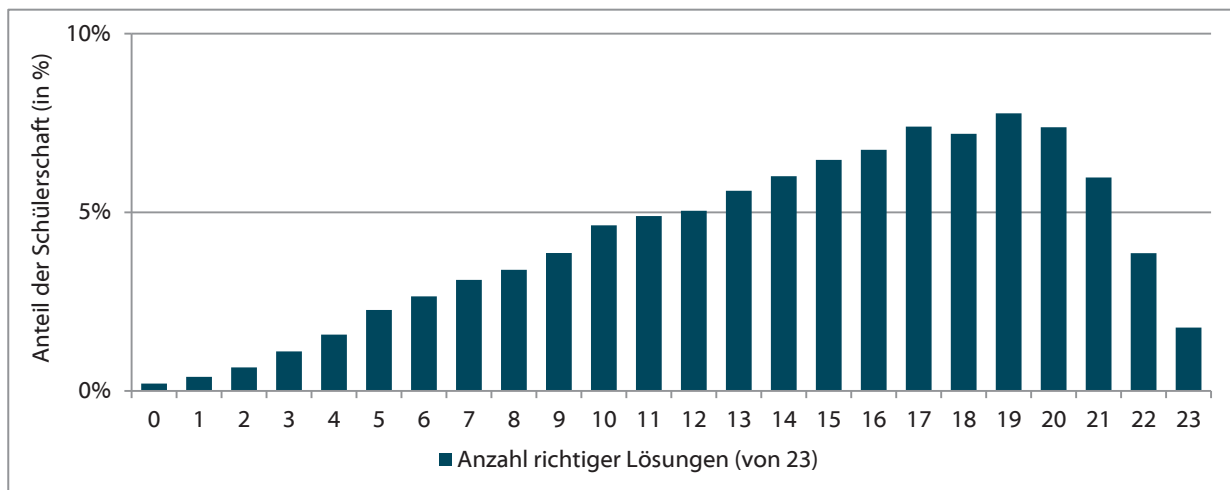


Abbildung 3: Verteilung der Testwerte in Punktsummen für Deutsch, Inhaltsbereich *Leseverstehen*

Detaillierte Analyse einzelner Aufgaben aus dem Bereich *Leseverstehen*

Blick auf die leichteste Aufgabe - Aufgabe 1.1 (literarischer Text)

1.1 Warum schleicht Wilfried um eine Schafherde herum?

- weil er Freunde sucht
- weil er Hunger hat
- um den Schäfer zu treffen
- um ein Würstchen zu bekommen


Die leichteste Aufgabe des Tests nach den Ergebnissen der Berliner Schülerinnen und Schüler zielte auf eine zentrale Information zum Motiv der Hauptfigur. Es musste auch Sprachwissen zum Einsatz gebracht werden. Die hohe Lösungshäufigkeit von 95 % verdeutlicht, dass die Aufgabe gut als Einstieg in den Test geeignet war. Bei dieser Aufgabe zeigten sich kaum Unterschiede zwischen Jungen und

Mädchen. Sie wurde von Kindern nichtdeutscher Herkunftssprache zu 91 % richtig gelöst. Die Lösungshäufigkeiten bei den Kindern deutscher Herkunftssprache liegt bei 96 %.

Blick auf die schwierigste Aufgabe - 1.8 (literarischer Text)

Die schwierigste Aufgabe des literarischen Textes wurde von 35 % der Berliner Schülerinnen und Schüler gelöst. 39 % der Kinder deutscher Herkunftssprache und 25 % der Kinder nichtdeutscher Herkunftssprache lösten diese Aufgabe korrekt. Die Schwierigkeit dieser Aufgabe liegt vermutlich in der schriftlichen Argumentation begründet, da diese im Unterricht der dritten Jahrgangsstufe noch vorrangig mündlich

1.8 Kemal sagt: „Der Wolf ist gar nicht so gefährlich, wie man am Anfang der Geschichte denkt.“
Stimmst du ihm zu? Begründe.

 _____

stattfindet. Für die Beantwortung der Frage musste der Text als Ganzes beurteilt und die Gefährlichkeit des Wolfes eingeschätzt werden.

Blick auf die leichteste Aufgabe - Aufgabe 2.4 (mehnteiliger Sach- und Gebrauchstext)

Bei der leichtesten Aufgabe des Sach- und Gebrauchstextes musste eine Information ermittelt werden, die fast wortgleich im Text zu finden war. Auch handelte es sich hierbei um eine Multiple-Choice-Aufgabe, in der aus verschiedenen Antwortoptionen auszuwählen war. Die Aufgabe wurde von 80 % der Berliner Schülerinnen und Schüler gelöst, wobei sich größere Unterschiede in den Ergebnissen der Kinder deutscher (85 %) und nichtdeutscher (69 %) Herkunftssprache abzeichneten.

Blick auf die schwierigste Aufgabe - Aufgabe 2.10 (mehnteiliger Sach- und Gebrauchstext)

2.10 Lisa ist in der 2. Klasse und will am liebsten entweder Schwimmen lernen oder sich verkleiden oder basteln.

Welche Angebote kommen in Frage?



Die Schwierigkeit der Aufgabe 2.10 liegt einerseits im offenen Format und andererseits in der Überprüfung der Stimmigkeit mehrerer Wünsche begründet. Es muss-

te die Antwort selbst produziert/aufgeschrieben sowie ein Abgleich mit verschiedenen Informationen durchgeführt werden. Den Berliner Schülerinnen und Schülern gelang dies zu 19 %. Allerdings wurde diese Aufgabe von 19 % der Berliner Kinder nicht bearbeitet. Betrachtet man in diesem Zusammenhang die letzte Aufgabe im Testheft, die von 14 % der Schülerinnen und Schüler nicht bearbeitet wurde, lässt sich vermuten, dass die Aufgabe nicht aus Zeitgründen, sondern eher aus Formatgründen unbearbeitet blieb.

Blick auf unbearbeitete Aufgaben

Betrachtet man die unbearbeiteten Aufgaben, lässt sich erkennen, dass wie in den letzten Jahren offene Aufgaben weniger bearbeitet wurden. Bei den jeweils letzten Aufgaben im Testheft darf allerdings der Zeitfaktor nicht außer Acht gelassen werden. Aufgabe 2.12

2.12 Kerim sagt: „Das ist ungerecht, dass man nicht seine erste Wahl bekommt. Vielleicht muss man dann beim Ausflug etwas machen, was man gar nicht so gerne mag.“

Stimmst du ihm zu? Begründe mit Hilfe des Textes.



wurde von 34 % der Berliner Schülerinnen und Schüler nicht bearbeitet. Diese Aufgabe war auf Kompetenzstufe V angesiedelt und verlangte eine begründete Meinungsäußerung. Im Vergleich zu der bereits beschriebenen Aufgabe 1.8 (ebenfalls Kompetenzstufe V) sollte hier ausführlich und nicht in Form einer Kurzantwort geantwortet werden. Die Kurzantwort wurde von 25 % der Berliner Kinder nicht gegeben. Bei Kindern nichtdeutscher Herkunftssprache fallen unter dem Aspekt der Nichtbearbeitung Aufgabe 1.10 (letzte Aufgabe im literarischen Text) und Aufgabe 2.12 (auf Kompetenzstufe V mit begründeter Meinung) auf.

Blick auf Aufgabenformate

Wie in den letzten Jahren wurden Multiple-Choice-Aufgaben häufiger bearbeitet als offene Aufgaben. Die im Test enthaltenen offenen Aufgaben lassen sich den Kompetenzstufen I, III und V zuordnen. Auffällig ist, dass auch die leichteren offenen Aufgaben zum Teil ebenso nicht bearbeitet wurden wie die schweren offenen Aufgaben (vgl. Aufgabe 2.10 auf Kompetenzstufe V und Aufgabe 2.7 auf Kom-

petenzstufe III, die jeweils von 19 % der Berliner Kinder nicht bearbeitet wurden). Bei der Interpretation der Ergebnisse muss die Lehrkraft prüfen, inwieweit Aufgaben zur schriftlichen Argumentation bzw. die Verschriftlichung von Antworten bereits gängige Unterrichtspraxis sind.

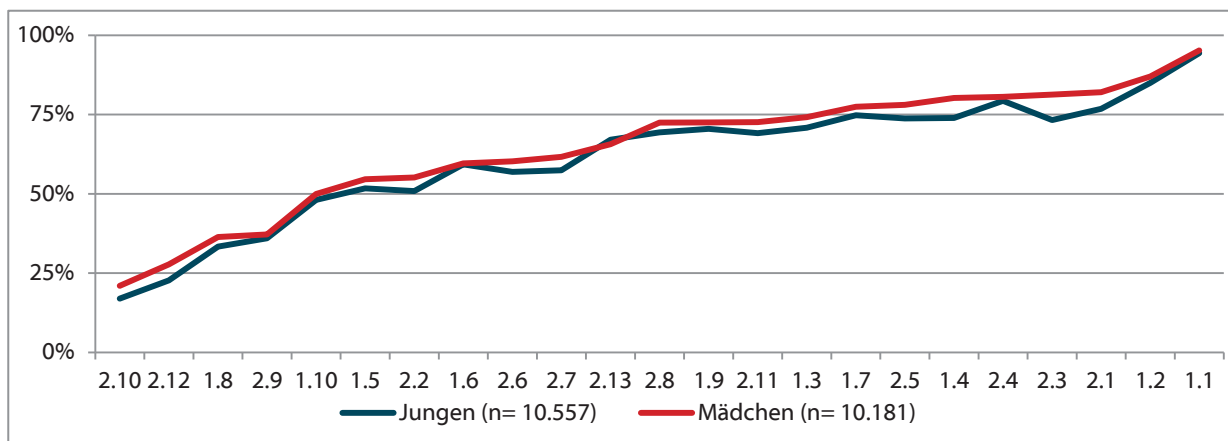


Abbildung 4: Itemspezifische Lösungshäufigkeiten in Deutsch – *Leseverstehen*, nach Geschlecht (Items geordnet nach der Schwierigkeit der Aufgaben für die Mädchen)

Blick auf Unterschiede zwischen Jungen und Mädchen

Der geschlechtsspezifische Vergleich verdeutlicht, dass Mädchen bis auf Aufgabe 2.13 leicht besser abschnitten als Jungen. Mit einer maximalen Differenz von fünf bis acht Prozentpunkten (z. B. bei Aufgabe 2.3) zeigt sich zwischen beiden Geschlechtern nur ein relativ geringer Unterschied in den Lösungshäufigkeiten.

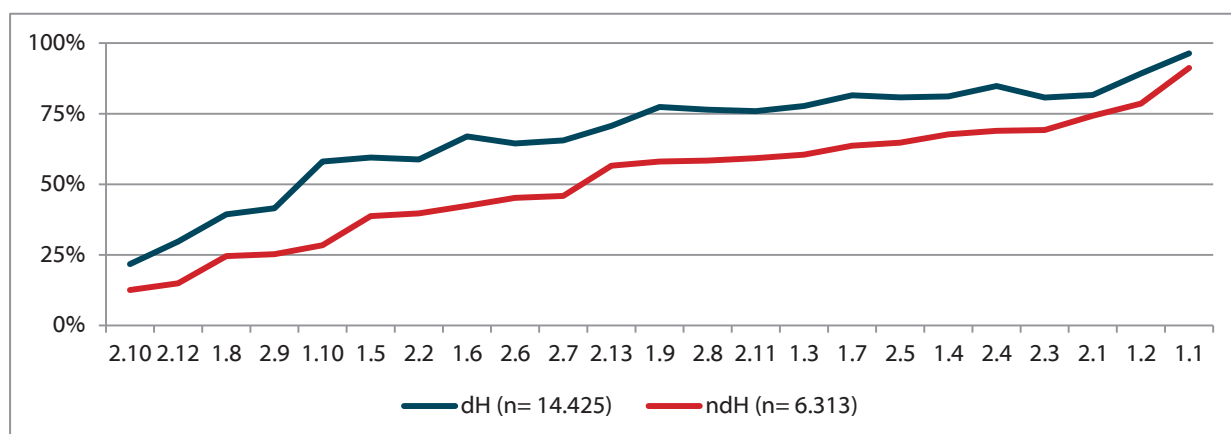


Abbildung 5: Itemspezifische Lösungshäufigkeiten in Deutsch – *Leseverstehen*, nach Herkunftssprache (Items geordnet nach der Schwierigkeit der Aufgaben für Kinder mit ndH)

Blick auf die Ergebnisse von Kindern deutscher und nichtdeutscher Herkunftssprache

Betrachtet man in Abbildung 5 die Unterschiede zwischen Kindern deutscher und nichtdeutscher Herkunftssprache, fallen diese im Gegensatz zur geschlechtsspezifischen Darstellung deutlich stärker aus, wobei sich vor allem bei den schwierigen Items nicht selten Differenzen von 15 bis zu 20 Prozentpunkten finden (z. B. 1.5, 1.6, 2.6). Bei Aufgabe 1.10 liegen die Lösungshäufigkeiten sogar 30 % auseinander.

inander. Diese Differenz kann auf den Aufgabentyp zurückzuführen sein, da dieses Item eine freie Antwortformulierung verlangte, was Kindern nichtdeutscher Herkunftssprache offenbar größere Schwierigkeiten bereitet. Durchschnittlich lösten Kinder nichtdeutscher Herkunftssprache 52 % aller Items richtig, wohingegen Schülerinnen und Schüler mit Deutsch als Muttersprache im Durchschnitt 68 % aller Aufgaben des Lesetests richtig lösten, was einer Differenz von durchschnittlich 3,6 mehr richtig beantworteten Items entspricht.

Abschließend werden die Items des Lesetests auch im Hinblick auf **Nichtbearbeitung** analysiert. Hierzu wird in Abbildung pro Aufgabe unterschieden, ob sie richtig, falsch oder nicht bearbeitet wurde. Zur Bearbeitung des Lesetests standen pro Text 20 Minuten, also insgesamt 40 Minuten zur Verfügung. Die dargestellte Verteilung zeigt, dass vor allem die Aufgaben am häufigsten nicht bearbeitet wurden, die der Kompetenzstufe V (Maximalstandard) zugeordnet wurden (Bsp. Aufgaben 1.8; 2.12).

Auffällig ist die hohe Nichtbearbeitung (27 %) von Aufgabe 1.10, die dem Regelstandard (Kompetenzstufe III) zugeordnet war und daher erwartungsgemäß häufiger bearbeitet werden sollte. Da diese Aufgabe jedoch am Ende des Aufgabenblocks 1 liegt und die Schülerinnen und Schüler an dieser Stelle warten sollten, bis sie zum Umblättern auf die nächste Seite aufgefordert wurden, könnten Unruhe und Ungeduld einige dazu veranlasst haben, die Aufgabe nicht mehr zu bearbeiten. Sonstige mögliche Auffälligkeiten wie eine erhöhte Nichtbearbeitung der Aufgaben gegen Ende des Testheftes lassen sich im Bereich *Leseverstehen* nicht feststellen.

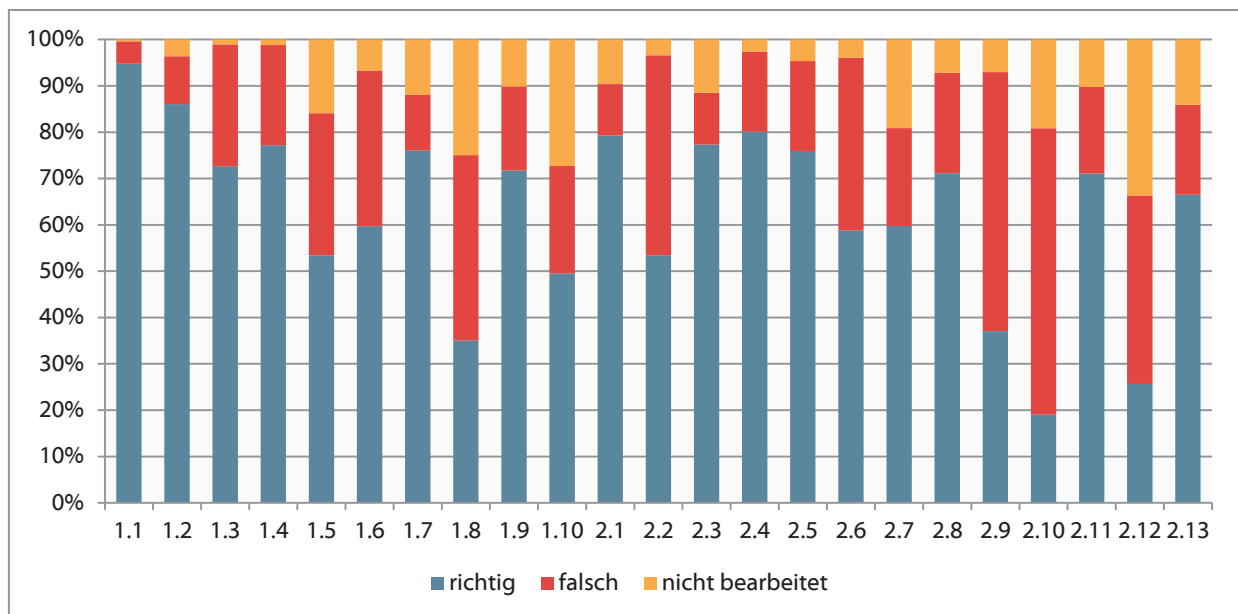


Abbildung 6: Bearbeitung der Aufgaben in Testreihenfolge im Testheft *Leseverstehen*

Auswertung nach Punkten der KMK-Bildungsstandards (BISTA-Punkte) im Bereich *Leseverstehen*

In Tabelle 4 sind die Ergebnisse im *Leseverstehen* auf Basis der BISTA-Punkte (vgl. Kapitel 3) differenziert nach Bezirk, Geschlecht und Herkunftssprache aufgelistet. Demnach lassen sich bei den verschiedenen Betrachtungsebenen erhebliche Unterschiede erkennen. Das Punktzahlspektrum erstreckte sich von 415 Punkten in Neukölln bis hin zu 537 Punkten in Pankow.

