Institut für Bildungsevaluation

Check P4 2026

Kantonsbericht zuhanden der Direktion für Bildung und Kultur, Kanton Zug

Zürich, 31. August 2026

Anschrift

Hofwiesenstrasse 349
8050 Zürich
Telefon +41 43 268 39 60
www.ibe-edu.ch

Zusammenfassung

Der Check P4 ermöglicht eine externe Standortbestimmung in den Fächern Deutsch, Englisch und Mathematik anhand standardisierter Online-Tests. Im März 2026 wurde der Check P4 in allen 4. Klassen der gemeindlichen Schulen im Kanton Zug durchgeführt. Es handelte sich um die zweite reguläre Durchführung nach der Pilotdurchführung im Mai 2024 und der ersten regulären Durchführung im März 2025. Die Ergebnisse dienen der individuellen Beurteilung sowie der Unterrichts- und Schulentwicklung. Darüber hinaus können sie für das Monitoring des kantonalen Bildungssystems genutzt werden, etwa zur Analyse von Leistungsunterschieden nach individuellen Merkmalen oder zur Einordnung der Ergebnisse anhand des Orientierungspunkts in der Mitte des 2. Zyklus des Lehrplans 21. Durch die wiederholte Durchführung des Check P4 lassen sich die Leistungen inzwischen über drei Jahre hinweg vergleichen.

Der vorliegende Bericht informiert über die Ergebnisse des Check P4 2026 auf kantonaler Ebene und ordnet sie im Vergleich zu den Durchführungen 2024 und 2025 ein. Die Ergebnisse bestätigen die besondere Stärke der Zuger Schülerinnen und Schüler im Fach Englisch: In Englisch Hören erreichen 91 Prozent und in Englisch Lesen 86 Prozent mindestens Kompetenzstufe V und damit in der Regel die bis zum Orientierungspunkt in der Mitte des 2. Zyklus vorgesehenen Kompetenzen. In Deutsch und Mathematik liegt dieser Anteil je nach Kompetenzbereich bei rund 59 bis 65 Prozent.

Im Vergleich zum Vorjahr blieben die mittleren Leistungen in allen drei Fächern weitgehend stabil. In Deutsch liegt der Mittelwert 2026 um 1 Punkt höher als 2025, in Englisch um 4 Punkte und in Mathematik um 8 Punkte. Diese Unterschiede sind sehr gering. Der Rückgang, der in Deutsch und Mathematik von der Pilotdurchführung 2024 zur ersten regulären Durchführung 2025 beobachtet wurde, setzte sich damit 2026 nicht fort. Er dürfte zumindest teilweise mit dem unterschiedlichen Durchführungszeitpunkt zusammenhängen: Die Pilotdurchführung 2024 fand im Mai statt, die regulären Durchführungen 2025 und 2026 hingegen im März. In Englisch zeigt sich über die drei Durchführungen hinweg ein anderes Muster: Die mittleren Ergebnisse stiegen von 2024 auf 2025 an und blieben 2026 auf einem ähnlichen Niveau.

Die Analysen nach individuellen Merkmalen zeigen überwiegend stabile Muster. In Deutsch erzielen Schülerinnen und Schüler mit Deutsch als Erstsprache höhere Ergebnisse als Schülerinnen und Schüler mit einer anderen Erstsprache. Dieser Unterschied hatte von 2024 auf 2025 zugenommen, blieb 2026 jedoch weitgehend stabil. Auch Mädchen erzielen in Deutsch in allen drei Jahren höhere Ergebnisse als Knaben; auch dieser Unterschied blieb von 2025 auf 2026 weitgehend stabil. In Englisch erreichen Schülerinnen und Schüler mit einer anderen Erstsprache über alle drei Durchführungen hinweg deutlich höhere Ergebnisse als Schülerinnen und Schüler mit Deutsch als Erstsprache. Dieses Muster dürfte damit zusammenhängen, dass ein Teil der Schülerinnen und Schüler mit einer anderen Erstsprache Englisch als Erstsprache spricht. Der Unterschied zwischen den Geschlechtern ist 2026 gegenüber den beiden Vorjahren grösser, bleibt gemessen an der Effektstärke jedoch gering. In Mathematik erzielen Schülerinnen und Schüler mit Deutsch als Erstsprache höhere Ergebnisse als Schülerinnen und Schüler mit einer anderen Erstsprache; dieser Unterschied hat über die drei Durchführungen tendenziell zugenommen, bleibt gemessen an der Effektstärke jedoch klein. Knaben erreichen in Mathematik höhere Ergebnisse als Mädchen; der Unterschied ist seit 2025 weitgehend stabil.

Besondere Vorsicht ist bei der Interpretation der Ergebnisse zu Schülerinnen und Schülern mit angepassten Lernzielen erforderlich. Die Zahl der Schülerinnen und Schüler mit registrierten angepassten Lernzielen ist 2026 gegenüber den Vorjahren deutlich gesunken. Gleichzeitig nahmen Schülerinnen und Schüler mit angepassten Lernzielen häufiger nicht am Check P4 teil oder schlossen den Test nicht ab. Die deutlich grösseren Leistungsunterschiede zwischen Schülerinnen und Schülern mit regulären und angepassten Lernzielen im Jahr 2026 lassen sich deshalb nicht eindeutig als Leistungsentwicklung interpretieren. Sie dürften zumindest teilweise mit einer veränderten Zusammensetzung der ausgewerteten Gruppe zusammenhängen. Für die Interpretation zukünftiger Ergebnisse ist eine möglichst einheitliche Handhabung und Registrierung angepasster Lernziele von Bedeutung.

Zusätzlich zu den Unterschieden zwischen Gruppen mit verschiedenen individuellen Merkmalen wurden auch Unterschiede zwischen Schulklassen und Gemeinden untersucht. In allen drei Fächern zeigten sich deutliche Unterschiede zwischen den Klassenmittelwerten. Rund 14 bis 15 Prozent der Unterschiede in den individuellen Ergebnissen entfallen auf Unterschiede zwischen Klassen. Dies unterstreicht die Bedeutung einer standardisierten, klassenübergreifend vergleichbaren Standortbestimmung als Ergänzung zu den stärker vom jeweiligen Klassenkontext geprägten Schulnoten. Zwischen den Gemeinden fallen die Unterschiede dagegen deutlich geringer aus: Nur ein kleiner Anteil der Leistungsunterschiede entfällt auf Unterschiede zwischen Gemeinden. Auch die Veränderungen der Gemeindemittelwerte von 2025 zu 2026 zeigen kein einheitliches Muster und sind insbesondere bei kleineren Gemeinden vorsichtig zu interpretieren, da die Ergebnisse einzelner Schülerinnen und Schüler den Gemeindemittelwert stärker beeinflussen können.

Ergänzend zum Check P4 stand den Schülerinnen und Schülern die Aufgabensammlung Mindsteps während des Schuljahres zur Verfügung. Im Schuljahr 2025/26 nutzte die Mehrheit der Schülerinnen und Schüler Mindsteps in allen drei Fächern mindestens einmal. Gegenüber dem Vorjahr stieg der Anteil der Nutzerinnen und Nutzer in Deutsch, Englisch und Mathematik. Gleichzeitig veränderte sich die Art der Nutzung: In Deutsch und Mathematik verschob sich die Nutzung stärker in Richtung des von Lehrpersonen gesteuerten Bereichs Mind, während die Nutzung von Steps in diesen beiden Fächern zurückging. In Englisch nahm die Nutzung sowohl von Mind als auch von Steps zu.