Die Ergebnisse nach geschlechtsspezifischer Betrachtung weisen zwischen den verschiedenen Bezirken deutliche Unterschiede auf. In ganz Berlin schnitten Mädchen besser ab als Jungen. Die BISTA-Punkte variierten zwischen 14 Punkten Unterschied in Spandau bis zu 35 Punkten im Bezirk Pankow.

Bei Betrachtung der Schülerleistungen nach Herkunftssprache auf Basis der BISTA-Punktzahlen sind noch eindeutigere Differenzen festzustellen. Hier schnitten Schülerinnen und Schüler mit Deutsch als Muttersprache in allen zwölf Berliner Bezirken weitaus besser ab als Kinder mit nichtdeutscher Herkunftssprache. Die Differenzen beliefen sich auf 53 Punkte in Lichtenberg und sind mit 143 Punkten im Bezirk Friedrichshain-Kreuzberg insgesamt am höchsten.

Tabelle 4: Durchschnittliche BISTA-Punkte im Bereich *Leseverstehen*
(alle Schüler/-innen, nach Gruppen differenziert, Bezirksergebnisse)

		Berlin		Geschlecht				Herkunftssprache			
		Gesamt (n= 20.738)		männlich (n= 10.557)		weiblich (n= 10.181)		Deutsch (n= 14.425)		nicht Deutsch (n= 6.313)	
Nr.	Bezirk	n	M	n	M	n	M	n	M	n	M
1	Mitte	2.017	421	1.014	412	1.003	429	711	487	1.306	385
2	Friedrichshain-Kreuzberg	1.611	477	825	467	786	487	889	541	722	398
3	Pankow	2.317	537	1.230	520	1.087	555	2.175	542	142	459
4	Charlottenburg-Wilmersdorf	1.591	513	824	496	767	530	1.016	544	575	457
5	Spandau	1.393	450	707	443	686	457	892	480	501	397
6	Steglitz-Zehlendorf	1.798	535	924	521	874	550	1.509	551	289	452
7	Tempelhof-Schöneberg	1.960	489	973	480	987	499	1.268	524	692	427
8	Neukölln	1.976	415	1.004	404	972	426	928	458	1.048	377
9	Treptow-Köpenick	1.433	501	691	487	742	514	1.353	507	80	403
10	Marzahn-Hellersdorf	1.556	442	793	433	763	452	1307	452	249	390
11	Lichtenberg	1.423	483	721	473	702	493	1.148	493	275	440
12	Reinickendorf	1.663	468	851	461	812	476	1229	496	434	391
Gesamt		M= 478	SD=153	M= 467	SD=150	M= 489	SD=155	M= 510	SD=149	M= 405	SD=134

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen im Bereich *Leseverstehen*

Neben der Darstellung von Lösungshäufigkeiten bzw. BISTA-Werten pro Aufgabe können Schülerleistungen einer bestimmten Kompetenzstufe (KS) zugeordnet werden. Unterschieden wird dabei zwischen Unter Mindeststandard (KS I), Mindeststandard (KS II), Regelstandard (KS III), Regelstandard Plus (KS IV) und Maximalstandard (KS V). Die Abbildung 7 zur **Kompetenzstufenverteilung** der Berliner Schülerinnen und Schüler im Inhaltsbereich *Leseverstehen* verdeutlicht, dass in Berlin über 70 % der Schülerinnen und Schüler mindestens die Kompetenzstufe II und somit die Mindeststandards für das Ende der Jahrgangsstufe 4 bereits in der 3. Jahrgangsstufe erreichten und z. T. Leistungen darüber hinaus zeigten. Die geschlechtsspezifische Betrachtung lässt bei der Verteilung der Kompetenzstufen lediglich Unterschiede auf Kompetenzstufe I und V erkennen, wobei Jungen (31 %) einen fünf Prozentpunkte größeren Anteil auf Kompetenzstufe I ausmachen als Mädchen (26 %). Dagegen ist mit 21 % ein etwas größerer Anteil der Mädchen auf dem höchsten Kompetenzstufenniveau zu verorten als bei den Jungen (17 %). Sehr deutlich treten die Differenzen hinsichtlich der Kompetenzstufenverteilung nach Herkunftssprache hervor. So erreichten fast die Hälfte der Kinder mit nichtdeutscher Her-

kunftssprache nicht die erforderlichen Mindeststandards im *Leseverstehen* für das Ende der Klassenstufe 4, wohingegen dieser Anteil bei Kindern mit Deutsch als Muttersprache nur 20 % ausmacht. Möglich wäre z. B., dass dieser Unterschied auf geringeres Lesetempo zurückzuführen ist oder es erhebliche Schwierigkeiten beim Verständnis und der Bearbeitung der Aufgaben gab.

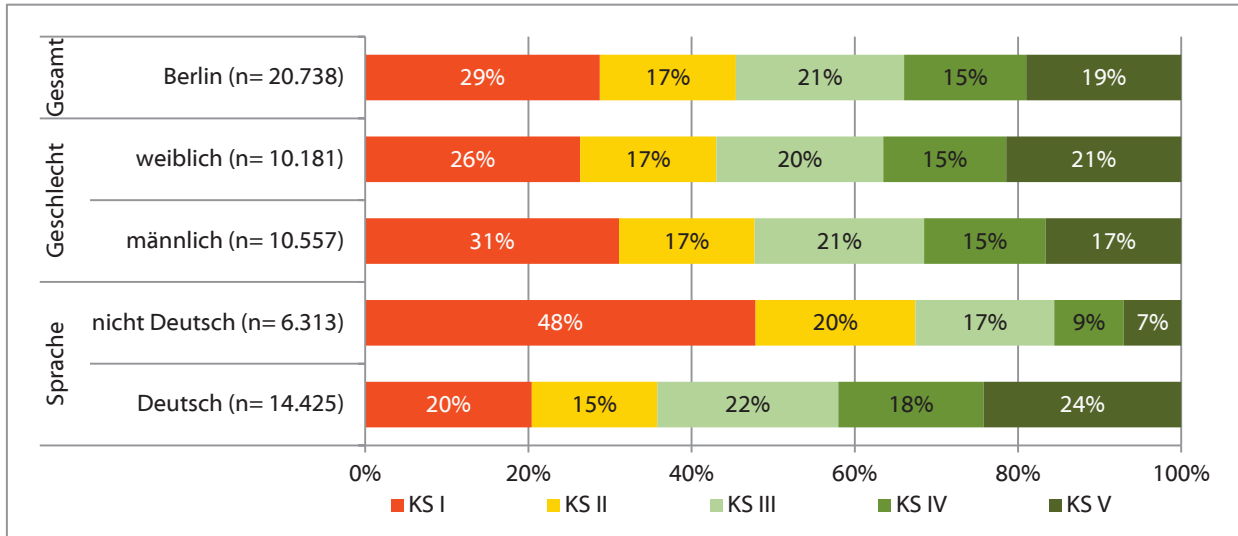


Abbildung 7: Verteilung der Kompetenzstufen im Bereich *Leseverstehen* (alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

Aus Abbildung 8 lassen sich die bezirksspezifischen Kompetenzstufenverteilungen entnehmen. Hierbei verdeutlicht sich der mit Blick auf die verschiedenen BISTA-Punktzahlen bereits gewonnene Eindruck, nämlich dass erhebliche Unterschiede in der Verteilung der Kinder auf die fünf verschiedenen Kompetenzstufen zwischen den einzelnen Bezirken deutlich werden. Demnach war in den Bezirken Neukölln und Mitte, wo durchschnittlich geringere BISTA-Punktzahlen erreicht wurden, ein relativ großer Anteil der Schülerinnen und Schüler auf den Kompetenzstufen I und II zu verorten. Die Bezirke Steglitz-Zehlendorf und Pankow konnten hingegen mit durchschnittlich höheren BISTA-Punktzahlen einen größeren Anteil der Kinder auf den höchsten Kompetenzstufenniveaus vorweisen.

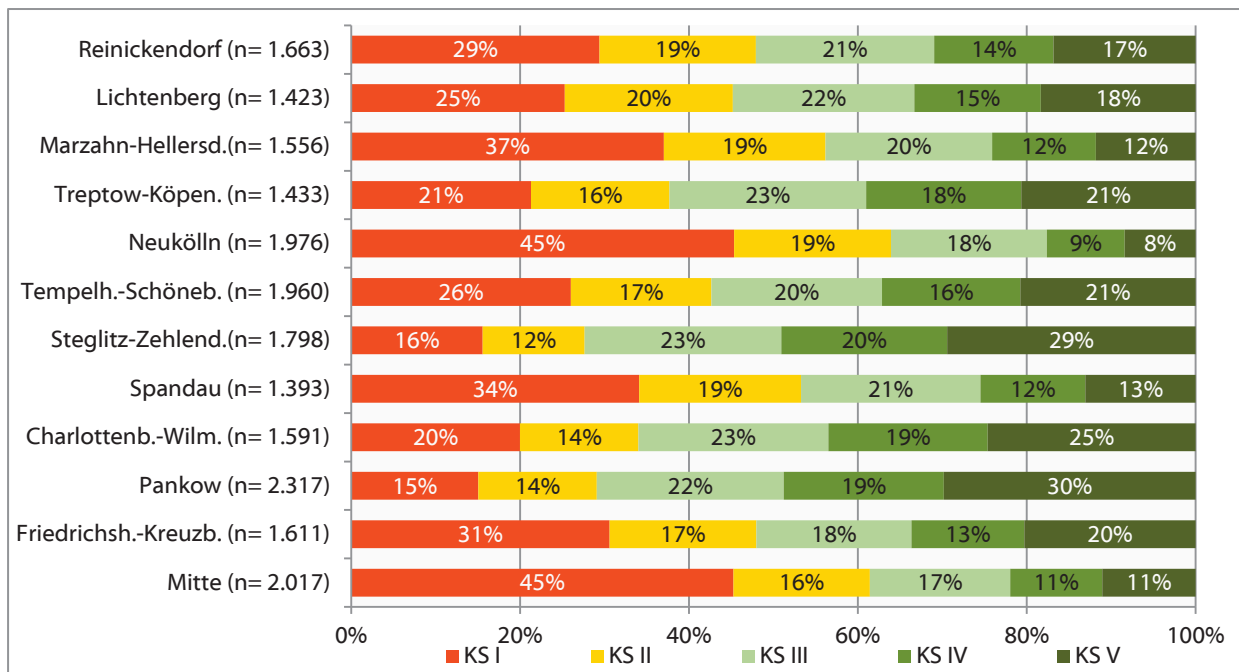


Abbildung 8: Bezirksspezifische Verteilung der Kompetenzstufen im Bereich *Leseverstehen*

3.2 Deutsch – Sprachgebrauch

Der für das Land Berlin obligatorische Test zum Inhaltsbereich Deutsch *Sprachgebrauch* fand am 10.05.2012 statt und bestand aus drei Aufgabenbereichen: Bei der ersten Aufgabe mussten sechs Fragen auf Wortschatz- und Grammatikebene zu einem kurzen lyrischen Text beantwortet werden. Die zweite Aufgabe umfasste drei Items zu einem diskontinuierlichen Sachtext in Form eines Kochrezepts, wobei der Aufgabeninhalt Grammatikkenntnisse der Schülerinnen und Schüler abfragte. Der dritte Aufgabenbereich enthielt acht Fragen zu einem kurzen kontinuierlichen Sachtext. Insgesamt betrug die Bearbeitungszeit des Tests 30 Minuten.

Lösungshäufigkeiten im Bereich *Sprachgebrauch*

Von den 19 Testitems wurden landesweit durchschnittlich 60 % gelöst, wobei Mädchen (62 %) insgesamt etwas besser abschnitten als Jungen (58 %).

Tabelle 5: Itemgenaue Lösungshäufigkeiten im Bereich *Sprachgebrauch*
(alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

	Aufgabe	Berlin (n= 20.754)	Jungen (n= 10.546)	Mädchen (n= 10.208)	Deutsch (n= 14.450)	ndH (n= 6.304)
1.1	Ende des Gedichts	57%	54%	59%	60%	50%
1.2	unterstreiche im Text	59%	57%	62%	66%	45%
1.3	Nomen trennen	76%	73%	78%	79%	68%
1.4	ähnliche unterstreichen	64%	64%	64%	70%	51%
1.5	ähnliche Bedeutung: graben	57%	56%	58%	62%	44%
1.6	Verb & Nomen	54%	53%	55%	60%	40%
2.1	Schüssel	45%	42%	49%	50%	36%
2.2	unterstreichen	53%	51%	55%	60%	38%
2.3	Verben einsetzen	56%	53%	58%	63%	39%
3.1	Verben einsetzen	19%	18%	20%	24%	7%
3.2	Lastenträger	33%	32%	35%	37%	26%
3.3	Adjektive einkreisen	61%	59%	63%	65%	50%
3.4	Gegenteile finden	79%	78%	80%	83%	70%
3.5.1	Steigerung: schwer	89%	89%	89%	92%	83%
3.5.2	Steigerung: wenig	90%	89%	90%	92%	85%
3.5.3	Steigerung: hart	83%	82%	84%	86%	75%
3.6	Oberbegriffe	60%	56%	64%	64%	51%
3.7	Plural ergänzen	71%	69%	73%	79%	54%
3.8	Eselohren	31%	28%	33%	37%	16%
	Gesamt	60%	58%	62%	65%	49%

Kinder mit Herkunftssprache Deutsch („Deutsch“) lösten durchschnittlich 65 % der Aufgaben im Bereich *Sprachgebrauch* korrekt, was deutlich über der durchschnittlichen Lösungshäufigkeit der Kinder nichtdeutscher Herkunftssprache (49 %) liegt. Der Anteil von Kindern nichtdeutscher Herkunftssprache machte in diesem Testbereich etwa 30 % aus.

Der Abbildung ist die grafisch dargestellte **Verteilung der erreichten Gesamtpunktzahlen** (von 19 möglichen Punkten) im Bereich *Sprachgebrauch* aller teilnehmenden Berliner Schülerinnen und Schüler zu entnehmen. Etwa ein Viertel der Schülerinnen und Schüler erzielten demnach eine Gesamtpunktzahl von mindestens 14 Punkten. Hingegen erreichte ungefähr ein Drittel aller Kinder maximal die Hälfte der vollen Punktzahl. Durchschnittlich wurden 11,4 der 19 Aufgaben korrekt gelöst, was einer mittleren Lösungshäufigkeit von 60 % entspricht.

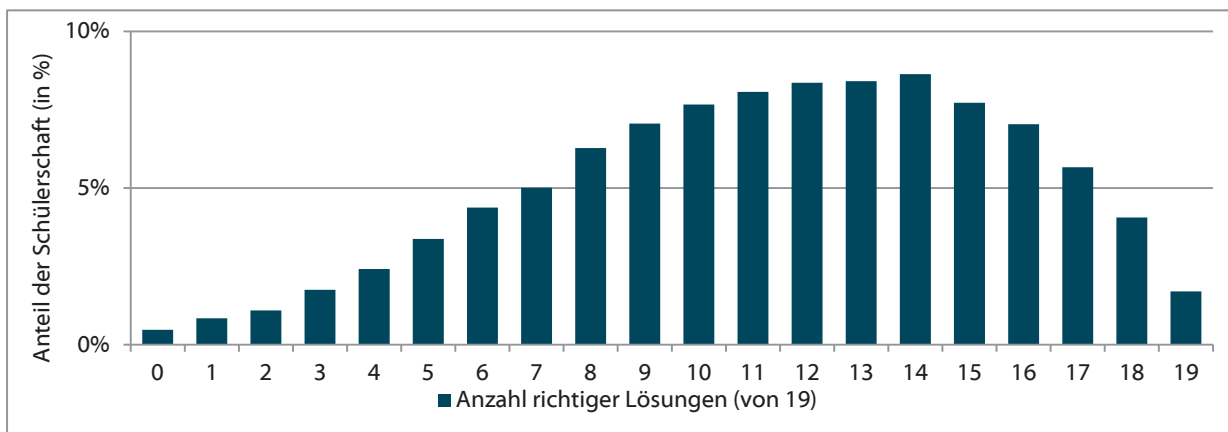


Abbildung 9: Verteilung der Testwerte in Punktschritten in Deutsch – *Sprachgebrauch*

Detaillierte Analyse einzelner Aufgaben aus dem Bereich *Sprachgebrauch*

Blick auf die leichteste Aufgabe - Aufgabe 3.5

3.5 Ergänze die Tabelle.

groß	<i>größer</i>	<i>am größten</i>
schwer		
wenig		
hart		

Aufgabe 3.5 erwies sich für die Berliner Schülerinnen und Schüler mit einer Lösungshäufigkeit von 82 bzw. 89 % als leichteste Aufgabe. Es mussten Adjektive nach einem vorgegebenen Muster gesteigert werden. Die drei Teilaufgaben wurden einzeln/zeilenweise ausgewertet. Bei der dritten Teilaufgabe musste die Veränderung des Stammvokals erkannt werden, was

für einige Kinder eine höhere Schwierigkeit darstellte, obwohl die Rechtschreibung keine Rolle spielte.

Blick auf die schwierigste Aufgabe - Aufgabe 3.1


Die einzige auf Kompetenzstufe V angesiedelte und damit schwierigste Aufgabe zielte auf die Bildung der Präteritumsform von im Infinitiv vorgegebenen Verben. Die Verben waren in Lückensätze unter Beachtung von Person und Numerus einzufügen. Erschwerend wirkte hier, dass die Aufgabe komplex ausgewertet wurde, d. h. dass sie nur als korrekt erfüllt gewertet werden konnte, wenn alle drei Verben in der richtigen Form gebildet waren. Diese Art der Auswertung war auch in den letzten Jahren auffällig, da in der Grundschule üblicherweise für jede erfüllte Teilaufgabe Punkte vergeben werden.

3.1 Im folgenden Text fehlt etwas. Setze die Verben in der Vergangenheit ein.

Beispiel:
Früher lebten (leben) vermutlich mehr Esel auf der Welt. In einer Welt ohne Autos _____ (ziehen) die Esel Karren und _____ (tragen) schwere Lasten. Auf fast jedem Bauernhof _____ (geben) es Esel.