Wie bereits in den Vorjahren zeigt sich ein positiver Zusammenhang zwischen der Häufigkeit der Mindsteps-Nutzung und den Ergebnissen im Check P4. Schülerinnen und Schüler, die mehr Aufgabenserien in Mindsteps bearbeiteten, erzielten im Durchschnitt höhere Ergebnisse. Die Zusammenhänge fallen jedoch insgesamt gering aus und erlauben keine Aussage über eine ursächliche Wirkung von Mindsteps auf die Leistungen. Möglich ist auch, dass weitere Merkmale, etwa das Engagement der Schülerinnen und Schüler oder der Lehrpersonen, die Unterrichtsgestaltung oder das Schulumfeld, sowohl die Nutzung von Mindsteps als auch die Ergebnisse im Check P4 beeinflussen.

Insgesamt zeigt der Check P4 2026 ein weitgehend stabiles Leistungsniveau gegenüber dem Vorjahr. Besonders hervorzuheben sind die weiterhin hohen Ergebnisse in Englisch, die deutlichen Unterschiede zwischen Klassen sowie die veränderte Nutzung von Mindsteps. Für das Bildungsmonitoring bleiben drei Punkte zentral: die weitere Beobachtung der Leistungsentwicklung über zusätzliche Durchführungen hinweg, die differenzierte Analyse von Leistungsunterschieden nach individuellen Merkmalen und Klassen sowie eine konsistente Erfassung angepasster Lernziele als Grundlage für vergleichbare Auswertungen.

Inhalt

1	**Einleitung**	**6**
1.1	Ausgangslage	6
1.2	Tests und Methoden	6
1.3	Teilnehmende Schülerinnen und Schüler	8
2	**Ergebnisse pro Fach im Jahr 2026**	**10**
2.1	Mittelwert und Streuung der Ergebnisse	10
2.2	Ergebnisse nach individuellen Merkmalen	10
2.3	Zusammenhang zwischen den Fächern	14
2.4	Ergebnisse nach Kompetenzstufen	15
2.5	Unterschiede zwischen Klassen	17
2.6	Unterschiede zwischen Gemeinden	18
3	**Ergebnisse pro Fach im Trend**	**20**
3.1	Trend seit 2024	20
3.2	Trends nach individuellen Merkmalen	22
3.3	Trends nach Kompetenzstufen	26
3.4	Veränderung der Gemeindemittelwerte	27
4	**Nutzung von Mindsteps und Ergebnisse im Check P4**	**29**
4.1	Nutzung von Mindsteps	29
4.2	Zusammenhang der Mindsteps-Nutzung und der Ergebnisse im Check P4	31

1 Einleitung

1.1 Ausgangslage

Der Check P4 ermöglicht eine externe Leistungsmessung in den Fächern Deutsch, Englisch und Mathematik anhand standardisierter Online-Tests. Die Ergebnisse dienen Schülerinnen und Schülern, Lehrpersonen und Schulleitungen als unabhängige Standortbestimmung. Sie können für die individuelle Beurteilung, die Unterrichts- und Schulentwicklung sowie für das Monitoring des kantonalen Bildungssystems genutzt werden.

Der Check P4 wurde im März 2026 in allen 4. Klassen der gemeindlichen Schulen im Kanton Zug durchgeführt. Es handelte sich um die zweite reguläre Durchführung nach der Pilotdurchführung im Mai 2024 und der ersten regulären Durchführung im März 2025.

Die individuellen Ergebnisse der Schülerinnen und Schüler stehen den teilnehmenden Lehrpersonen auf der Onlineplattform «Check Dein Wissen» zur Verfügung. Schulleitungen haben Zugriff auf die aggregierten Klassen- und Schulergebnisse.

Der vorliegende Bericht präsentiert die Ergebnisse der drei bisherigen Durchführungen des Check P4 mit einem Schwerpunkt auf der Durchführung 2026. Kapitel 1.2 beschreibt die standardisierten Online-Tests und die verwendeten Methoden, Kapitel 1.3 die teilnehmenden Schülerinnen und Schüler. Kapitel 2 präsentiert die Ergebnisse nach Fächern und untersucht Unterschiede nach individuellen Merkmalen, Zusammenhänge zwischen den Fächern, die erreichten Kompetenzstufen sowie Unterschiede zwischen Klassen und Gemeinden. Kapitel 3 vergleicht die Ergebnisse 2026 mit jenen der bisherigen Durchführungen. Kapitel 4 untersucht die Nutzung der digitalen Aufgabensammlung Mindsteps und deren Zusammenhang mit den Ergebnissen im Check P4 im Jahr 2026 sowie über die drei bisherigen Durchführungen hinweg.

1.2 Tests und Methoden

Tabelle 1: Fächer, Kompetenzbereiche, Anzahl Aufgaben und Testdauer

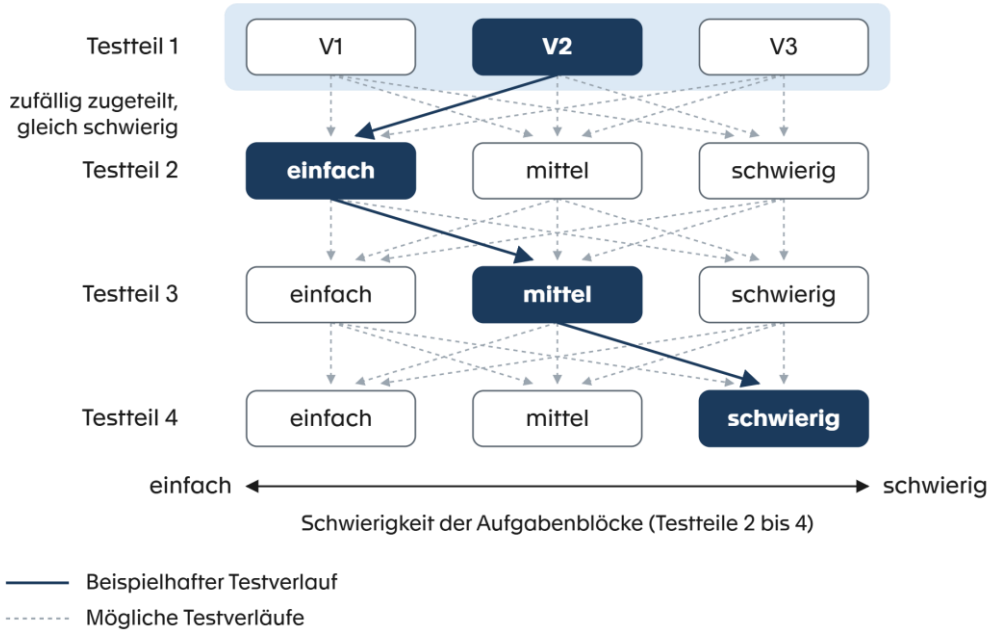
Fach	Kompetenzbereich	Anzahl Aufgaben	Testdauer
Deutsch	Lesen	32	40 min
	Sprache im Fokus	29–32*	20 min
Englisch	Hören	32	40 min
	Lesen	32	40 min
Mathematik	Zahl und Variable	16	
	Form und Raum	16	80 min
	Grössen, Funktionen, Daten und Zufall	16	

Die drei Kompetenzbereiche in Mathematik werden in einem gemeinsamen Test geprüft. In Deutsch und Englisch wird für jeden Kompetenzbereich ein eigener Test durchgeführt. Die Tests umfassen 29 bis 48 Aufgaben, die sich auf vier aufeinanderfolgende Testteile verteilen. Tabelle 1 zeigt die geprüften Kompetenzbereiche, die Anzahl Aufgaben und die Testdauer. Die Bearbeitungszeit ist für jeden Testteil begrenzt.

Alle Tests folgen einem adaptiven Testdesign. Dabei wird die Schwierigkeit der Aufgaben an die Fähigkeiten der Schülerinnen und Schüler angepasst. Dies ermöglicht eine möglichst präzise und verlässliche Messung der individuellen Fähigkeiten und reduziert das Risiko einer deutlichen Über- oder Unterforderung.

Abbildung 1 veranschaulicht das adaptive Testdesign. Im ersten Testteil wird den Schülerinnen und Schülern zufällig einer von drei gleich schwierigen Aufgabenblöcken (V1, V2 oder V3) zugewiesen. In den folgenden Testteilen wählt ein Algorithmus anhand des bisherigen Antwortverhaltens jeweils einen von drei unterschiedlich schwierigen Aufgabenblöcken (einfach, mittel oder schwierig) aus. Grundsätzlich bearbeiten alle Schülerinnen und Schüler gleich viele Aufgaben innerhalb derselben Testzeit. Eine Ausnahme bildet der Test Deutsch Sprache im Fokus: Hier enthält der einfachste Aufgabenblock jeweils eine Aufgabe weniger. Innerhalb eines Aufgabenblocks können Aufgaben übersprungen oder bereits gegebene Antworten überarbeitet werden.

Abbildung 1: Adaptives Testdesign



Anhand der Antworten werden mithilfe der Item-Response-Theorie die Fähigkeiten der Schülerinnen und Schüler geschätzt und auf der Check-Skala von 0 bis 1200 Punkten dargestellt. Die Skala bildet den Lernfortschritt von der 3. Klasse der Primarschule bis

zur 3. Klasse der Sekundarstufe I ab und umfasst damit den 2. und 3. Zyklus des Lehrplans 21. Die Ergebnisse des Check P4 sind grundsätzlich mit den Ergebnissen anderer Checks sowie mit früheren und künftigen Durchführungen vergleichbar.[1]

1.3 Teilnehmende Schülerinnen und Schüler

Insgesamt nahmen 1'217 Schülerinnen und Schüler aus 85 Klassen in 30 Schulen am Check P4 2026 teil. Je nach Test liegt die Anzahl teilnehmender Schülerinnen und Schüler zwischen 1'190 und 1'211. Über alle Tests betrachtet, reicht die Anzahl Teilnehmender pro Klasse von 2 bis 21.

Für ein aussagekräftiges Bildungsmonitoring soll der Check P4 im Kanton Zug möglichst flächendeckend in allen 4. Klassen der gemeindlichen Schulen durchgeführt werden. Um die Teilnahme beurteilen zu können, erhob das Amt für gemeindliche Schulen bei den Gemeinden zusätzlich die Anzahl der Schülerinnen und Schüler, die nicht am Check P4 teilnahmen. Diese Angaben wurden mit den im Check P4 verfügbaren Ergebnissen abgeglichen.

Insgesamt waren 1'269 Schülerinnen und Schüler für den Check P4 angemeldet. Tabelle 2 zeigt die Teilnahme pro Kompetenzbereich auf Grundlage der Check-Daten sowie der Angaben des Amts für gemeindliche Schulen. Die Ausschöpfung wurde jeweils auf Basis der 1'269 angemeldeten Schülerinnen und Schüler berechnet. Gemäss den Angaben des Amts nahmen je nach Test 37 bis 57 Schülerinnen und Schüler nicht teil. Dies entspricht einer Ausschöpfung von rund 96 Prozent. Auf Basis der im Check P4 verfügbaren Ergebnisse liegt die Ausschöpfung bei rund 95 Prozent. Wie bereits bei der Durchführung 2025 ist die Ausschöpfung auf Basis der Check-Daten in Deutsch Sprache im Fokus am niedrigsten (93.8%) und in Englisch Hören am höchsten (95.4%).

Tabelle 2: Teilnahme am Check P4 nach Kompetenzbereich und Datenquelle

		Daten Check P4 [a]				Daten Amt	
Fach	Kompetenzbereich	Teil.	NA	NT	Auss.	NT	Auss.
Deutsch	Lesen	1205	2	62	95.0%	57	95.5%
	Sprache im Fokus	1190	0	79	93.8%	56	95.6%
Englisch	Hören	1211	6	52	95.4%	42	96.7%
	Lesen	1199	1	69	94.5%	37	97.1%
Mathematik		1202	13	54	94.7%	49	96.1%

Anmerkungen: Teil.: Anzahl Teilnahmen; NA: Anzahl nicht abgeschlossener Tests; NT: Anzahl Nichtteilnahmen; Auss.: Ausschöpfung auf Basis von *N* = 1'269 angemeldeten Schülerinnen und Schülern.
[a] Zwei Schülerinnen und Schüler, die den Check P4 erst nach dem vorgesehenen Durchführungszeitraum absolvierten, werden als Nichtteilnahmen gezählt.

[1] Weiterführende Informationen sind unter www.ibe-edu.ch/methoden zu finden.

Zwischen den Angaben des Amts für gemeindliche Schulen und den im Check P4 verfügbaren Ergebnissen bestehen bei allen Tests Abweichungen. Diese lassen sich teilweise durch nicht abgeschlossene Tests erklären: Wird ein Test nicht abgeschlossen, liegt kein Testergebnis vor. Solche Fälle werden in den Check-Daten separat ausgewiesen, während sie von den Schulen möglicherweise als Teilnahmen gemeldet wurden.

Das Amt für gemeindliche Schulen erhob zudem die Gründe für die Nichtteilnahme. Tabelle 3 zeigt, dass bei allen Tests «andere Gründe» am häufigsten genannt wurden. Je nach Test entfallen zwischen 41 und 57 Prozent der Nichtteilnahmen auf diese Kategorie.

Tabelle 3: Gründe für die Nichtteilnahme am Check P4

Fach	Kompetenzbereich	NT	Überd. LZA	Vorüb. LZA	Krankheit	Andere Gründe
Deutsch	Lesen	57	8	15	9	25
	Sprache im Fokus	56	8	12	11	25
Englisch	Hören	42	5	4	11	22
	Lesen	37	5	4	7	21
Mathematik		49	8	9	12	20

Anmerkungen: NT: Anzahl Nichtteilnahmen; Überd. LZA: überdauernde Lernzielanpassung; Vorüb. LZA: vorübergehende Lernzielanpassung.