Blick auf Aufgabenformate

3.8 Tim sagt: „In deinem Buch sind ja Eselsohren!“
 Maria antwortet: „Das kann nicht sein, denn es ist ein Buch über Fische.“
 Was ist mit dem Wort **Eselsohren** gemeint?



Im Bereich *Sprachgebrauch* kamen verschiedene Aufgabenformate zum Einsatz, die aber auch in den im Unterricht gängigen Lehrwerken wiederzufinden sind und somit zu den klassischen Unterrichtsaufgaben gehören. Aufgaben 1.2, 2.2 und 3.8 wurden häufig nicht bearbeitet. Hier handelt es sich um zwei Aufgaben zum Unter-

streichen von zusammengesetzten Nomen und eine offene Aufgabe, die auf das Erklären einer Wortbedeutung zielt.

Blick auf Unterschiede zwischen Jungen und Mädchen

In Abbildung 10 sind die Lösungshäufigkeiten pro Item im Test Deutsch, Inhaltsbereich *Sprachgebrauch*, nach Geschlecht geordnet dargestellt. Sortiert wurden die Aufgaben hierbei nach der Schwierigkeit der Items für die Mädchen. Auch im Bereich *Sprachgebrauch* lösten Mädchen die Aufgaben besser als Jungen, wenn auch die Unterschiede zum Teil sehr gering ausfallen. Lediglich bei Aufgabe 3.6 besteht ein größerer Unterschied. Die Aufgabe verlangte das Zuordnen von Oberbegriffen. Es waren mehr Oberbegriffe „im Angebot“ als zur Aufgabenlösung nötig, was erschwerend wirkte.

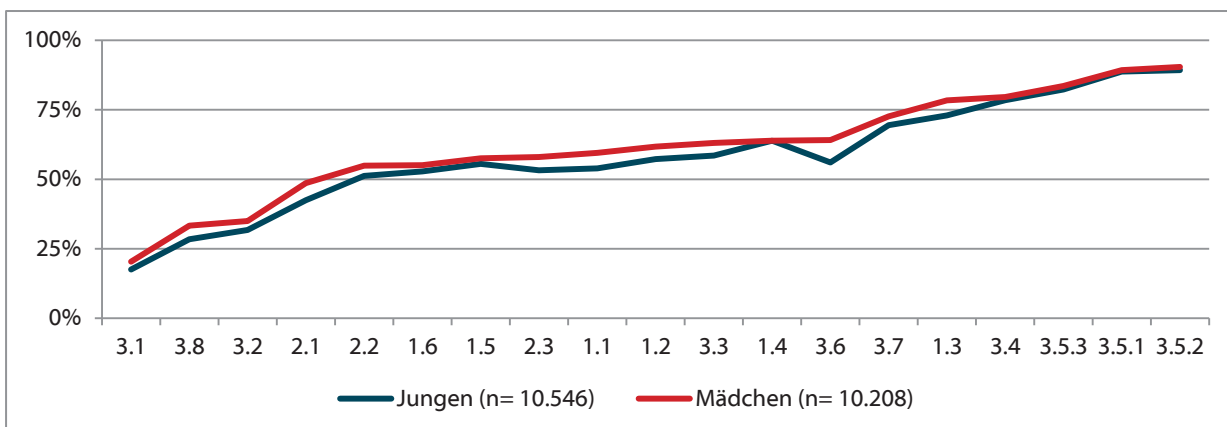


Abbildung 10: Itemspezifische Lösungshäufigkeiten in Deutsch – *Sprachgebrauch*, nach Geschlecht

Blick auf die Ergebnisse von Kindern deutscher und nichtdeutscher Herkunftssprache

In Abbildung werden die Unterschiede zwischen Kindern deutscher und nichtdeutscher Herkunftssprache veranschaulicht. Diese fallen erheblich größer aus als bei der geschlechterspezifischen Darstellung, wobei die Differenzen in den Lösungshäufigkeiten hierbei nicht selten 20 Prozent und mehr betragen. Die Spanne der Lösungshäufigkeiten bei Kindern mit der Muttersprache Deutsch reicht von 92 % bis 24 %. Bei Kindern nichtdeutscher Herkunftssprache erstreckt sie sich zwischen 85 % und 7 %. Die geringste Lösungshäufigkeit zeigt sich bei Kindern deutscher sowie nichtdeutscher Herkunftssprache bei Aufgabe 3.1 (siehe auch Blick auf die schwerste Aufgabe). Drei Aufgaben fallen auf, die insbesondere von Kindern nichtdeutscher Herkunftssprache häufig nicht bearbeitet wurden: Aufgabe 1.2 wurde von 34 % dieser Kinder nicht bearbeitet. Sie verlangte das Unterstreichen ungewöhnlicher zusammengesetzter Nomen in einem lyrischen Text. Aufgabe 2.2, von 31 % nicht bearbeitet, zielte ebenfalls auf das Erkennen von zusammengesetzten Nomen in einem Text.

Aufgabe 3.8 (Kompetenzstufe IV) wurde von 34 % der Kinder nichtdeutscher Herkunftssprache nicht bearbeitet. Sie war die letzte Aufgabe im Testheft und zielt auf das Klären eines sprachlichen Missverständnisses ab. Die genannten drei Aufgaben fielen auch Kindern deutscher Herkunftssprache schwer.

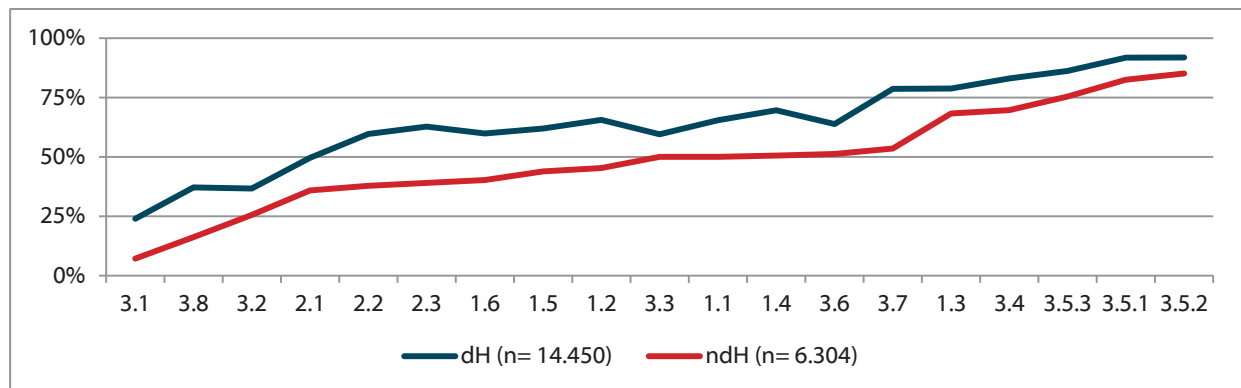


Abbildung 11: Itemspezifische Lösungshäufigkeiten in Deutsch – *Sprachgebrauch*, nach Herkunftssprache (Items geordnet nach Schwierigkeit der Aufgaben für die Kinder mit ndH)

Im Hinblick auf **Nichtbearbeitung** der einzelnen Aufgaben wurden die Items am häufigsten nicht bearbeitet, die den tendenziell höheren Kompetenzstufen zugeordnet wurden. Hinweise dafür, dass der Test zu lang bzw. die Bearbeitungszeit zu kurz war, lassen sich aus der Abbildung 12 nicht erkennen.

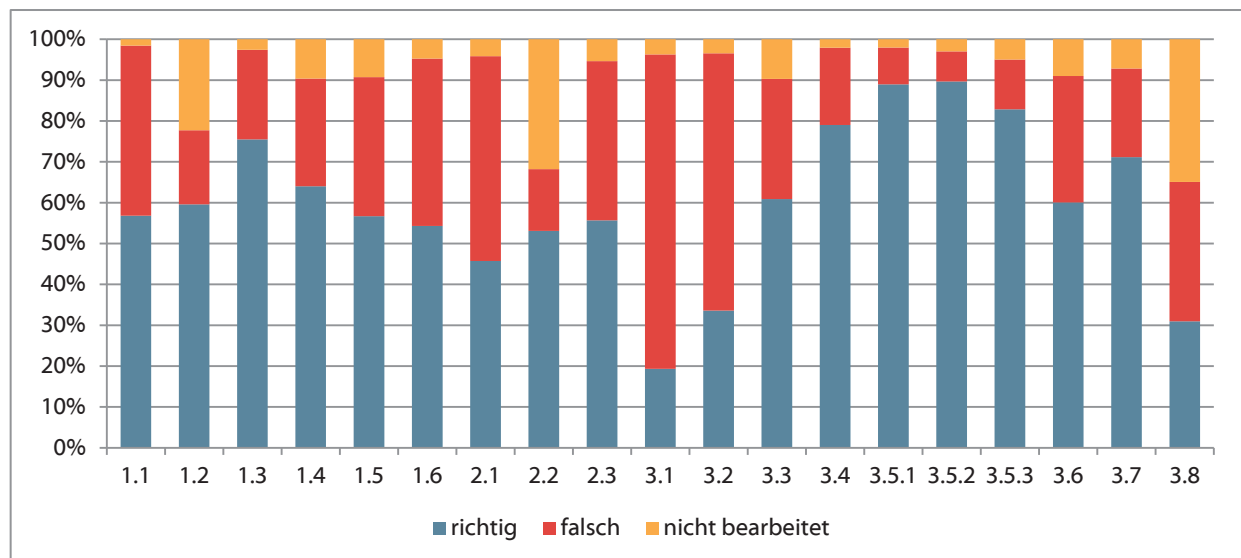


Abbildung 12: Bearbeitung der Aufgaben in Testreihenfolge im Testheft *Sprachgebrauch*

Auswertung nach Punkten der KMK-Bildungsstandards (BISTA-Punkte) im Bereich *Sprachgebrauch*

In Tabelle 6 sind die Ergebnisse im Bereich *Sprachgebrauch* auf Basis der BISTA-Punkte (vgl. Kapitel 3) differenziert nach Bezirk, Geschlecht und Herkunftssprache abzulesen. Hierbei lassen sich bei den verschiedenen Betrachtungsebenen erhebliche Unterschiede erkennen. Das Spektrum der durchschnittlichen BISTA-Punktzahlen reichte in diesem Inhaltsbereich von 431 Punkten im Bezirk Mitte bis hin zu 517 BISTA-Punkten in Steglitz-Zehlendorf.

Tabelle 6: Durchschnittliche BISTA-Punkte im Bereich *Sprachgebrauch*
(alle Schüler/-innen, nach Gruppen differenziert, Bezirksergebnisse)

Nr.	Bezirk	Berlin		Geschlecht				Herkunftssprache			
		Gesamt (n= 20.754)		männlich (n= 10.546)		weiblich (n= 10.208)		Deutsch (n= 14.450)		nicht Deutsch (n= 6.304)	
		n	M	n	M	n	M	n	M	n	M
1	Mitte	2.015	431	1.014	424	1.001	439	715	480	1.300	405
2	Friedrichshain-Kreuzberg	1.612	472	826	466	786	477	890	519	722	413
3	Pankow	2.309	516	1.223	503	1.086	530	2.167	520	142	448
4	Charlottenburg-Wilmersdorf	1.599	493	828	481	771	506	1.019	516	580	452
5	Spandau	1.399	459	699	453	700	466	894	480	505	422
6	Steglitz-Zehlendorf	1.799	517	920	501	879	534	1.516	530	283	448
7	Tempelhof-Schöneberg	1.966	490	978	482	988	498	1.271	519	695	436
8	Neukölln	1.966	434	991	428	975	441	932	467	1.034	405
9	Treptow-Köpenick	1.432	495	697	483	735	506	1.350	498	82	437
10	Marzahn-Hellersdorf	1.554	457	794	447	760	467	1.308	464	246	418
11	Lichtenberg	1.434	483	729	474	705	492	1.156	491	278	448
12	Reinickendorf	1.669	461	847	454	822	468	1.232	483	437	401
	Gesamt	M= 476	SD= 116	M= 467	SD= 113	M= 485	SD= 118	M= 500	SD= 111	M= 421	SD= 108

Die Ergebnisse nach geschlechtsspezifischer Betrachtung weisen zwischen den verschiedenen Bezirken deutliche Unterschiede auf. Insgesamt schnitten die Mädchen besser ab als die Jungen. Die BISTA-Punkte variierten bezirkweise zwischen beiden Geschlechtern. Friedrichshain-Kreuzberg lässt mit 11 Punkten Unterschied die geringste Differenz erkennen, wohingegen in Steglitz-Zehlendorf mit 33 Punkten die größten Unterschiede auftraten.

Zieht man nun die Herkunftssprache als differenzierendes Kriterium der Leistungen im Bereich *Sprachgebrauch* auf Basis der BISTA-Punktzahlen heran, sind noch weitaus größere Differenzen auszumachen. Kinder mit Deutsch als Muttersprache schnitten demnach in ausnahmslos allen Bezirken Berlins besser ab als Schülerinnen und Schüler nichtdeutscher Muttersprache. Am deutlichsten waren diese Differenzen in Friedrichshain-Kreuzberg mit 106 BISTA-Punkten Unterschied, wohingegen die geringste Spanne im Bezirk Lichtenberg mit 43 Punkten auftrat.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen im Bereich *Sprachgebrauch*

Neben der Darstellung von Lösungshäufigkeiten bzw. BISTA-Werten pro Aufgabe können Schülerleistungen einer bestimmten Kompetenzstufe (KS) zugeordnet werden. Unterschieden wird dabei zwischen Unter Mindeststandard (KS I), Mindeststandard (KS II), Regelstandard (KS III), Regelstandard Plus (KS IV) und Maximalstandard (KS V).

In Abbildung 13 sind die Kompetenzstufenverteilungen der Berliner Schülerinnen und Schüler im Bereich Deutsch *Sprachgebrauch* dargestellt. Landesweit erreichten über 80 % aller Kinder mindestens Kompetenzstufe II und erfüllten somit die Mindeststandards für das Ende der Klassenstufe 4. Auch hier gab es bei der geschlechtsspezifischen Betrachtung geringe Differenzen zwischen Jungen und Mädchen. Lediglich auf Kompetenzstufe I und V waren Unterschiede von 3 bzw. 5 Prozentpunkten auszumachen, wobei die Mädchen mit 14 % auf Kompetenzstufe V und 18 % auf KS I etwas besser einzuordnen waren als die Jungen mit einem Anteil von 9 % auf KS V und 21 % auf der niedrigsten Kompetenzstufe.

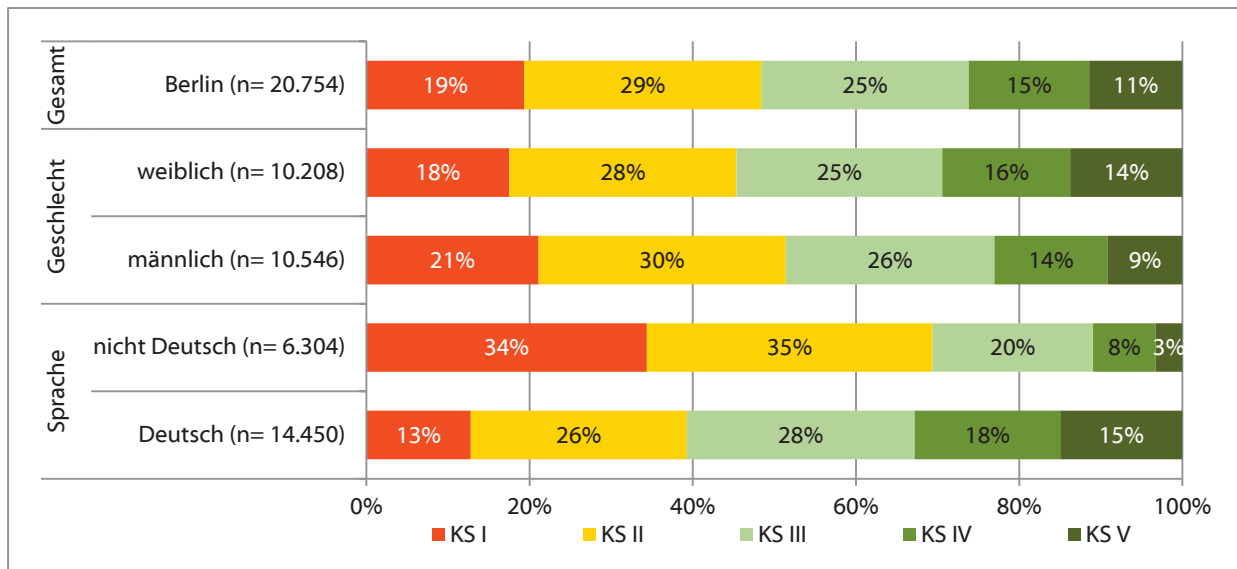


Abbildung 13: Verteilung der Kompetenzstufen im Bereich *Sprachgebrauch* (alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

Bei der Betrachtung der Kompetenzstufenverteilungen nach Herkunftssprache sticht klar der große Anteil von Kindern mit nichtdeutscher Herkunftssprache auf Kompetenzstufe I hervor. So erfüllten mit 34 % über ein Drittel dieser Kinder nicht die Mindeststandards in Deutsch *Sprachgebrauch* für das Ende der Klassenstufe 4. Möglich wäre z. B., dass dieser Unterschied auf geringeres Lesetempo zurückzuführen ist oder es erhebliche Schwierigkeiten beim Verständnis und der Bearbeitung der Aufgaben gab.

Aus Abbildung 14 gehen die bezirksspezifischen Kompetenzstufenverteilungen hervor. Hierbei verdeutlichen sich die Tendenzen der BISTA-Punktetabelle. So finden sich in Bezirken mit geringeren durchschnittlichen BISTA-Punktzahlen auch größere Anteile von Kindern auf Kompetenzstufe I und II wieder. In Bezirken mit relativ hohen BISTA-Werten gibt es dagegen größere Anteile von Schülerinnen und Schülern, die die Kompetenzstufen IV und V erreichten. In den Bezirken Neukölln (29 %) und Mitte (34 %) finden sich demnach die höchsten Anteile mit Kindern auf Kompetenzstufe I, wohingegen in den Bezirken Steglitz-Zehlendorf (18 %) und Pankow (17 %) die höchsten Anteile der Schülerinnen und Schüler auf Kompetenzstufe V auszumachen sind.

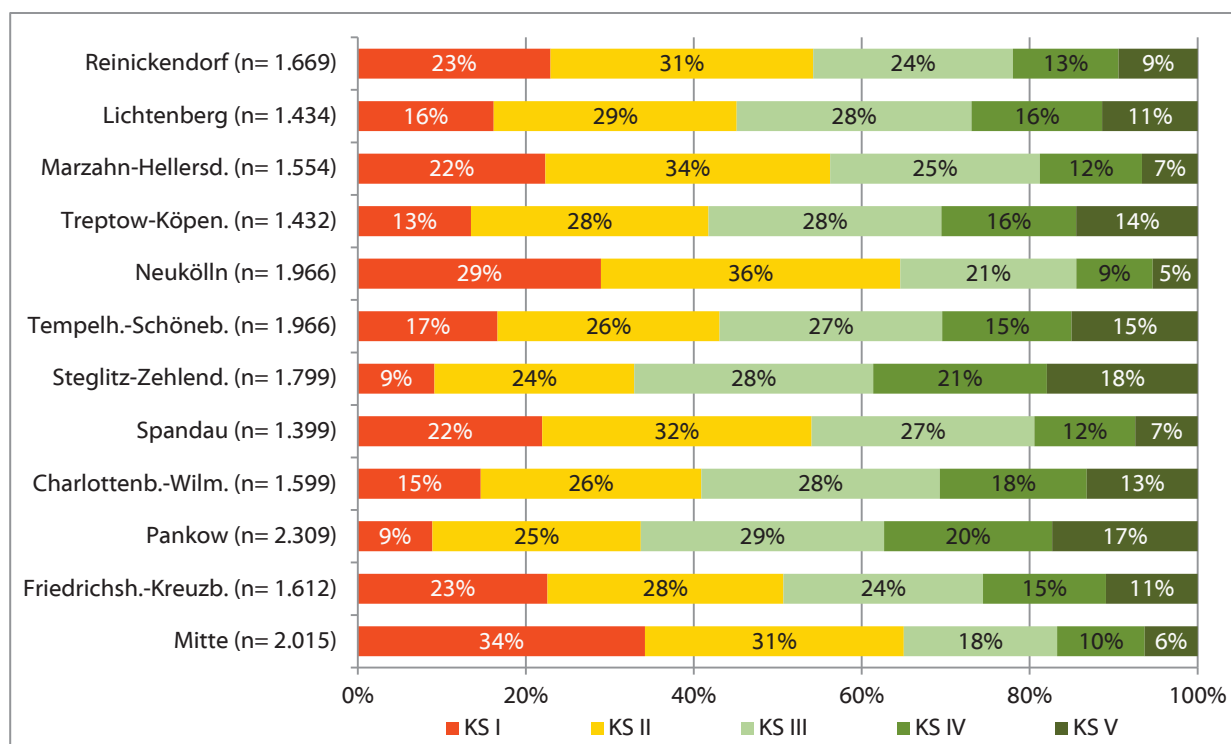


Abbildung 14: Bezirksspezifische Verteilung der Kompetenzstufen im Bereich *Sprachgebrauch*

3.3 Mathematik

Analog zu den Ergebnisdarstellungen für das Fach Deutsch folgen die Ergebnisse für Mathematik auf Basis der Kompetenzstufen, wobei die Fähigkeiten wiederum unterschieden werden nach Unter Mindeststandard (KS I), Mindeststandard (KS II), Regelstandard (KS III), Regelstandard Plus (KS IV) und Maximalstandard (KS V). Weiterführende detaillierte Beschreibungen der Kompetenzstufen für die beiden spezifischen getesteten Inhaltsbereiche (*Muster und Strukturen* sowie *Größen und Messen*) finden sich in den didaktischen Handreichungen (<http://www.isq-bb.de/portal/>).

Der Test zu den beiden Mathematik-Inhaltsbereichen fand am 15.05.2012 statt.

3.3.1 Mathematik – *Muster und Strukturen*

Lösungshäufigkeiten im Bereich *Muster und Strukturen*

Tabelle 7: Itemgenaue Lösungshäufigkeiten im Bereich *Muster und Strukturen*
(alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

	Aufgabe	Berlin (n= 20.724)	Jungen (n= 10.592)	Mädchen (n= 10.132)	Deutsch (n= 14.349)	ndH (n= 6.375)
1a	Welche Aufgabe I	88%	89%	88%	91%	83%
1b	Welche Aufgabe II	68%	68%	69%	71%	62%
2a	Gleichung finden I	81%	79%	83%	82%	78%
2b	Gleichung finden II	86%	85%	88%	87%	84%
3	Fußgänger	52%	58%	46%	57%	42%
4	Finde die Regel	74%	74%	74%	78%	66%
5	Preise	57%	63%	50%	60%	49%
6	Gummibärchen	32%	39%	25%	35%	24%
7	Stellenwert/Punkte	77%	75%	78%	80%	69%
8	Musterfolge	78%	79%	76%	81%	71%
9	Rechentürme	44%	44%	44%	47%	36%
10	Käfer	49%	55%	42%	53%	39%
11a	Preiszuordnung I	44%	49%	39%	47%	36%
11b	Preiszuordnung II	34%	40%	26%	37%	27%
11c	Preiszuordnung III	37%	44%	30%	41%	29%
11d	Preiszuordnung IV	26%	32%	21%	29%	20%
12	Bleistifte	20%	23%	17%	23%	13%
13	Käse	32%	38%	25%	35%	23%
14	Quadratmuster	65%	65%	64%	68%	56%
15	Zahlenreihe	50%	55%	44%	55%	38%
16	Katzenfutter	50%	52%	48%	58%	33%
17a	Zahlenpaare I	60%	64%	55%	64%	51%
17b	Zahlenpaare II	66%	70%	61%	69%	57%
	Gesamt	55%	58%	52%	59%	47%

Aus Tabelle 7 kann man die durchschnittlichen Lösungshäufigkeiten für Berlin im Inhaltsbereich *Muster und Strukturen* entnehmen. Demnach wurden in Berlin im Durchschnitt 55 % des gesamten Inhaltsbereichs korrekt gelöst. Jungen schnitten mit einer durchschnittlichen Lösungshäufigkeit von 58 % insgesamt etwas besser ab als Mädchen mit 52 %. Ein deutlich größerer Unterschied ist auch hier zwischen Schülerinnen und Schülern mit Deutsch als Muttersprache (59 %) und Kindern nichtdeutscher Herkunftssprache (47 %) mit einer Differenz von rund zwölf Prozent auszumachen.

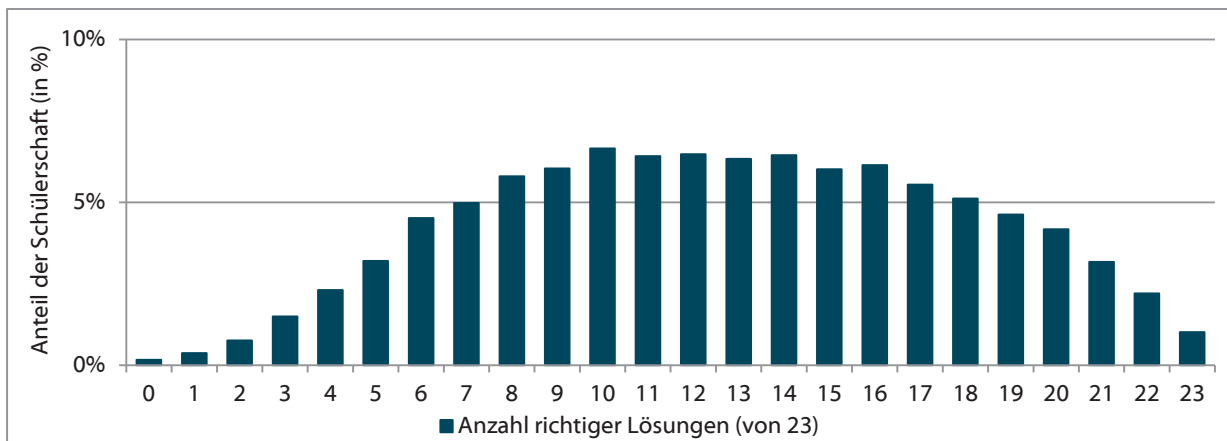


Abbildung 15: Verteilung der Testwerte in Punktschritten in Mathematik – *Muster und Strukturen*

In Abbildung 15 ist die prozentuale Verteilung der erreichten Punktschritte dargestellt. Nur etwa ein Prozent aller Schülerinnen und Schüler erreichten demnach die Höchstpunktzahl von 23 Punkten. Knapp die Hälfte aller Kinder erreichte eine Gesamtpunktzahl von bis zu 12 Punkten, was etwas mehr als 50 Prozent des gesamten Testbereichs *Muster und Strukturen* ausmacht. Rund acht Prozent hingegen erzielten lediglich eine Punktzahl von maximal fünf Punkten und lösten somit weniger als ein Viertel des Testbereichs. Im Schnitt lösten die Schülerinnen und Schüler 12,6 der 23 Aufgaben, was einer mittleren Lösungshäufigkeit von 55 % entspricht.

Detaillierte Analyse einzelner Aufgaben aus dem Bereich *Muster und Strukturen*

Blick auf die leichteste Aufgabe – Aufgabe 1a

Aufgabe 1

Wie heißt die nächste Aufgabe?

a) $36 + 41 = 77$
 $46 + 31 = 77$
 $56 + 21 = 77$

=

Am häufigsten richtig gelöst wurde mit 88 % die Aufgabe 1a, bei der die Gesetzmäßigkeit einer strukturierten Aufgabenfolge erkannt und um eine Aufgabe fortgesetzt werden musste. Der erste Summand vergrößert sich von Aufgabe zu Aufgabe jeweils um 10 und gleichzeitig verkleinert sich der zweite Summand von Aufgabe zu Aufgabe ebenfalls jeweils um 10, sodass die Summen der Aufgaben unverändert 77 bleiben.

Es handelt sich hierbei um ein Aufgabenformat, das häufig unter dem Namen „Entdeckerpäckchen“ oder „schönes Päckchen“ bekannt ist.

Diese strukturierten Aufgabenfolgen sind inzwischen im Unterricht und in den Lehrwerken etabliert und den Schülern bekannt. Das Finden und Aufschreiben der fehlenden Zahlen stellt keine Schwierigkeit dar, da der Unterschied zwischen den Summanden keinen großen Anspruch an die Rechenfertigkeit erhebt und musterbedingt das Ergebnis immer gleich bleibt. Aus diesen Gründen wurde die Aufgabe der Kompetenzstufe I zugeordnet.

Interessant ist an dieser Stelle Aufgabe 1b, die von 68 % der Kinder richtig gelöst wurde. Es handelte sich um das gleiche Aufgabenformat einer strukturierten Aufgabenfolge, bei der allerdings aufgrund der sich verändernden Ergebnisse mehr gerechnet werden musste. Aufgrund dessen wurde Aufgabe 1b der Kompetenzstufe II zugeordnet. In dieser Aufgabe ist der erste Summand jeder Aufgabe ein Produkt gleicher Faktoren. Von Aufgabe zu Aufgabe vergrößern sich die Faktoren jeweils um 1 und auch der zweite Summand vergrößert sich um jeweils 1. Hierbei waren die Zahlen so gewählt, dass die Struktur schnell erkannt werden konnte. Trotzdem lösten nur 68 % der Kinder diese Aufgabe richtig. Die Ursache könnte in einer fehlerhaften Errechnung der Summen durch die Schülerinnen und Schüler liegen. Möglich ist aber auch, dass sie, ohne zu rechnen, versuchten ein Muster in der Abfolge der Zahlen im Ergebnis zu erkennen und dabei die Richtigkeit der Aufgabe als Gleichung nicht mehr überprüften. Als Beispiel könnte $6 \cdot 6 + 4 = 38$ (als scheinbar fortlaufendes Muster im Ergebnis, statt 40) genannt werden. Woran es im Einzelfall liegt, kann nur von den betreuenden Lehrerinnen und Lehrern beurteilt werden.


Blick auf die schwierigste Aufgabe - Aufgabe 12

Aufgabe 12

Ein Bleistift kostet im Laden 1 Euro.
Für ein 10er-Päckchen muss man 9 Euro bezahlen.

Trage jeweils den günstigsten Preis ein.

Anzahl der Stifte	8	12	20
Preis	€	€	€



Grafik: © IQB

Die Aufgabe 12 wurde mit 20 % am seltensten richtig gelöst. Die Aufgabe bestand aus drei Teilaufgaben, die alle korrekt gelöst werden mussten, damit die Aufgabe als richtig in die Bewertung einging. Sie ist der Kompetenzstufe IV und dem Anforderungsbereich III zugeordnet. Auf den ersten Blick erscheint die Aufgabe gar nicht so anspruchsvoll. Die Schwierigkeit liegt bei dieser Aufgabe darin, den Text sorgfältig und korrekt zu lesen und zu erkennen, dass

bei 12 und 20 Stiften das Sonderangebot berücksichtigt werden muss, um den geforderten „günstigsten“ Preis zu ermitteln.

Nur 13 % der Kinder nichtdeutscher Herkunftssprache konnten diese Aufgabe richtig lösen, was ein weiteres Indiz dafür sein könnte, dass hier eine hohe Lese- und Sprachkompetenz gefordert ist.

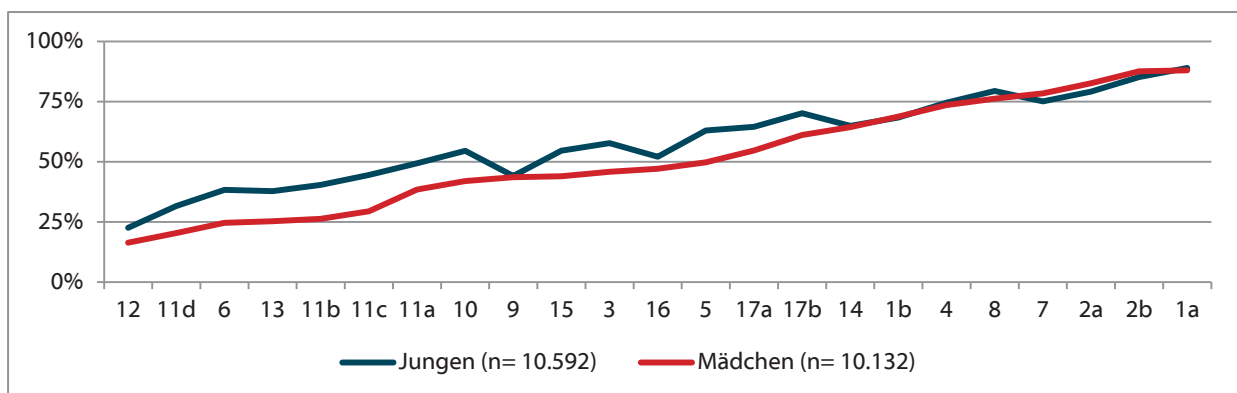


Abbildung 16: Itemspezifische Lösungshäufigkeiten in Mathematik – *Muster und Strukturen*, nach Geschlecht (Items geordnet nach der Schwierigkeit der Aufgaben für die Mädchen)

Blick auf Unterschiede zwischen Jungen und Mädchen

In Abbildung 16 sind die Lösungshäufigkeiten im Bereich *Muster und Strukturen* nach Geschlecht abgebildet, aufsteigend sortiert nach den meistgelösten Items der Mädchen. Die Grafik zeigt, dass sich mit Zunahme der Itemschwierigkeit eine erkennbare Differenz zwischen den Lösungshäufigkeiten

beider Geschlechter herausbildet. Bei den leichteren Aufgaben sind die Lösungshäufigkeiten noch fast auf gleicher Ebene, wobei Mädchen bei Aufgabe 2a und 2b sogar zwei bis drei Prozentpunkte besser abschnitten als Jungen. Bis auf Aufgabe 9, bei welcher eine Folge von Additionsgleichungen vervollständigt werden sollte, zeigen sich bei den mittelschwierigen bis schwierigen Items jedoch zum Teil Differenzen von 10 bis 15 Prozentpunkten. Besonders hoch waren die Unterschiede beispielsweise bei den Aufgaben 11a bis 11 d, bei welchen Divisions- und Multiplikationskenntnisse der Schüler/-innen in Form von kleinen Textaufgaben abgefragt wurden. Diese Aufgaben wurden von Jungen weitaus häufiger korrekt gelöst als von Mädchen.

Dass Mädchen und Jungen in der Aufgabe 9 (Rechentürme), die der Kompetenzstufe III und dem Anforderungsbereich II zugeordnet wurde, gleich abschnitten, zeigt, dass sich das Aufgabenformat der strukturierten Aufgabenreihe („Entdeckerpäckchen“) zum Auffinden von Mustern und Strukturen für Jungen und Mädchen gleichermaßen etabliert hat.

Blick auf die Ergebnisse von Kindern deutscher und nichtdeutscher Herkunftssprache

Zieht man nun die Lösungshäufigkeiten nach Herkunftssprache zur genaueren Analyse heran, lässt sich eine noch eindeutigerere Differenz als bei der geschlechtsspezifischen Betrachtung beobachten (vgl. Abbildung). Zwischen Schülerinnen und Schülern deutscher und nichtdeutscher Herkunftssprache lag die durchschnittliche Leistungsdifferenz bei etwa 12 % zugunsten der Schülerinnen und Schüler deutscher Herkunftssprache (vgl. auch Tabelle 7).