Bei der Anmeldung zum Check P4 wurden für die Schülerinnen und Schüler auf der Onlineplattform «Check Dein Wissen» das Geburtsdatum, das Geschlecht, die Erstsprache (Deutsch oder andere Erstsprache) sowie Fächer mit angepassten Lernzielen erfasst.

Die teilnehmenden Schülerinnen und Schüler sind im Durchschnitt 10.6 Jahre alt (SD = 0.41; Stichdatum: 31. März 2026). Der Anteil der Mädchen beträgt 53.0 Prozent, 62.6 Prozent der Schülerinnen und Schüler haben Deutsch als Erstsprache. Angepasste Lernziele wurden für 37 Schülerinnen und Schüler (3.0%) in Deutsch, für 3 (0.2%) in Englisch und für 24 (2.0%) in Mathematik registriert.

Die Anzahl der Schülerinnen und Schüler mit registrierten angepassten Lernzielen ist gegenüber 2025 in allen drei Fächern gesunken. Besonders deutlich ist die Abnahme in Deutsch von 7.8 auf 3.0 Prozent und in Mathematik von 6.9 auf 2.0 Prozent. Dies deutet auf eine veränderte Handhabung bei der Registrierung angepasster Lernziele hin. In Englisch lässt sich die Entwicklung aufgrund der sehr kleinen Fallzahlen kaum interpretieren.

Schülerinnen und Schüler mit angepassten Lernzielen nahmen zudem häufiger nicht am Check P4 teil. Eine weiterführende Analyse pro Test zeigt, dass in Deutsch Lesen 1 Prozent der Schülerinnen und Schüler mit regulären Lernzielen nicht teilnahm oder kein Testergebnis erhielt, gegenüber 30 Prozent der Schülerinnen und Schüler mit angepassten Lernzielen. Auch in Deutsch Sprache im Fokus (2 Prozent zu 35 Prozent) und Mathematik (1 Prozent zu 38 Prozent) zeigen sich deutliche Unterschiede. Schülerinnen

und Schüler mit angepassten Lernzielen sind in den verfügbaren Testergebnissen daher vermutlich unterrepräsentiert.

2 Ergebnisse pro Fach im Jahr 2026

Die Ergebnisse werden im Folgenden mehrheitlich pro Fach dargestellt. Dazu wurde für jede Schülerin und jeden Schüler der Mittelwert aus den Ergebnissen der Kompetenzbereiche eines Fachs gebildet. Ein Ergebnis pro Fach wurde nur berechnet, wenn für alle Kompetenzbereiche dieses Fachs ein Ergebnis vorliegt.

2.1 Mittelwert und Streuung der Ergebnisse

Tabelle 4 zeigt die durchschnittlichen Ergebnisse der Schülerinnen und Schüler pro Fach. Neben dem Mittelwert (M) ist die Standardabweichung (SD) ausgewiesen. Sie beschreibt die Streuung der Ergebnisse und kann vereinfacht als durchschnittliche Abweichung vom Mittelwert interpretiert werden.

Die Ergebnisse der verschiedenen Fächer sind nicht direkt miteinander vergleichbar, da die Tests keine gemeinsame Metrik haben. Vergleiche mit Check-Ergebnissen aus anderen Kantonen zeigen jedoch, dass die Schülerinnen und Schüler des Kantons Zug in Englisch vergleichsweise hohe Ergebnisse erzielen.[2] Dieses Muster zeigte sich bereits bei den beiden vorangehenden Durchführungen. Eine mögliche Erklärung ist der vergleichsweise hohe Anteil an Kindern internationaler Fach- und Führungskräfte (Expats), die zu Hause häufig Englisch sprechen.

Tabelle 4: Mittelwert und Streuung der Ergebnisse pro Fach

Fach	M	SD	N
Deutsch	532	128	1174
Englisch	649	119	1180
Mathematik	531	153	1192

2.2 Ergebnisse nach individuellen Merkmalen

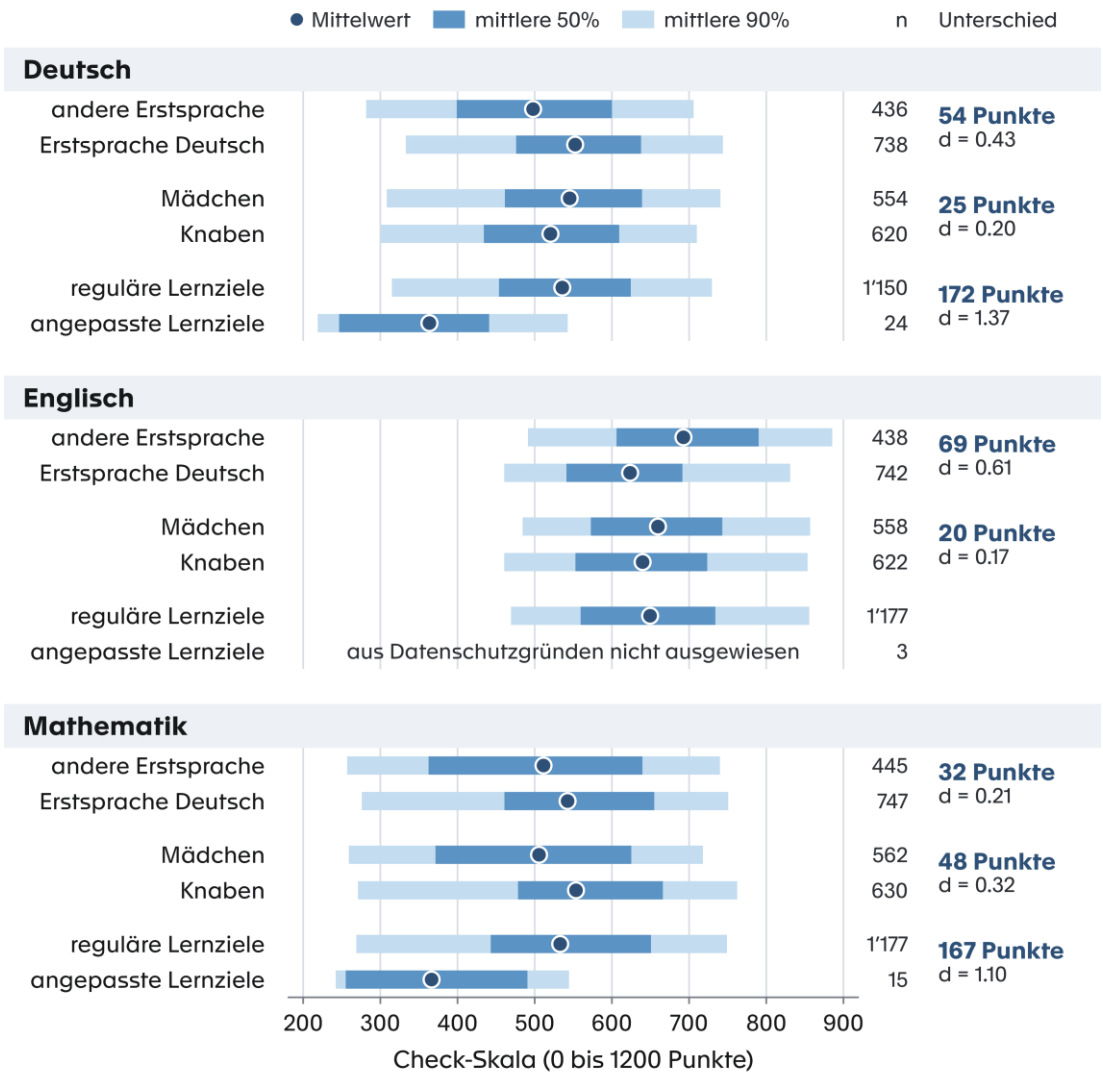
In Kombination mit den erfassten individuellen Merkmalen geben die Ergebnisse des Check P4 Aufschluss über Leistungsunterschiede zwischen verschiedenen Gruppen. Abbildung 2 zeigt die Mittelwerte (Punkte) sowie die mittleren 50 Prozent (dunkelblaue Balken) und die mittleren 90 Prozent (hellblaue Balken) der Ergebnisse pro Fach, jeweils getrennt nach Erstsprache, Geschlecht und Lernzielen. Zusätzlich sind die Anzahl Schülerinnen und Schüler (n), die Differenz der Mittelwerte in Punkten und die Effektstärke *d* ausgewiesen. Die Kennwerte zu den einzelnen Gruppen sind in Tabelle 5 aufgeführt.