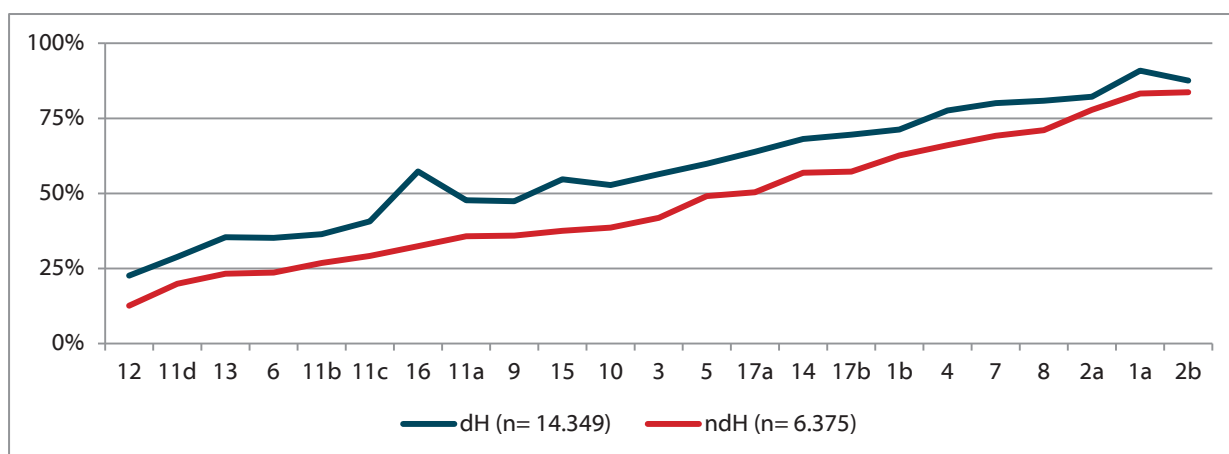


Abbildung 17: Itemspezifische Lösungshäufigkeiten in Mathematik – *Muster und Strukturen*, nach Herkunftssprache (Items geordnet nach der Schwierigkeit der Aufgaben für die Kinder mit ndH)

Nachdem die Unterschiede bei den leichtesten Items (2a, 2b) noch bei etwa vier Prozent lagen, stieg die Differenz schnell auf konstante 8 bis 14 Prozentpunkte an. Besonders ausgeprägte Differenzen zwischen Kindern mit deutscher und nichtdeutscher Herkunftssprache fallen mit 17 % bei der Aufgabe „Zahlenreihe“ (15) und mit 25 % bei der Aufgabe „Katzenfutter“ (16) auf.

Bei Aufgabe 16 war eine einfache Sachaufgabe zur Proportionalität zu lösen. Die Schwierigkeit dieser Aufgabe liegt nicht nur darin, dass ein hohes Maß an Modellierungsfähigkeit gefordert war. Es war auch eine hohe Lesekompetenz nötig, denn die vorliegende Aufgabe beschrieb in einer Sachsituation eine antiproportionale Beziehung in einem Text, der inhaltlich genau verstanden werden musste. Schließlich galt es zu erfassen, dass sich zwar die Anzahl der zu versorgenden Katzen verdoppelt, gleichzeitig aber die Anzahl der Tage halbiert wird.

Auch im Bereich *Muster und Strukturen* sollen die Aufgaben im Hinblick auf deren **Nichtbearbeitung** als Hinweis auf zu umfangreiche Testhefte betrachtet werden.

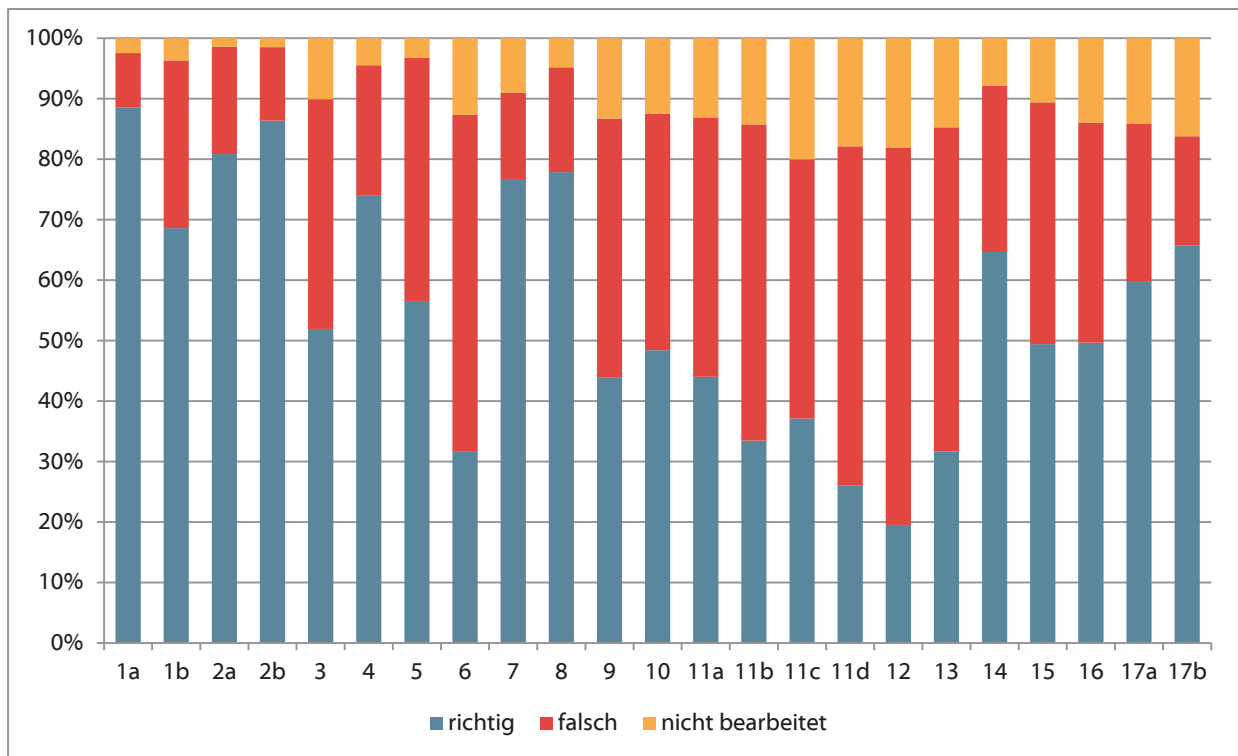


Abbildung 18: Bearbeitung der Aufgaben in Testreihenfolge in Mathematik – *Muster und Strukturen*

In Abbildung wird die Lösung der Aufgaben nach „richtig“, „falsch“ und „nicht bearbeitet“ differenziert dargestellt. Häufig nicht bearbeitet wurden insbesondere diejenigen Aufgaben, die sich durch eine hohe Schwierigkeit auszeichneten. Es ist anzunehmen, dass diese von den Kindern oft gleich übersprungen wurden und somit dann vielleicht auch wegen des Zeitdrucks letztlich zu den Aufgaben wurden, die sich hier durch einen hohen Anteil an Nichtbearbeitungen kennzeichnen lassen.

Auswertung nach Punkten der KMK-Bildungsstandards (BISTA-Punkte) im Bereich *Muster und Strukturen*

Bei den bezirksspezifischen Ergebnissen auf Basis der BISTA-Punkte werden die Differenzen zwischen den Schülerleistungen im Inhaltsbereich *Muster und Strukturen* sichtbar. Mit einer durchschnittlichen BISTA-Punktzahl von 476 Punkten lag Steglitz-Zehlendorf vor allen anderen Berliner Bezirken. Mitte erreichte mit 409 Punkten die geringste BISTA-Punktzahl für diesen Testbereich.

Tabelle 8: Durchschnittliche BISTA-Punkte im Bereich *Muster und Strukturen*
(alle Schüler/-innen, nach Gruppen differenziert, Bezirksergebnisse)

Nr.	Bezirk	Berlin		Geschlecht				Herkunftssprache			
		Gesamt (n= 20.724)		männlich (n= 10.592)		weiblich (n= 10.132)		Deutsch (n= 14.349)		nicht Deutsch (n= 6.375)	
		n	M	n	M	n	M	n	M	n	M
1	Mitte	2.010	409	1.026	420	984	399	711	437	1.299	394
2	Friedrichsh.-Kreuzb.	1.581	430	827	445	754	413	872	458	709	394
3	Pankow	2.285	468	1.218	481	1.067	453	2.144	470	141	435
4	Charlottenb.-Wilmersd.	1.592	461	821	476	771	445	1.014	475	578	438
5	Spandau	1.400	429	713	447	687	409	901	440	499	409
6	Steglitz-Zehlendorf	1.847	476	936	487	911	465	1.507	483	340	447
7	Tempelhof-Schöneb.	1.983	458	996	473	987	441	1.270	474	713	429
8	Neukölln	1.933	411	982	422	951	399	913	429	1.020	394
9	Treptow-Köpenick	1.431	452	696	464	735	441	1.349	455	82	411
10	Marzahn-Hellersdorf	1.540	432	794	448	746	414	1.293	436	247	409
11	Lichtenberg	1.443	448	729	463	714	432	1.159	448	284	448
12	Reinickendorf	1.679	439	854	454	825	423	1.216	454	463	400
	Gesamt	M= 443	SD= 94	M= 457	SD= 99	M= 428	SD= 87	M= 457	SD= 92	M= 410	SD= 88

Bei der geschlechtsspezifischen Betrachtung wird in allen Bezirken ein Leistungsvorsprung der Jungen gegenüber den Mädchen deutlich. Am meisten bestätigt sich diese Beobachtung in Spandau, wo die Jungen mit 38 BISTA-Punkten Differenz besser abschnitten. Der geringste Unterschied konnte im Bezirk Mitte mit 21 BISTA-Punkten ermittelt werden.

Noch größer fielen die Unterschiede jedoch bei den Ergebnissen nach Herkunftssprache differenziert aus. In fast allen Bezirken - den Bezirk Lichtenberg, wo Kinder mit nichtdeutscher Herkunftssprache die gleiche durchschnittliche Punktzahl erreichten, ausgenommen - konnten hierbei Schülerinnen und Schüler mit Deutsch als Muttersprache bessere Ergebnisse erzielen. Die auffälligste Leistungsdifferenz war im Bezirk Friedrichshain-Kreuzberg mit 64 Punkten Unterschied zu beobachten.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen in *Muster und Strukturen*

Eine Kurzbeschreibung der mathematischen Kompetenzstufen findet sich in Kapitel 1, weitere Informationen finden sich in den didaktischen Handreichungen. Folgend werden die Verteilungen auf diese Kompetenzstufen berichtet.

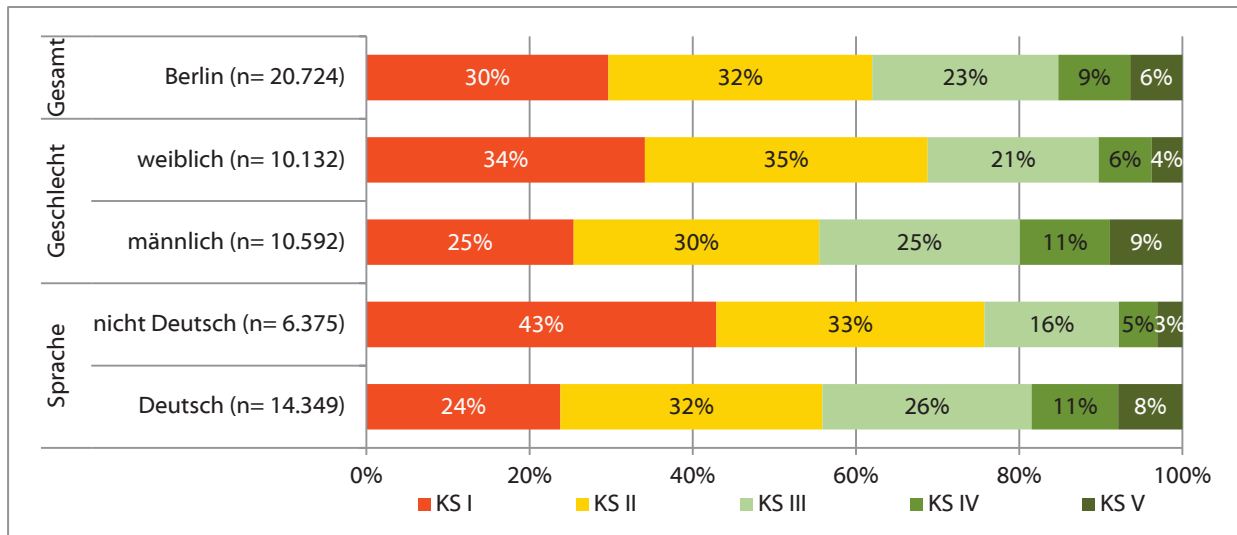


Abbildung 19: Verteilung der Kompetenzstufen in *Muster und Strukturen* (alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

Wie aus Abbildung 19 zu entnehmen ist, erreichten mit 30 Prozent knapp ein Drittel der Berliner Schülerinnen und Schüler im Inhaltsbereich *Muster und Strukturen* die Kompetenzstufe I und erfüllten somit nicht die Mindeststandards für das Ende der Klassenstufe 4. Bei den Mädchen machte der Anteil auf dieser Kompetenzstufe 34 Prozentpunkte aus, wohingegen bei den Jungen lediglich 25 % unter den Mindeststandards lagen. Auf den Kompetenzstufen II und III sind nur leichte Unterschiede zwischen Jungen und Mädchen auszumachen. Sie bildeten mit 30 und 25 % bei den Jungen und 35 und 21 % bei den Mädchen den größten Anteil in der Kompetenzstufenverteilung für diesen Teil des Mathematiktests. Für die Kompetenzstufen IV und V zeigt sich eine etwas größere Differenz zwischen beiden Geschlechtern, wobei die Jungen mit elf und neun Prozent einen um zehn Prozentpunkte größeren Anteil auf den beiden höchsten Kompetenzniveaus ausmachen als die Mädchen (KS IV: 6 %; KS V: 4 %).

Weitaus größer fallen die Differenzen bei den Kompetenzstufenverteilungen nach Herkunftssprache betrachtet aus: Demnach waren auf Kompetenzstufe I 43 % der Kinder mit nichtdeutscher Herkunftssprache zu verorten, wohingegen mit 24 % nur ein Viertel der Kinder mit Deutsch als Muttersprache die Mindeststandards für das Ende der Klassenstufe 4 nicht erfüllten. 33 % der Kinder mit nichtdeutscher Herkunftssprache und 32 % der Kinder mit Deutsch als Muttersprache erreichten Kompetenzstufe II, wohingegen der Anteil der deutschen Muttersprachler sowohl auf Kompetenzstufe III als auch auf den höchsten beiden Kompetenzniveaus erkennbar größer war. So erreichten elf und acht Prozent KS IV und V, wobei es bei den Kindern mit nichtdeutscher Herkunftssprache lediglich fünf und drei Prozent waren.

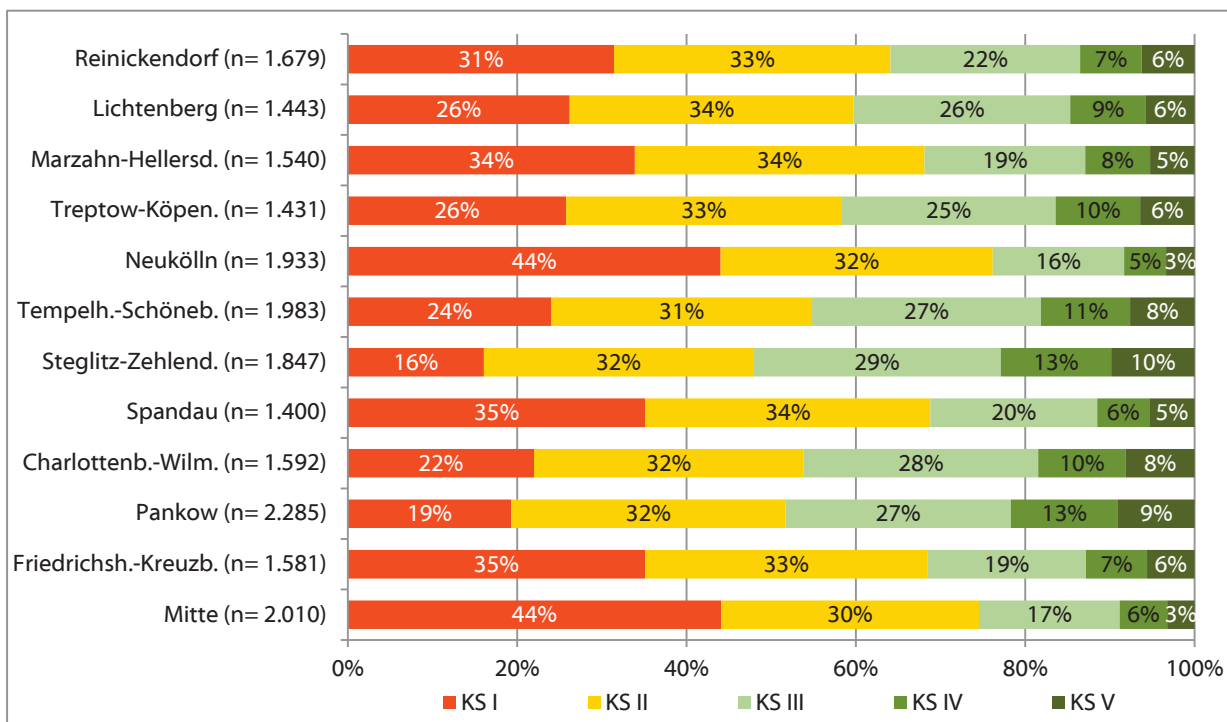


Abbildung 20: Bezirksspezifische Verteilung der Kompetenzstufen in *Muster und Strukturen*

In Abbildung 20 sind die **bezirksspezifischen Kompetenzstufenverteilungen** aller Berliner Schülerinnen und Schüler abgebildet. Demnach waren die Anteile der Kinder auf Kompetenzstufe I, welche die Mindeststandards für das Ende der Klassenstufe 4 nicht erfüllen, besonders in den Bezirken Mitte und Neukölln mit jeweils 44 % vergleichsweise hoch. In Steglitz-Zehlendorf machten diesen Anteil lediglich 16 % der Kinder aus. Hingegen fiel der Anteil von Schülerinnen und Schülern auf Kompetenzstufe V in diesem Bezirk mit 10 Prozent am höchsten aus. Komplementär dazu bildeten Mitte und Neukölln mit jeweils nur drei Prozent der Kinder auf dem höchsten Kompetenzstufenniveau das Schlusslicht Berlins für den Bereich *Muster und Strukturen*.

3.3.2 Mathematik – Größen und Messen

Lösungshäufigkeiten im Bereich *Größen und Messen*

Der Inhaltsbereich *Größen und Messen* umfasste insgesamt 31 Aufgaben, wobei eine maximale Gesamtpunktzahl von 31 Punkten erreicht werden konnte. Landesweit wurde der gesamte Testbereich im Durchschnitt zu 49 % korrekt gelöst. Die Jungen schnitten auch in diesem Inhaltsbereich mit einer durchschnittlichen Lösungshäufigkeit von 53 Prozentpunkten etwas besser ab als die Mädchen (44 %). Für Kinder mit Deutsch als Muttersprache ergab sich eine durchschnittliche Lösungshäufigkeit von 52 %, wohingegen Schülerinnen und Schüler mit nichtdeutscher Herkunftssprache durchschnittlich 40 % aller Aufgaben korrekt lösten.

Tabelle 9: Itemgenaue Lösungshäufigkeiten im Bereich *Größen und Messen*
(alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

	Aufgabe	Berlin (n= 20.724)	Jungen (n= 10.592)	Mädchen (n= 10.132)	Deutsch (n= 14.349)	ndH (n= 6.375)
18	Eier ausbrüten	66%	69%	62%	69%	58%
19	Käse	37%	43%	30%	41%	27%
20	Briefmarke	82%	84%	81%	85%	77%
21	Einkauf	50%	55%	45%	54%	41%
22	Komma	45%	48%	42%	47%	41%
23a	Ausflug I	28%	32%	24%	30%	22%
23b	Ausflug II	19%	22%	16%	21%	14%
24	Balkenwaage Tassen	49%	52%	46%	55%	35%
25	Längen	18%	23%	13%	21%	13%
26a	Einheiten I	79%	79%	79%	82%	71%
26b	Einheiten II	80%	82%	78%	84%	71%
26c	Einheiten III	80%	81%	78%	84%	71%
26d	Einheiten IV	75%	78%	73%	80%	65%
26e	Einheiten V	59%	61%	58%	65%	47%
26f	Einheiten VI	70%	74%	67%	76%	58%
26g	Einheiten VII	63%	67%	59%	69%	48%
27	Wettlauf	20%	23%	17%	24%	12%
28	Zeit	21%	25%	15%	23%	16%
29	Längenvergleich	88%	88%	88%	90%	84%
30a	Gewichte I	28%	34%	22%	31%	21%
30b	Gewichte II	55%	64%	45%	61%	41%
30c	Gewichte III	55%	59%	50%	59%	45%
31	Balkenwaage Apfel	50%	56%	44%	55%	38%
32a	Größen umwandeln I	54%	58%	49%	60%	40%
32b	Größen umwandeln II	62%	67%	57%	68%	50%
32c	Größen umwandeln III	24%	30%	17%	27%	18%
33	Theaterbesuch	5%	7%	3%	6%	3%
34	Freund besuchen	39%	42%	36%	43%	29%
35a	Zeitangaben I	30%	37%	22%	33%	22%
35b	Zeitangaben II	40%	48%	31%	44%	30%
35c	Zeitangaben III	35%	41%	28%	40%	23%
	Gesamt	49%	53%	44%	52%	40%

Der Abbildung lässt sich die **Verteilung der erreichten Gesamtpunktzahlen** entnehmen. Da im Testbereich *Größen und Messen* eine höhere Gesamtpunktzahl als in den anderen Inhaltsbereichen erreicht werden konnte, fallen die prozentualen Anteile für die einzelnen Punktsummen geringer aus als in den vorherigen Punktsummenanalysen. Die volle Punktzahl von insgesamt 31 erreichten demnach

lediglich 0,2 % aller teilnehmenden Schülerinnen und Schüler. Etwa 28 % der Kinder erreichten eine Gesamtpunktzahl von mindestens 20 Punkten, und etwa 50 Prozent aller Kinder konnten mehr als die Hälfte der zu erreichenden Gesamtpunktzahl erzielen. Schüler/-innen, die weniger als 10 Punkte in diesem Inhaltsbereich erlangten, machen einen Anteil von rund 23 Prozentpunkten aus. Durchschnittlich lösten die Schülerinnen und Schüler 14,9 von 31 Aufgaben richtig, was der in Tabelle 9 abgebildeten Lösungshäufigkeit von durchschnittlich 49 % entspricht.

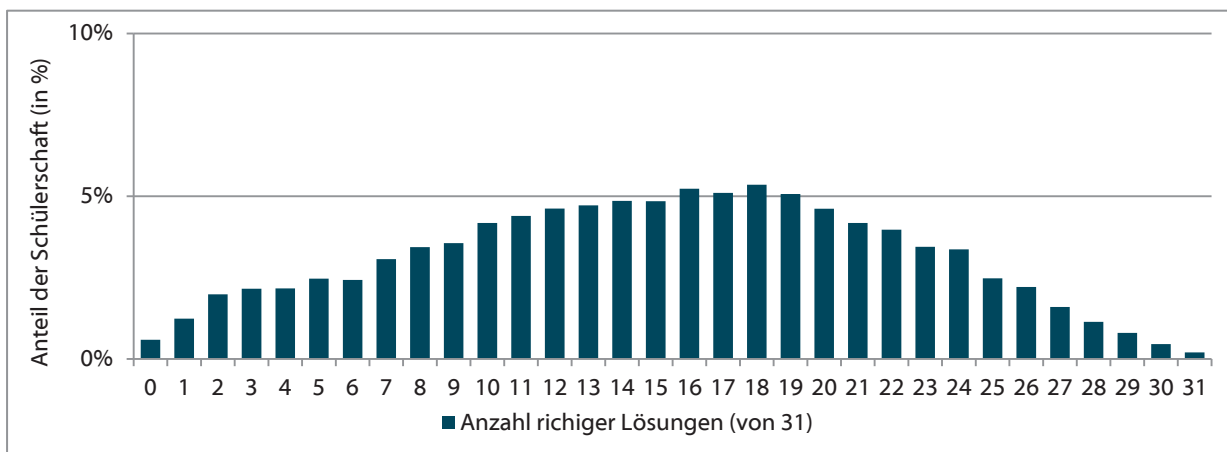
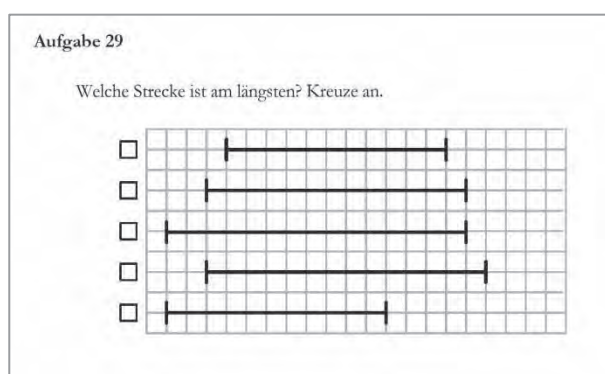


Abbildung 21: Verteilung der Testwerte in Punktsummen in Mathematik – Größen und Messen

Detaillierte Analyse einzelner Aufgaben aus dem Bereich *Größen und Messen*

Blick auf die leichteste Aufgabe – Aufgabe 29

Am häufigsten richtig gelöst wurde die Aufgabe 29 mit 88 %. In dieser Aufgabe musste die Länge von fünf Strecken miteinander verglichen werden, um die längste Strecke zu benennen. Obwohl diese weder links- noch rechtsbündig eingezeichnet waren, war die Aufgabe offensichtlich nicht schwierig, da durch ein unterlegtes Karomuster auch Kindern mit weniger ausgeprägtem räumlichem Vorstellungsvermögen die richtige Lösung durch Abzählen der Quadrate zwischen Anfangs- und Endpunkt der Strecke möglich war. Aus diesem Grund ist die Aufgabe entsprechend der Kompetenzstufe I und dem Anforderungsbereich I zugeordnet.



Blick auf die schwierigste Aufgabe – Aufgabe 33

Die Aufgabe 33 wurde mit 7 % am seltensten richtig gelöst. Sie ist gleichzeitig die Aufgabe, die mit 43 % am häufigsten nicht bearbeitet wurde. Hier handelt es sich um eine problemhaltige Sachaufgabe aus dem Größenbereich Geld. Um das Sachproblem zu verstehen, müssen die Kinder über Alltagsbegriffe wie Schultheater, Theaterkasse, Eintrittspreise in Abhängigkeit vom Alter (Erwachsener und Kind) verfügen und die Informationen aus zwei Darstellungen (Text und Tabelle) in Beziehung setzen und miteinander kombinieren. Die Problematik besteht darin, dass die Kinder für die Lösung der Aufgabe über kein gängiges Rechenverfahren verfügen, da sie Gleichungen noch nicht anwenden können. Außerdem muss bei dieser Aufgabe „rückwärts“ gearbeitet werden. Die Kinder können das Er-

gebnis durch Probieren ermitteln, wobei sie über Strategien verfügen müssen, die nach Möglichkeit auch noch systematisch sein sollten, um nicht zu viel Zeit für diese Aufgabe zu verwenden.

Aufgabe 33

Es wurden für 47,50 € Karten an der Theaterkasse gekauft.



Grafik © IQB

Für wie viele Erwachsene und wie viele Kinder wurden Karten gekauft?

Es wurden _____ Karten für Erwachsene und _____ Karten für Kinder gekauft.

Das Zeitproblem kann ein Hinweis darauf sein, dass diese Aufgabe von 43 % gar nicht bearbeitet oder beendet wurde. Sie wurde der Kompetenzstufe IV und dem Anforderungsbereich II zugeordnet.

Blick auf die Unterschiede zwischen Jungen und Mädchen

In Abbildung sind die Lösungshäufigkeiten für den Bereich *Größen und Messen* nach dem Geschlecht der Schülerinnen und Schüler dargestellt (vgl. Tabelle 9, Spalten 4 und 5). Die Items wurden nach absteigendem Schwierigkeitsgrad für die Mädchen geordnet. Demnach schnitten Jungen insgesamt besser ab als Mädchen, wobei die Differenzen weitestgehend gering ausfielen. Die größten Unterschiede zeigen sich bei den mittelschweren bis schweren Aufgaben, wohingegen bei den leichtesten Items die geringsten Differenzen auszumachen sind. So hatten Mädchen beispielsweise beim Lösen der Aufgaben 30a, 30b, 35a und 35b deutlich größere Probleme als Jungen. Bei der Aufgabe 35 b „Zeitangaben II“ handelte es sich um das Berechnen eines Zeitpunkts (Ende einer Film-DVD). Die dazu nötigen Angaben waren in Form einer Textaufgabe gegeben. Ob das schlechtere Abschneiden der Mädchen daran liegt, dass ihr Alltagsbezug zu DVD-Filmen geringer ist oder ob es daran liegt, dass sie mehr Schwierigkeiten mit dem nicht dekadischen Aufbau des Größenbereichs Zeit haben, kann hier nicht festgestellt werden, sollte aber bei der weiteren Arbeit im Unterricht bedacht werden. Bei den Aufgaben 29, 26a und 20 waren hingegen die Lösungshäufigkeiten bei beiden Geschlechtern fast identisch.

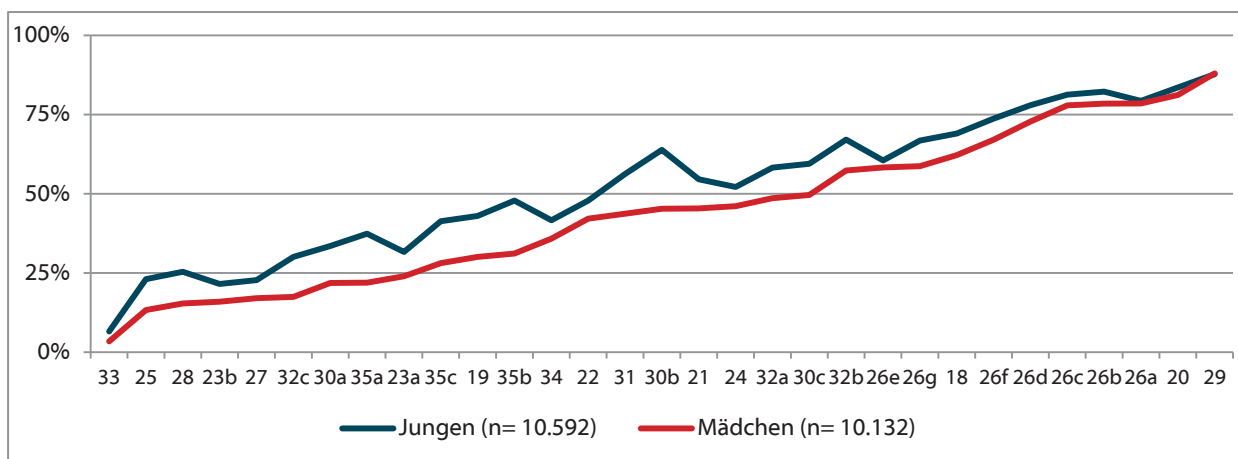


Abbildung 22: Itemspezifische Lösungshäufigkeiten in Mathematik – *Größen und Messen*, nach Geschlecht (Items geordnet nach der Schwierigkeit der Aufgaben für die Mädchen)

Blick auf die Ergebnisse von Kindern deutscher und nichtdeutscher Herkunftssprache

Weitaus deutlicher sind die Unterschiede zwischen den Lösungshäufigkeiten nach Herkunftssprache der Kinder aufgeschlüsselt. Der Abbildung lassen sich die Lösungshäufigkeiten sortiert nach absteigendem Schwierigkeitsgrad für Kinder mit nichtdeutscher Herkunftssprache entnehmen.

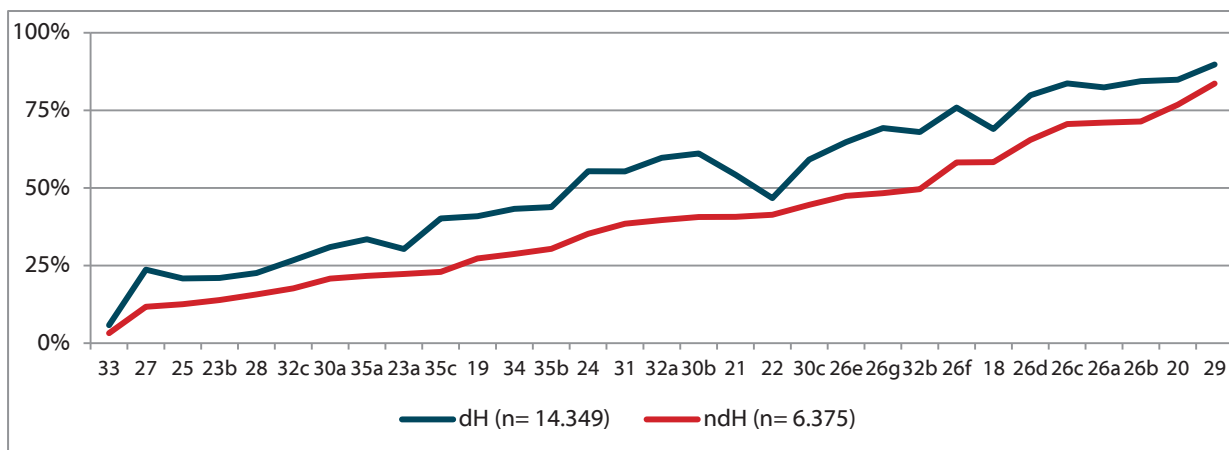
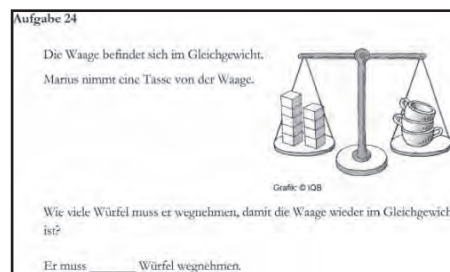


Abbildung 23: Itemspezifische Lösungshäufigkeiten in Mathematik – *Größen und Messen*, nach Sprachgebrauch (Items geordnet nach der Schwierigkeit der Aufgaben für die ndH-Kinder)

Insbesondere im mittleren Schwierigkeitsbereich (z. B. Aufgaben 24, 32a, 30b) finden sich eindeutige Differenzen zum Teil von 17 bis zu 21 Prozentpunkten. Auffällig ist jedoch, dass sich die Lösungshäufigkeiten beider Kategorien der Herkunftssprache bei den besonders leichten sowie bei den besonders schwer lösbaren Items leicht annähern. Dies lässt den Rückschluss zu, dass bei den schwierigen Aufgaben (z.B. Aufgabe 33, 28, 23b) vorrangig Probleme mit dem Format (z. B. offenes Format) oder der Aufgabenstellung bestanden

haben könnten, weniger auf sprachlicher Ebene. Bei Aufgabe 23b ‚Ausflug II‘ wurde beispielsweise eine offene Antwort in Form einer Beschreibung des Rechenweges verlangt, was scheinbar auch eine hohe Anzahl von Kindern mit Deutsch als Muttersprache überfordert hat.



Zum Abschluss des Kapitels zur Analyse der Lösungshäufigkeiten von Aufgaben im Inhaltsbereich *Größen und Messen* sollen die Aufgaben im Hinblick auf die **Nichtbearbeitung** (vgl. Abbildung 24) als Hinweis für zu umfangreiche Testmaterialien geprüft werden. Zum Testende hin lag die Bearbeitungsquote noch zwischen rund 71 % und 78 %, sodass davon ausgegangen werden kann, dass für einen Großteil der Schülerinnen und Schüler der Umfang der Aufgaben zu bewältigen war. Erwartungsgemäß wurden die Aufgaben am häufigsten nicht bearbeitet, die der Kompetenzstufe IV oder höher zugeordnet wurden (Bsp. 23b, 33).