[2] IBE (2025). Check P5 2025: Ergebnisbericht für den Bildungsraum Nordwestschweiz. https://bildungsraum-nw-api.webcloud7.ch/schwerpunkte/volksschule/ergebnisberichte_checks/1a49bd5c02ef4b4d98ad29675654e4fc/kantonsbericht-check-p5-2025.pdf

Da die Ergebnisse der verschiedenen Fächer nicht direkt vergleichbar sind, wird das Ausmass der Gruppenunterschiede zusätzlich anhand der Effektstärke *d* beurteilt. Sie gibt die Differenz der Mittelwerte im Verhältnis zur Streuung der Ergebnisse an. Nach Cohen gelten Effektstärken von *d* = 0.2 als klein, *d* = 0.5 als mittel und *d* = 0.8 als gross.[3]

Zur Einschätzung der praktischen Relevanz können die Differenzen in Punkten zudem mit dem erwarteten Lernfortschritt verglichen werden. Erfahrungswerte des Instituts für Bildungsevaluation zeigen für die Primarstufe einen durchschnittlichen Lernfortschritt pro Semester von rund 47 Punkten in Deutsch, 33 Punkten in Englisch und 50 Punkten in Mathematik.

Abbildung 2: Ergebnisse nach individuellen Merkmalen



Die Balken zeigen die mittleren 50 bzw. 90 Prozent der Ergebnisse einer Gruppe. Unterschied: Differenz der Mittelwerte in Punkten auf der Check-Skala; d: Effektstärke nach Cohen.

[3] Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.

Zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache und jenen mit einer anderen Erstsprache zeigen sich in allen drei Fächern Unterschiede. In Deutsch erzielen Schülerinnen und Schüler mit Deutsch als Erstsprache durchschnittlich 54 Punkte mehr. Dies entspricht einem kleinen bis mittleren Effekt (d = 0.43) und etwas mehr als dem erwarteten Lernfortschritt eines Semesters.

Tabelle 5: Ergebnisse pro Fach gruppiert nach individuellen Merkmalen

Fach	M	SD	n	M	SD	n
Erstsprache	Andere Erstsprache			Deutsch als Erstsprache		
Deutsch	498	129	436	552	123	738
Englisch	693	121	438	624	110	742
Mathematik	511	155	445	543	150	747
Geschlecht	Mädchen			Knaben		
Deutsch	545	129	554	520	126	620
Englisch	660	116	558	640	120	622
Mathematik	506	150	562	554	152	630
Lernziele	regulär			angepasst		
Deutsch	536	126	1150	364	118	24
Englisch	650	119	1177	n < 5 – nicht ausgewiesen		
Mathematik	533	152	1177	366	129	15

Anmerkung: Gruppen mit weniger als fünf Schülerinnen und Schülern werden aus Datenschutzgründen nicht ausgewiesen.

In Englisch zeigt sich der Unterschied in umgekehrter Richtung: Schülerinnen und Schüler mit einer anderen Erstsprache erzielen durchschnittlich 69 Punkte mehr als jene mit Deutsch als Erstsprache. Dies entspricht einem mittleren Effekt (d = 0.61) und etwa dem erwarteten Lernfortschritt eines Schuljahres. Dieses Muster zeigte sich bereits bei den beiden vorangehenden Durchführungen des Check P4 (vgl. Kapitel 3.2). Eine mögliche Erklärung ist, dass ein Teil der Schülerinnen und Schüler mit einer anderen Erstsprache Englisch als Erstsprache spricht.

Mit Blick auf andere Datenquellen zu schulischen Leistungen in der Schweiz ist dieses Ergebnismuster dennoch bemerkenswert. Die Überprüfung der Grundkompetenzen (ÜGK) 2017 zeigte für den Kanton Zug, dass Schülerinnen und Schüler im 8. Schuljahr, die zu Hause nicht Deutsch sprechen, in Englisch gleich gute oder etwas tiefere Ergebnisse erzielten als jene mit Deutsch als Erstsprache.[4] Die Ergebnisse sind allerdings nur eingeschränkt mit dem Check P4 vergleichbar. Zum einen unterscheiden sich die erfassten sprachlichen Merkmale: Im Check P4 wird die Erstsprache erfasst, in der ÜGK die zu Hause gesprochene Sprache. Zum anderen weist die ÜGK aus, ob die Grundkompetenzen erreicht werden, während der Check P4 Unterschiede auf einer kontinuierlichen Leistungsskala abbildet. Unterschiede zwischen Gruppen können bei der ÜGK daher weniger deutlich hervortreten, wenn ein grosser Anteil beider Gruppen die

[4] Konsortium ÜGK (2019). *Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2017: Sprachen 8. Schuljahr.* EDK & SRED. https://doi.org/10.18747/PHSG-coll3/id/385

Grundkompetenzen erreicht. Darüber hinaus ist denkbar, dass sich der im Check P4 festgestellte Unterschied im Verlauf der Schullaufbahn verändert.

In Mathematik erzielen Schülerinnen und Schüler mit Deutsch als Erstsprache durchschnittlich 32 Punkte mehr als jene mit einer anderen Erstsprache. Dies entspricht einem kleinen Effekt (d = 0.21). Die Richtung des Unterschieds stimmt mit den Ergebnissen der ÜGK 2016 überein: Im 11. Schuljahr erreichten Schülerinnen und Schüler, die zu Hause Deutsch sprechen, die Grundkompetenzen in Mathematik häufiger als jene, die zu Hause eine andere Sprache sprechen.[5]

Auch nach Geschlecht zeigen sich Unterschiede in den Ergebnissen des Check P4. In Deutsch erzielen Mädchen durchschnittlich 25 Punkte mehr als Knaben. Dies entspricht einem kleinen Effekt (d = 0.20) und etwa dem erwarteten Lernfortschritt eines halben Semesters. In Englisch beträgt der Vorsprung der Mädchen 20 Punkte. Der Unterschied ist mit einer Effektstärke von d = 0.17 gering. In Mathematik zeigt sich ein umgekehrtes Muster: Knaben erzielen durchschnittlich 48 Punkte mehr als Mädchen. Dies entspricht einem kleinen Effekt (d = 0.32) und ungefähr dem erwarteten Lernfortschritt eines Semesters.