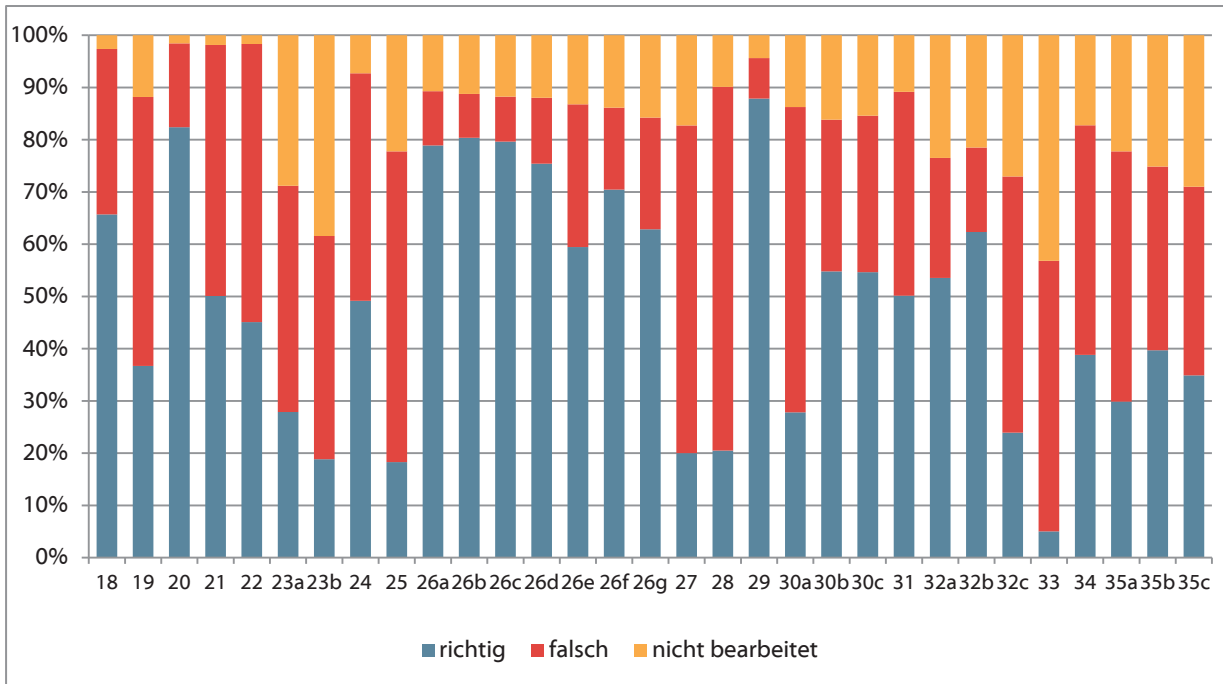


Abbildung 24: Bearbeitung der Aufgaben in Testreihenfolge in Mathematik – Größen und Messen

Auswertung nach Punkten der KMK-Bildungsstandards (BISTA-Punkte) im Bereich *Größen und Messen*

In Tabelle 10 werden die Leistungen der Berliner Schülerschaft für den Bereich *Größen und Messen* nach BISTA-Punkten bezirksspezifisch veranschaulicht. Hierbei variieren die Ergebnisse zwischen den Bezirken sehr stark von durchschnittlich 333 Punkten in Mitte bis zu 419 Punkten in Steglitz-Zehlendorf.

Mit Blick auf die Leistungsdifferenzen zwischen Jungen und Mädchen stechen vor allem Tempelhof-Schöneberg mit 50 Punkten und Reinickendorf mit 54 Punkten Unterschied zugunsten der Jungen hervor. Die geringste Differenz zwischen beiden Geschlechtern findet sich mit 28 BISTA-Punkten im Bezirk Steglitz-Zehlendorf.

Markanter unterschieden sich auch für diesen Inhaltsbereich die Leistungen der Kinder nach ihrer Herkunftssprache. Hier bildeten die Ergebnisse mit 20 BISTA-Punkten Unterschied in Lichtenberg und bis zu 83 Punkten Differenz in Friedrichshain-Kreuzberg eine sehr große Spanne im Leistungsspektrum von *Größen und Messen*.

Tabelle 10: Durchschnittliche BISTA-Punkte im Bereich *Größen und Messen*
(alle Schüler/-innen, nach Gruppen differenziert, Bezirksergebnisse)

Nr.	Bezirk	Berlin		Geschlecht				Herkunftssprache			
		Gesamt (n= 20.724)		männlich (n= 10.592)		Weiblich (n= 10.132)		Deutsch (n= 14.349)		nicht Deutsch (n= 6.375)	
		n	M	n	M	n	M	n	M	n	M
1	Mitte	2.010	333	1.026	350	984	315	711	367	1.299	314
2	Friedrichshain-Kreuzberg	1.581	366	827	388	754	341	872	403	709	320
3	Pankow	2.285	415	1.218	432	1.067	396	2.144	418	141	369
4	Charlottenburg-Wilmersdorf	1.592	402	821	419	771	383	1.014	418	578	373
5	Spandau	1.400	362	713	383	687	341	901	378	499	333
6	Steglitz-Zehlendorf	1.847	419	936	433	911	405	1.507	427	340	383
7	Tempelhof-Schöneberg	1.983	389	996	414	987	364	1.270	412	713	348
8	Neukölln	1.933	336	982	353	951	318	913	361	1.020	313
9	Treptow-Köpenick	1.431	397	696	416	735	379	1.349	400	82	343
10	Marzahn-Hellersdorf	1.540	364	794	386	746	341	1.293	370	247	332
11	Lichtenberg	1.443	391	729	409	714	373	1.159	395	284	375
12	Reinickendorf	1.679	365	854	392	825	338	1.216	385	463	314
	Gesamt	M= 379	SD= 115	M= 398	SD= 117	M= 358	SD= 109	M= 398	SD= 110	M= 335	SD= 114

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen in *Größen und Messen*

Die inhaltliche Beschreibung der Kompetenzstufen im Fach Mathematik findet sich in Kapitel 1. Neben dieser Kurzbeschreibung lässt sich in den didaktischen Handreichungen (<http://www.isq-bb.de/portal/>, S. 9) eine detaillierte Beschreibung der Kompetenzstufen für den Bereich *Größen und Messen* nachlesen.

In Abbildung 25 lässt sich die **Kompetenzstufenverteilung** dieses Inhaltsbereichs ablesen. Hierbei ist insbesondere der hohe Anteil der Berliner Schülerinnen und Schüler auf Kompetenzstufe I markant. Demnach erfüllten nur knapp die Hälfte der Kinder in Berlin die für diesen Bereich festgelegten Mindeststandards für das Ende der Klassenstufe 4. Diese Beobachtung wird noch verstärkt, wenn man die Kompetenzstufenverteilungen geschlechtsspezifisch betrachtet: So waren bei den Mädchen sogar 59 % auf diesem Kompetenzstufenniveau zu verorten, wohingegen die Jungen (43 % auf KS I) weitaus besser abschnitten. Auch auf den höchsten Kompetenzstufen werden die Verschiebungen zwischen beiden Geschlechtern klar sichtbar. Während der Anteil auf den Kompetenzstufen IV und V bei Mädchen lediglich fünf Prozent ausmachte, sind es zwölf Prozent aller Jungen, die sich auf den beiden höchsten Kompetenzniveaus bewegten (KS IV: 8 %; KS V: 4 %).

Mit Blick auf die Kompetenzstufenverteilungen nach Herkunftssprache wird die Ausprägung, welche einen erheblich hohen Anteil von Schülerinnen und Schülern auf Kompetenzstufe I aufweist, jedoch am deutlichsten. Mehr als zwei Drittel (68 %) aller Kinder nichtdeutscher Herkunftssprache lagen im Inhaltsbereich *Größen und Messen* unter den Mindeststandards für das Ende der Klassenstufe 4. Bei den Kindern mit Deutsch als Muttersprache machte dieser Anteil 43 % aus. Auch auf den folgenden Kompetenzstufen ist eine starke Verschiebung bei dieser Betrachtung erkennbar. So ordneten sich auf Kompetenzstufe II und III 46 % der Kinder mit Deutsch als Muttersprache ein, wohingegen der Anteil bei Kindern mit nichtdeutscher Herkunftssprache 18 % weniger betrug (28 %). Die höchste Kompetenzstufe erreichten bei Kindern mit Deutsch als Herkunftssprache immerhin noch drei Prozent, wohingegen Kinder nichtdeutscher Herkunftssprache nur zu einem Prozent auf diesem Niveau vertreten waren.

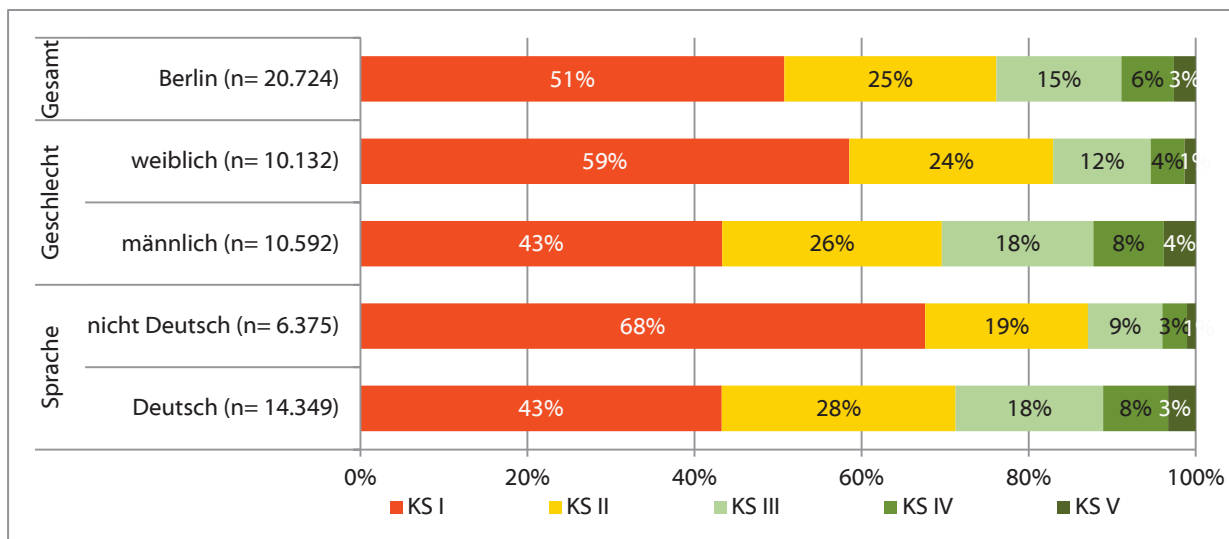


Abbildung 25: Verteilung der Kompetenzstufen in *Größen und Messen* (alle Schüler/-innen, nach Geschlecht und Herkunftssprache differenziert)

Für den Bereich *Größen und Messen* werden die Kompetenzstufenverteilungen in Abbildung bezirksspezifisch dargestellt. Grundsätzlich lässt sich für alle Bezirke ein erheblich höherer Anteil von Kindern feststellen, welche die Kompetenzstufe II nicht erreicht haben.

Die höchsten Anteile weisen hierbei Neukölln mit 68 % und der Bezirk Mitte mit 67 % der Schülerinnen und Schüler auf Kompetenzstufe I auf. Am besten konnten auch in diesem Inhaltsbereich die Bezirke Steglitz-Zehlendorf und Pankow mit verhältnismäßig geringen 35 % bzw. 37 % Prozent auf diesem Kompetenzniveau abschneiden.

Kompetenzstufe V konnte im Bereich *Größen und Messen* nur von einem relativ geringen Anteil von bis zu vier Prozent der Kinder in den Bezirken Steglitz-Zehlendorf, Pankow und Charlottenburg-Wilmersdorf erreicht werden, wobei auf diesem Kompetenzniveau die Prozentwerte nicht so stark divergierten. Den geringsten Anteil machten hier die Bezirke Neukölln, Mitte und Marzahn-Hellersdorf mit jeweils einem Prozent aus.

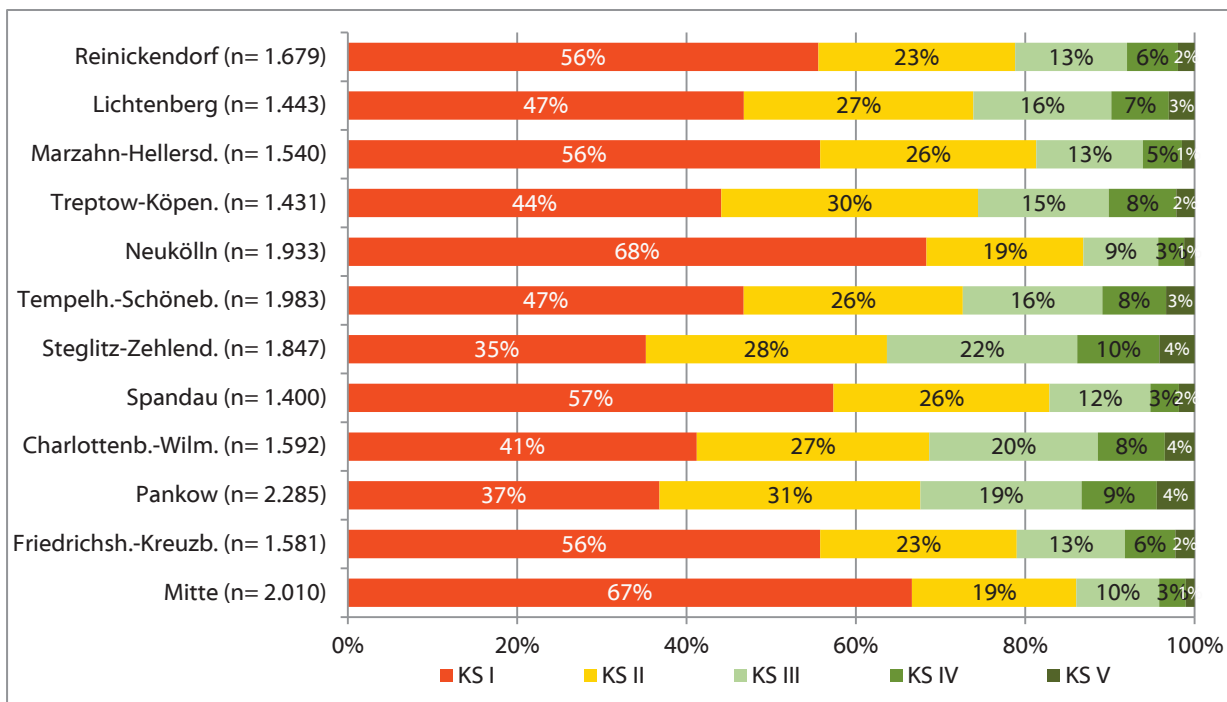


Abbildung 26: Bezirksspezifische Verteilung der Kompetenzstufen in *Größen und Messen*

4. Übergreifende Analysen

Zusätzlich zu den Darstellungen der Grundinformationen zu den beiden Testteilen im Fach Deutsch und den beiden Testteilen im Fach Mathematik wurden als Erweiterung der Berichterstattung zu VERA 3 2012, analog zum Vorjahr, übergreifende Analysen berechnet. In diesem Kapitel werden die Ergebnisse abschließend im Hinblick auf das Abschneiden der Kontextgruppen sowie die Zusammenhänge zwischen den erbrachten Leistungen in den verschiedenen Testteilen untersucht.

4.1 Soziale Herkunft – VERA-Ergebnisse nach Kontextgruppen

Im folgenden Kapitel werden die Ergebnisse von VERA 3 in diesem Jahr im Hinblick auf die Bedeutsamkeit der Zusammensetzung der Schülerschaft untersucht. Zur Differenzierung wurde auf zwei Maße zurückgegriffen, die auch für die Zuteilung der Mittel für die Sprachförderung an Schulen durch die Senatsverwaltung für Bildung, Jugend und Wissenschaft herangezogen werden: der Anteil der Schülerinnen und Schüler nichtdeutscher Herkunftssprache (ndH) und der Anteil der Schülerinnen und Schüler mit Lernmittelbefreiung (LMB) in der dritten Jahrgangsstufe.

Zur Kontextgruppenbildung wurde dabei jedes der beiden Maße am Median geteilt, d. h. jeweils so zerlegt, dass zwei Gruppen mit je 50 % der Schulen entstanden (Anteil der Schüler/-innen nichtdeutscher Herkunftssprache: niedrig oder hoch; Anteil der Schüler/-innen mit Lernmittelbefreiung: niedrig oder hoch). Anschließend wurden die Merkmale gekreuzt, sodass drei Vergleichsgruppen entstanden. Die Kontextgruppe 1 (KG 1) umfasst die N=116 Schulen, deren ndH- und LMB-Anteil relativ niedrig ausgeprägt ist, während in der KG 3 die N=170 Schulen zu finden sind, die sich durch einen hohen ndH- sowie LMB-Anteil kennzeichnen lassen. Da sie sich nicht eindeutig einer dieser beiden Gruppen zuordnen lassen, finden sich in KG 2 jene N=86 Schulen, deren ndH- oder LMB-Anteil hoch ist, während das andere Merkmal niedrig ausgeprägt ist (Tabelle 11).

Tabelle 11: Kontextgruppenbildung aufgrund der Merkmale Anteil ndH und Anteil LMB

		Anteil der Schülerschaft nichtdeutscher Herkunftssprache (ndH)	
		niedrig (0-24,1 %)	hoch (24,1-100 %)
Anteil der Schülerschaft mit Lernmittelbefreiung (LMB)	niedrig (0-30,2 %)	KG 1 (116 Schulen)	KG 2 (50 Schulen)
	hoch (30,2-100 %)	KG 2 (36 Schulen)	KG 3 (170 Schulen)

Die Verteilung der Schulen auf eine der drei Kontextgruppen unterscheidet sich zwischen den Bezirken erheblich (vgl. Abbildung). Während in Pankow beispielsweise nur eine Schule der Kontextgruppe 3 (Schulen mit einem hohen Anteil an Schülerinnen und Schülern nichtdeutscher Herkunftssprache sowie einem hohen Anteil der Schülerschaft mit Lernmittelbefreiung) zuzuordnen war, zählten in den Bezirken Mitte und Neukölln über 80 % aller Schulen zu dieser Kontextgruppe 3.