Ein Leistungsvorsprung von Knaben gegenüber Mädchen in Mathematik wird in der Schweiz regelmässig festgestellt. In der ÜGK 2016 erreichten Knaben im 11. Schuljahr die Grundkompetenzen in Mathematik häufiger als Mädchen.[5] Auch in PISA 2022 erzielten Knaben in der Schweiz im Durchschnitt höhere Mathematikleistungen als Mädchen.[6] Der Vergleich mit dem Check P5 2025 im Bildungsraum Nordwestschweiz zeigt, dass der Geschlechterunterschied dort gegen Ende der 5. Klasse mit 21 Punkten kleiner ausfiel als gegen Ende der 4. Klasse im Kanton Zug mit 48 Punkten.[7]

Schülerinnen und Schüler mit angepassten Lernzielen erzielen in Deutsch und Mathematik deutlich tiefere Ergebnisse als Schülerinnen und Schüler mit regulären Lernzielen. In Deutsch beträgt die Differenz 172 Punkte (d = 1.37), in Mathematik 167 Punkte (d = 1.10). Die Unterschiede entsprechen damit in beiden Fächern grossen Effekten. In Englisch liegen lediglich für drei Schülerinnen und Schüler mit angepassten Lernzielen Ergebnisse vor. Aufgrund der kleinen Fallzahl werden diese Ergebnisse aus Datenschutzgründen nicht ausgewiesen.

Die Ergebnisse zu den angepassten Lernzielen sind insgesamt vorsichtig zu interpretieren. Auch in Deutsch und Mathematik sind die Gruppen mit angepassten Lernzielen klein. Zudem zeigt Kapitel 1.3, dass Schülerinnen und Schüler mit angepassten Lernzielen häufiger nicht am Check P4 teilnahmen oder keinen vollständigen Test abschlossen. Die verfügbaren Ergebnisse bilden diese Gruppe daher möglicherweise nicht vollständig ab.

---

[5] Konsortium ÜGK (2019). *Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2016: Mathematik 11. Schuljahr.* EDK & SRED. https://doi.org/10.18747/PHSG-coll3/id/386

[6] Erzinger, A. B., Pham, G., Prosperi, O., & Salvisberg, M. (2023). PISA 2022: Die Schweiz im Fokus. Universität Bern. https://www.pisa-schweiz.ch/PISA2022-DieSchweizimFokus.pdf

[7] IBE (2025). Check P5 2025: Ergebnisbericht für den Bildungsraum Nordwestschweiz. https://bildungsraum-nw-api.webcloud7.ch/schwerpunkte/volksschule/ergebnisberichte_checks/1a49bd5c02ef4b4d98ad29675654e4fc/kantonsbericht-check-p5-2025.pdf

2.3 Zusammenhang zwischen den Fächern

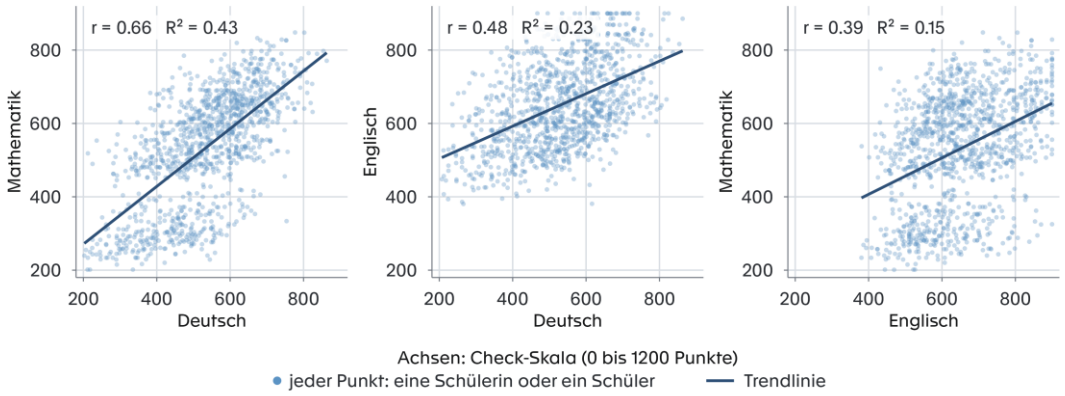
Abbildung 3 zeigt die Zusammenhänge zwischen den Ergebnissen in Deutsch, Englisch und Mathematik. Die blauen Punkte veranschaulichen, welche Kombinationen von Ergebnissen in jeweils zwei Fächern vorkommen. Je mehr Punkte in einem Bereich liegen, desto häufiger tritt die entsprechende Kombination auf. Zusätzlich sind die Korrelation (r) und der Determinationskoeffizient (R^2) ausgewiesen.

Die Korrelation (r) beschreibt die Stärke und Richtung des linearen Zusammenhangs zwischen den Ergebnissen in zwei Fächern. Sie kann Werte zwischen –1 und +1 annehmen. Ein Wert nahe 0 weist auf einen geringen linearen Zusammenhang hin, Werte nahe –1 oder +1 auf einen starken Zusammenhang. Der Determinationskoeffizient (R^2) gibt an, welcher Anteil der Unterschiede in einem Fach durch die Ergebnisse im jeweils anderen Fach erklärt werden kann. Er kann als Prozentwert ausgedrückt werden. Für Deutsch und Englisch beträgt R^2 beispielsweise 0.23. Das bedeutet, dass 23 Prozent der Unterschiede in den Englisch-Ergebnissen statistisch durch die Deutsch-Ergebnisse erklärt werden können – und umgekehrt.

Der stärkste Zusammenhang zeigt sich zwischen den Ergebnissen in Deutsch und Mathematik (R^2 = 0.43). Schülerinnen und Schüler mit höheren Ergebnissen in Deutsch erzielen somit tendenziell auch höhere Ergebnisse in Mathematik. Ein möglicher Grund für den vergleichsweise starken Zusammenhang ist die Bedeutung sprachlicher Kompetenzen für das Bearbeiten von Mathematikaufgaben. Deutsch ist nicht nur Unterrichtssprache, sondern auch die Sprache, in der die Mathematikaufgaben formuliert sind. Insbesondere bei Aufgaben in einem sprachlichen Kontext können gute Sprachkompetenzen das Verständnis der Aufgabenstellung unterstützen.

Die Zusammenhänge zwischen Deutsch und Englisch sowie zwischen Englisch und Mathematik sind deutlich schwächer. Der Zusammenhang zwischen Deutsch und Englisch könnte teilweise auf allgemeine sprachliche Fähigkeiten zurückzuführen sein. Insgesamt zeigen die Ergebnisse, dass Leistungen in verschiedenen Schulfächern nicht unabhängig voneinander sind. Aus den dargestellten Zusammenhängen lässt sich jedoch nicht ableiten, dass die Leistungen in einem Fach die Leistungen in einem anderen Fach verursachen.