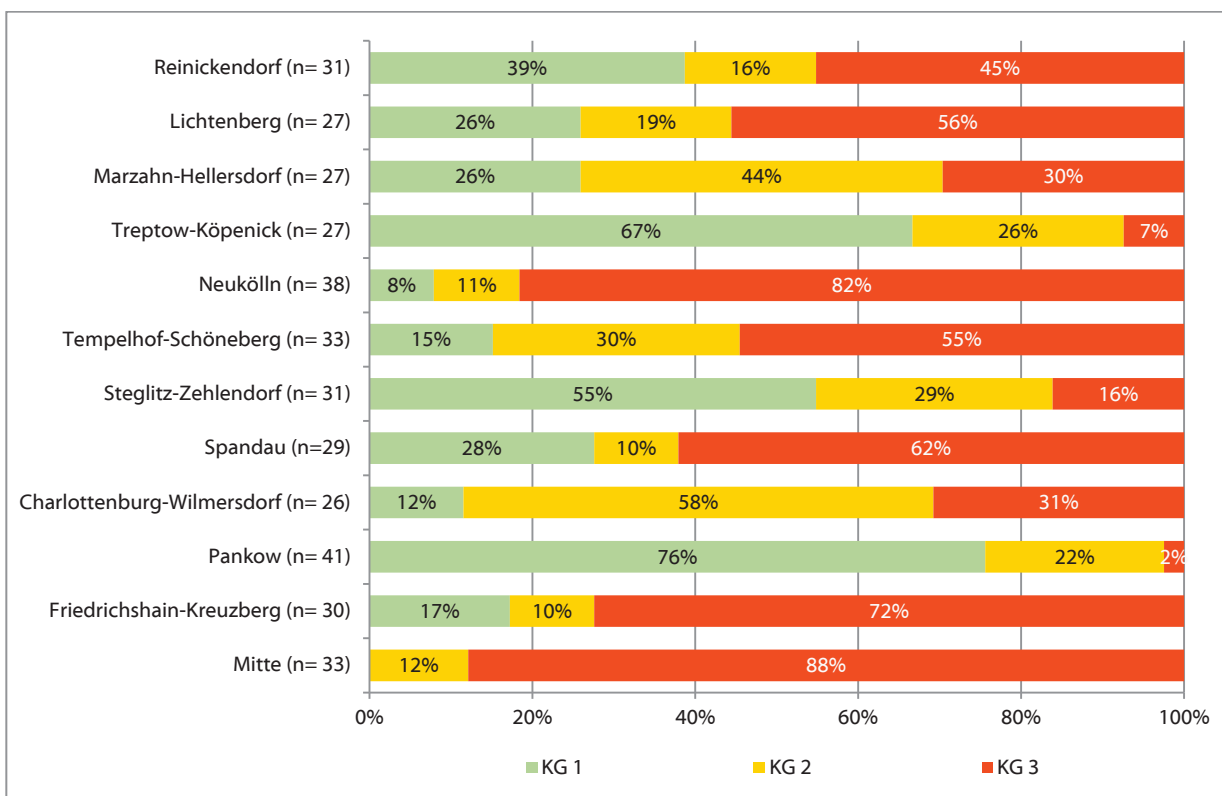


Abbildung 27: Bezirksspezifische Verteilung der Kontextgruppen (KG) in Berlin

Die Kompetenzstufenverteilungen im *Leseverstehen* unterschieden sich zwischen den drei Kontextgruppen erheblich (vgl.

Abbildung 28). In der Kontextgruppe 1 war erwartungsgemäß der Anteil der Schülerinnen und Schüler, die mit mindestens Kompetenzstufe II bereits den Mindeststandard für die Jahrgangsstufe 4 erreichten mit 85 % am größten, aber auch in den anderen beiden Kontextgruppen gelang es bereits vielen Schülerinnen und Schülern, den Mindeststandard zu erreichen (KG 2: 72 %; KG 3: 55 %).

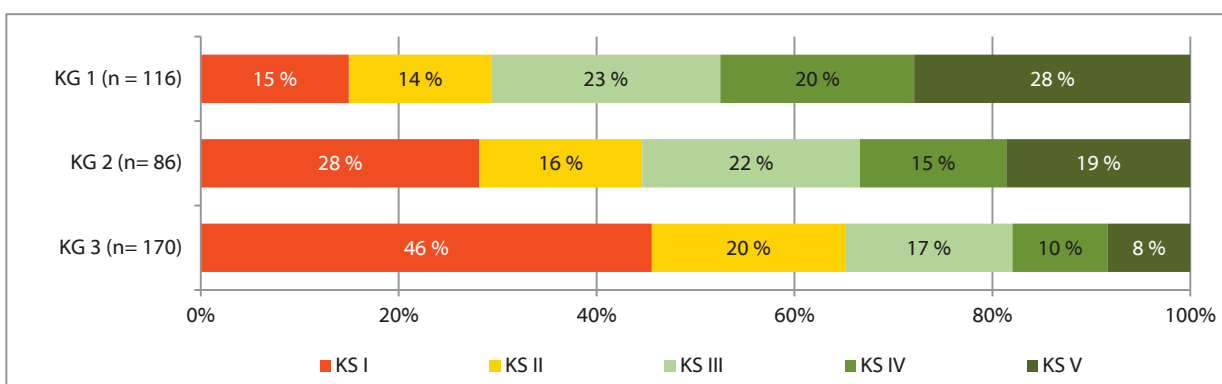


Abbildung 28: Kompetenzstufenverteilungen in *Leseverstehen*, nach Kontextgruppen

Rund 35 Prozent der Schülerinnen und Schüler aus den Schulen der KG 3 erreichten mindestens Kompetenzstufe III und übertrafen damit bereits am Ende der Jahrgangsstufe 3 die in den Bildungsstandards für das Ende der Jahrgangsstufe 4 formulierten Mindestenerwartungen. Eingehend betrachtet werden sollten gerade auch im Hinblick auf individuelle Fördermaßnahmen in der Jahrgangsstufe 4 die Leistungen der Schülerinnen und Schüler, die bei VERA 3 die Kompetenzstufe I erreichten. Ihr An-

teil lag in den Schulen der KG 1 bei 15 %, machte in KG 2 und KG 3 aber ca. 28 % bzw. knapp 46 % der Schülerschaft aus.

Für den Bereich *Sprachgebrauch* stellt sich die Verteilung ähnlich dar wie im *Leseverstehen* (vgl. Abbildung). Jedoch zeigen die Ergebnisse, dass an beiden Extremen der Verteilung (Kompetenzstufen I und V) vor allem für die Kontextgruppen 1 und 2 weniger Schulen zu finden waren als beispielsweise im *Leseverstehen*. Das bedeutet, dass sich die Schulen in diesem Bereich häufiger in den Kompetenzstufen 2 bis 4 verorten ließen. Dennoch erreichten in der Kontextgruppe 3 auch hier knapp über 30 % der Schülerinnen und Schüler nicht den Mindeststandard (Kompetenzstufe II) für den Bereich *Sprachgebrauch*.

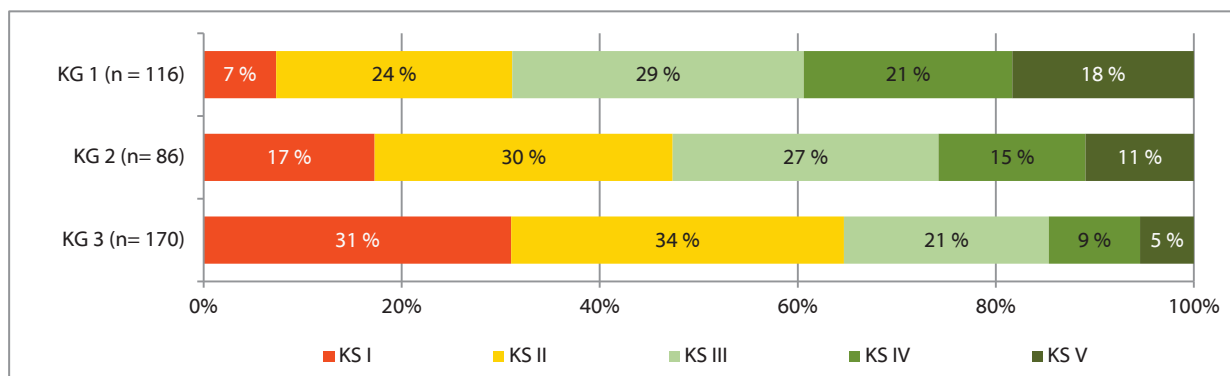


Abbildung 29: Kompetenzstufenverteilungen in *Sprachgebrauch* nach Kontextgruppen

Im Bereich *Muster und Strukturen* des Faches Mathematik erreichten 48 % der Schülerinnen und Schüler von Schulen der Kontextgruppe 1 mindestens die Kompetenzstufe III (vgl.

Abbildung 30). Insgesamt 80 % dieser Schulen erreichten mindestens den Mindeststandard. Für die Kontextgruppen 2 und 3 trifft das auf 73 % bzw. 56 % der Schulen zu. Weiterhin erreichten in der Kontextgruppe 3 lediglich 24 % der Schulen mindestens den Regelstandard für diesen Bereich.

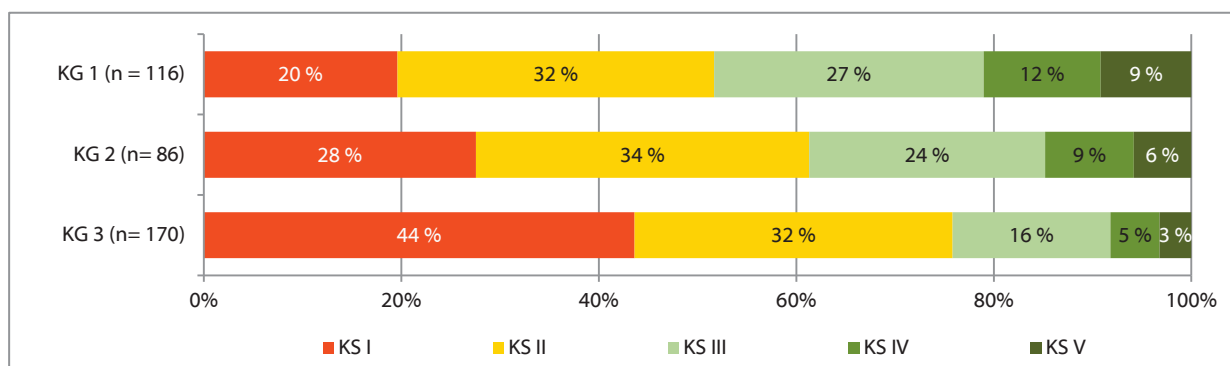


Abbildung 30: Kompetenzstufenverteilungen in *Muster und Strukturen*, nach Kontextgruppen

Der Bereich *Größen und Messen* fiel über alle Kontextgruppen und Testdomänen hinweg am schlechtesten aus (vgl. Abbildung 31). Hier lag der Anteil der Schülerinnen und Schüler, die unter dem Mindeststandard verortet wurden, bei gut einem Drittel in Kontextgruppe 1 bis über zwei Drittel in Kontextgruppe 3. Jedoch erreichten auch hier 34 % der Schulen aus Kontextgruppe 1 mindestens den Regelstandard.

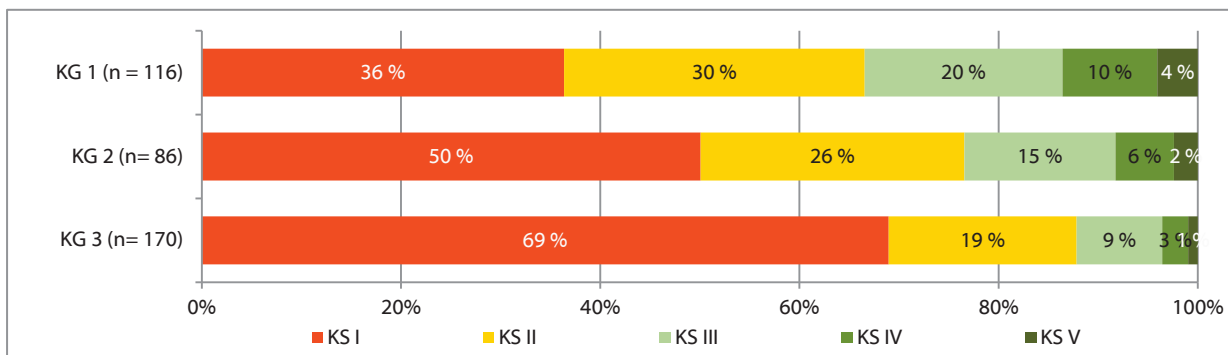


Abbildung 31: Kompetenzstufenverteilungen in *Größen und Messen*, nach Kontextgruppen

4.2 Korrelationen und Kreuztabellierungen der Einzeltests

Die Korrelationen der erreichten Punktwerte der Leistungstests in Deutsch und Mathematik zeigen in allen Fällen mittlere Beziehungen zwischen den Fähigkeiten der Schülerinnen sowie Schüler: Beispielsweise fallen die Beziehungen zwischen *Leseverstehen bzw. Sprachgebrauch* und *Muster und Strukturen* mit $r=.56$ und $r=.57$ moderat aus. Erwartungsgemäß sind die Zusammenhänge zwischen den Bereichen eines Faches etwas höher als über die Fächer hinweg. So liegt der Korrelationskoeffizient zwischen *Leseverstehen* und *Sprachgebrauch* bei $r=.72$ und zwischen *Muster und Strukturen* und *Größen und Messen* bei $r=.77$ (vgl. Tabelle 12). Die positiven Korrelationen zwischen den verschiedenen Testteilen weisen darauf hin, dass viele Schüler/-innen über die Testteile hinweg eher gute bzw. schlechte Leistungen erbracht haben, sie also häufig über Leistungsbereiche hinweg ähnliche Kompetenzstufen erreichten.

Tabelle 12: Interkorrelationen der Inhaltsbereiche bei VERA 3 2011

	Deutsch - <i>Leseverstehen</i>	Deutsch - <i>Sprachgebrauch</i>	<i>Muster und Strukturen</i>
<i>Sprachgebrauch</i>	.72		
<i>Muster und Strukturen</i>	.56	.57	
<i>Größen und Messen</i>	.59	.59	.77

Veranschaulicht man die korrelativen Beziehungen zwischen den Leistungstests über die Kreuztabellierung der Kompetenzstufenzugehörigkeiten, so zeigt sich, dass die meisten Schülerinnen und Schüler in der 3. Jahrgangsstufe eine hohe Konsistenz in den Leistungen über verschiedene Anforderungsbereiche bzw. Fächer hinweg aufwiesen. Zur Veranschaulichung werden die Werte der Anteile, die über 10 % liegen, fett hervorgehoben und Anteile von unter 5 % grau gefärbt.

Tabelle 13: Inhaltsübergreifendes Verhältnis der KS-Gruppen

		Deutsch – <i>Leseverstehen</i>					Deutsch – <i>Sprachgebrauch</i>					Mathematik – <i>Muster und Strukturen</i>				
		KS I	KS II	KS III	KS IV	KS V	KS I	KS II	KS III	KS IV	KS V	KS I	KS II	KS III	KS IV	KS V
<i>Sprachgebrauch</i>	KS I	15%	3%	1%	0%	0%										
	KS II	10%	7%	7%	3%	1%										
	KS III	3%	5%	8%	6%	5%										
	KS IV	1%	1%	3%	4%	6%										
	KS V	0%	0%	1%	2%	8%										
<i>M & S</i>	KS I	17%	6%	4%	2%	1%	13%	11%	4%	1%	0%					
	KS II	8%	7%	8%	5%	4%	5%	11%	9%	4%	2%					
	KS III	3%	3%	6%	5%	6%	1%	5%	7%	5%	4%					
	KS IV	1%	1%	2%	2%	4%	0%	1%	3%	2%	3%					
	KS V	0%	0%	1%	1%	4%	0%	1%	2%	2%	3%					
<i>G & M</i>	KS I	23%	10%	9%	4%	3%	16%	18%	10%	4%	2%	27%	18%	4%	1%	0%
	KS II	4%	4%	7%	5%	6%	2%	7%	8%	5%	3%	2%	11%	10%	2%	1%
	KS III	1%	2%	3%	4%	6%	0%	3%	5%	4%	4%	0%	3%	7%	3%	2%
	KS IV	0%	0%	1%	2%	3%	0%	1%	2%	2%	2%	0%	0%	2%	2%	2%
	KS V	0%	0%	0%	1%	2%	0%	0%	1%	1%	1%	0%	0%	0%	1%	2%

Erwartungsgemäß war die Konsistenz innerhalb eines Faches am größten. So erreichten 42 % der Schülerinnen und Schüler im Vergleich der beiden Deutschtests und 49 % der Schülerinnen und Schüler in beiden Bereichen im Fach Mathematik identische Kompetenzstufen (siehe grau hinterlegte Zellen). Beim Blick auf die fachübergreifenden Verteilungen ist zu erkennen, dass der Anteil der Schülerinnen und Schüler, die in den beiden jeweiligen Testbereichen die gleiche Kompetenzstufe erreichten, zwischen 32 und 36 % liegt. Innerhalb eines Faches erreichten 25 % der Schülerinnen und Schüler im *Sprachgebrauch* und 32 % im *Leseverstehen* eine höhere Kompetenzstufe als im jeweils anderen Bereich. 43 % der Schülerschaft erreichten im Inhaltsbereich *Muster und Strukturen* eine höhere Kompetenzstufe als in *Größen und Messen*, für lediglich 8 % der Schülerschaft zeigte sich das Gegenteil. Hierbei ist jedoch anzumerken, dass sich nur in seltenen Fällen große Abweichungen zwischen den erreichten Kompetenzstufen zeigten. So gab es beispielsweise keinen einzigen Schüler, der in *Muster und Strukturen* die Kompetenzstufe I belegt hatte, in *Größen und Messen* hingegen die Kompetenzstufe V.

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

www.isq-bb.de