Abbildung 3: Zusammenhang zwischen den Ergebnissen der drei Fächer



2.4 Ergebnisse nach Kompetenzstufen

Auf der Check-Skala beschreiben Kompetenzstufen, welche Aufgaben Schülerinnen und Schüler in der Regel sicher lösen können beziehungsweise welche Kompetenzen sie beherrschen.[8] Mit dem Check P4 werden die Kompetenzstufen II bis IX erfasst. Schülerinnen und Schüler einer bestimmten Kompetenzstufe beherrschen in der Regel auch die Kompetenzen der darunterliegenden Stufen. Kompetenzen höherer Stufen werden hingegen noch nicht oder nur teilweise beherrscht.

Die im Lehrplan 21 definierten Orientierungspunkte und Grundansprüche lassen sich den Kompetenzstufen des Check P4 zuordnen. Für die 4. Klasse ist insbesondere der Orientierungspunkt in der Mitte des 2. Zyklus relevant. Er bezeichnet die Kompetenzen, die bis zum Ende der 4. Klasse bearbeitet werden sollen. Schülerinnen und Schüler, die im Check P4 mindestens Kompetenzstufe V erreichen, beherrschen die bis zum Orientierungspunkt vorgesehenen Kompetenzen in der Regel.

Die ebenfalls im Lehrplan 21 festgelegten Grundansprüche definieren Kompetenzen, die die Schülerinnen und Schüler bis zum Ende des 2. Zyklus erreichen sollen. Im Gegensatz dazu handelt es sich beim Orientierungspunkt um einen Anhaltspunkt für die Unterrichtsplanung der Lehrpersonen. Es ist zudem zu beachten, dass der Check P4 im März und somit vor Ende der 4. Klasse stattfand. Deshalb ist es möglich, dass einzelne bis zum Orientierungspunkt vorgesehene Inhalte zum Zeitpunkt des Check P4 noch nicht behandelt wurden.

Tabelle 6 zeigt die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen II bis IX. Zusätzlich ist der Anteil ausgewiesen, der mindestens Kompetenzstufe V erreicht und damit die bis zum Orientierungspunkt vorgesehenen Kompetenzen in der Regel beherrscht. Die Entwicklung dieses Anteils über die drei bisherigen Durchführungen wird in Kapitel 3.3 dargestellt.

Tabelle 6: Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Fach	Kompetenzbereich	Kompetenzstufen (%)								
		II	III	IV	V	VI	VII	VIII	IX	**≥ V**
Deutsch	Lesen	7.8	12.5	17.9	30.5	21.7	7.5	2.2		**61.9**
	Sprache im Fokus	4.0	14.3	21.8	24.3	25.4	8.6	1.5		**59.8**
Englisch	Hören		1.0	8.1	26.5	28.9	22.6	8.6	4.2	**90.8**
	Lesen		0.9	13.4	24.9	29.6	16.7	11.0	3.5	**85.7**
Mathematik	Zahl und Variable	13.8	9.8	15.4	24.9	27.9	6.0	2.3		**61.1**
	Form und Raum	6.1	14.0	21.1	16.3	24.9	11.4	5.3	0.9	**58.8**
	Grössen, Funktionen, Daten und Zufall	14.2	10.5	10.5	28.8	21.5	11.3	3.0	0.3	**64.9**

Anmerkung: Ab Kompetenzstufe V werden die bis zum Orientierungspunkt in der Mitte des 2. Zyklus vorgesehenen Kompetenzen in der Regel beherrscht.

[8] Kompetenzbeschreibungen unter https://support.ibe-edu.ch/kompetenzbeschreibungen-lernende.

In Deutsch und Mathematik erreichen je nach Kompetenzbereich rund 59 bis 65 Prozent der Schülerinnen und Schüler mindestens Kompetenzstufe V. In Englisch sind die Anteile deutlich höher: In Englisch Hören erreichen 90.8 Prozent und in Englisch Lesen 85.7 Prozent mindestens Kompetenzstufe V. Damit beherrscht in Englisch ein grosser Teil der Schülerinnen und Schüler die bis zum Orientierungspunkt vorgesehenen Kompetenzen. In Deutsch und Mathematik trifft dies auf rund drei Fünftel der Schülerinnen und Schüler zu.

In Deutsch Lesen erreichen 61.9 Prozent der Schülerinnen und Schüler mindestens Kompetenzstufe V. Auf dieser Stufe können sie verschiedene einfache Texte verstehen, deren Themen und Wortschatz ihrem Weltwissen entsprechen, beispielsweise kurze Sachtexte, kindgerechte Reportagen oder Ausschnitte aus Kinderbüchern. Kompetenzen höherer Stufen, wie das Entnehmen von Hauptaussagen und Detailinformationen aus anspruchsvolleren Texten, werden noch nicht oder nur teilweise beherrscht.

In Deutsch Sprache im Fokus erreichen 59.8 Prozent mindestens Kompetenzstufe V. Dazu gehören unter anderem das Bestimmen typischer Nomen, Verben und Adjektive, das Anwenden grundlegender Rechtschreibregeln sowie das Setzen von Kommas in Aufzählungen. Kompetenzen höherer Stufen, etwa das sichere Bestimmen von Zeitformen oder das Setzen von Kommas zwischen Verbgruppen, werden hingegen noch nicht oder nur teilweise beherrscht.

In Englisch erreichen besonders viele Schülerinnen und Schüler mindestens Kompetenzstufe V: 90.8 Prozent in Hören und 85.7 Prozent in Lesen. In Englisch Hören können Schülerinnen und Schüler auf Kompetenzstufe V beispielsweise kurze Gespräche über vertraute Themen verstehen und grundlegende Informationen aus einfachen Hörtexten entnehmen. Längere Gespräche und Hörtexte werden hingegen noch nicht oder nur teilweise verstanden.

In Englisch Lesen können Schülerinnen und Schüler auf Kompetenzstufe V einfache, kurze Texte und Informationsmaterialien (z. B. Comics, SMS oder Rezepte) verstehen und grundlegende Aussagen daraus entnehmen. Authentische Texte (z. B. Reportagen) und persönliche Texte (z. B. Briefe oder Blogs) werden hingegen noch nicht oder nur teilweise verstanden, selbst wenn sie klar strukturiert sind.

In Mathematik erreichen in allen drei Kompetenzbereichen rund 59 bis 65 Prozent der Schülerinnen und Schüler mindestens Kompetenzstufe V. Im Kompetenzbereich Zahl und Variable sind es rund 61 Prozent. Schülerinnen und Schüler auf Kompetenzstufe V beherrschen unter anderem das schriftliche Addieren und Subtrahieren, können Rechenwege für Addition und Subtraktion notieren und Zahlen bis 1 Million ordnen. Noch nicht oder nicht sicher können sie beispielsweise Zahlen mit vier Wertziffern addieren und subtrahieren (z. B. 183 + 2401) oder Dezimalzahlen ordnen, addieren und subtrahieren.

Im Kompetenzbereich Form und Raum erreichen rund 59 Prozent der Schülerinnen und Schüler mindestens Kompetenzstufe V. Auf dieser Stufe können sie unter anderem Symmetrieachsen erkennen und Spiegelbilder von Figuren skizzieren. Noch nicht oder nicht sicher können sie mehrere Symmetrieachsen in einer Figur erkennen oder die Fläche von Rechtecken berechnen.

Im Kompetenzbereich Grössen, Funktionen, Daten und Zufall erreichen rund 65 Prozent der Schülerinnen und Schüler mindestens Kompetenzstufe V. Auf dieser Stufe können

sie Grössen in andere Masseinheiten umwandeln, damit einfache Rechnungen durchführen und einfache Textaufgaben mit Grössen lösen. Noch nicht oder nicht sicher können sie proportionale Beziehungen anwenden (z. B. Fr./kg oder l/km) oder Funktionswerte aus Graphen ablesen.

Zusammenfassend ist festzuhalten, dass in Deutsch und Mathematik etwa 60 Prozent der Schülerinnen und Schüler die Kompetenzen bis zum Orientierungspunkt des 2. Zyklus beherrschen. In Englisch liegt dieser Anteil mit über 85 Prozent und damit deutlich höher. Dies zeigt erneut, dass das Fach Englisch eine besondere Stärke der Schülerinnen und Schüler im Kanton Zug ist.

2.5 Unterschiede zwischen Klassen

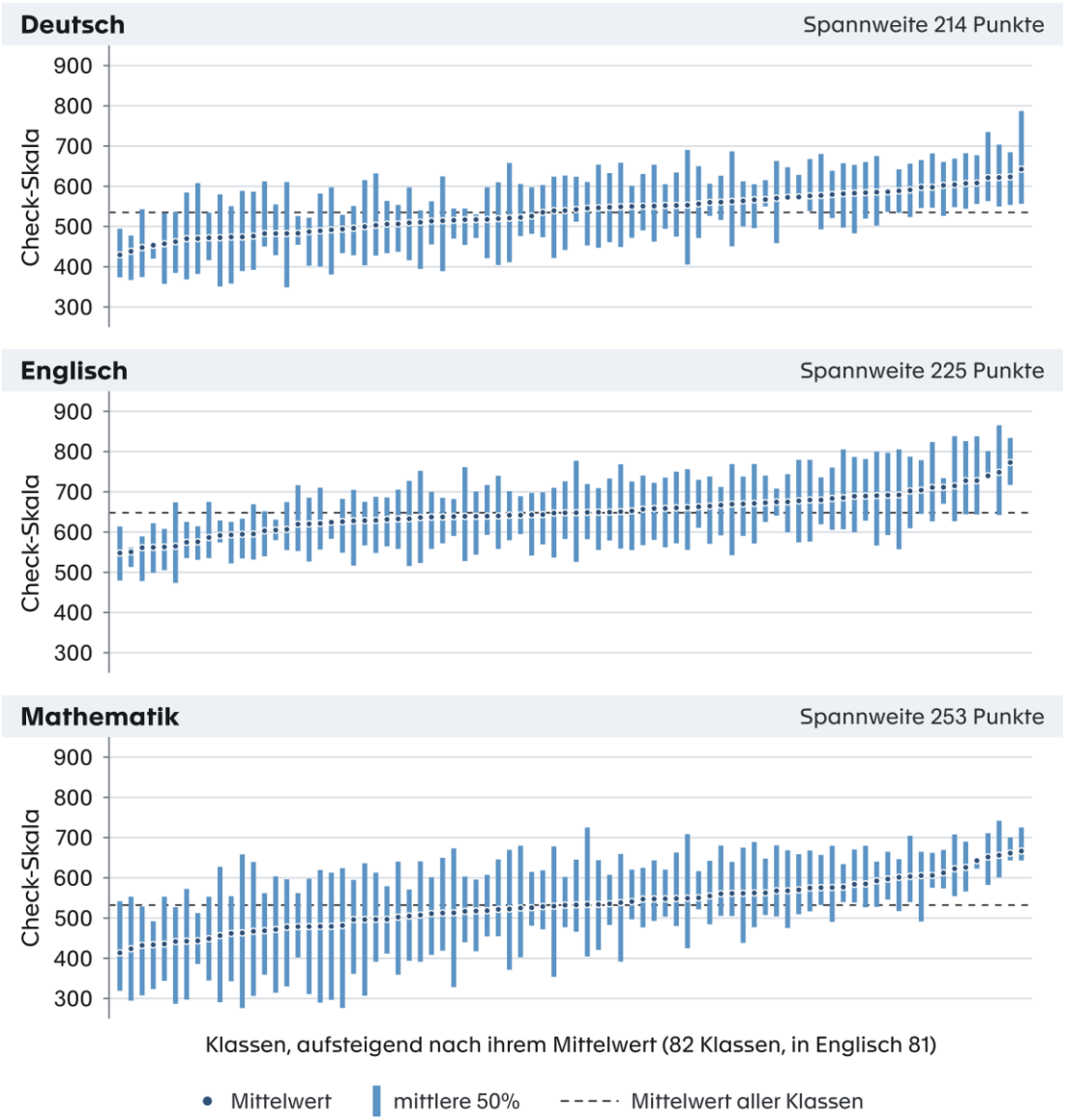
Abbildung 4 zeigt die Ergebnisse der teilnehmenden Klassen. Für jede Klasse wurde der Mittelwert der Ergebnisse der Schülerinnen und Schüler mit regulären Lernzielen berechnet. Zudem wurde der Bereich, in dem die mittleren 50 Prozent der Ergebnisse der Schülerinnen und Schüler liegen, abgebildet. Klassen mit weniger als fünf Ergebnissen wurden nicht dargestellt. Dadurch wurden drei Klassen in Deutsch und Mathematik sowie vier Klassen in Englisch von der Auswertung ausgeschlossen. Für die verbleibenden 82 beziehungsweise 81 Klassen liegen jeweils Ergebnisse von 5 bis 21 Schülerinnen und Schülern vor; durchschnittlich sind es über alle Fächer hinweg 14.3 Ergebnisse pro Klasse.

Die Unterschiede zwischen den Klassen fallen deutlich aus. Die Differenz zwischen dem niedrigsten und dem höchsten Klassenmittelwert beträgt in Deutsch 214 Punkte, in Englisch 225 Punkte und in Mathematik 253 Punkte. Gleichzeitig streuen die Ergebnisse innerhalb der einzelnen Klassen stark. Die dargestellten Bereiche überlappen sich zwischen vielen Klassen deutlich. Auch Klassen mit unterschiedlichen Mittelwerten weisen somit häufig Schülerinnen und Schüler mit ähnlichen Ergebnissen auf.

Um die Unterschiede zwischen den Klassen zu quantifizieren, wurde berechnet, welcher Anteil der Streuung der individuellen Ergebnisse auf Unterschiede zwischen den Klassen entfällt. Je höher dieser Anteil ist, desto stärker unterscheiden sich die durchschnittlichen Ergebnisse zwischen den Klassen. Der Varianzanteil der Klassen beträgt 14.4 Prozent in Deutsch, 15.0 Prozent in Englisch und 14.1 Prozent in Mathematik. Die Werte sind damit in allen drei Fächern ähnlich hoch. Rund 14 bis 15 Prozent der Unterschiede in den individuellen Ergebnissen entfallen auf Unterschiede zwischen den Klassen; die übrigen Unterschiede bestehen innerhalb der Klassen. Die Werte liegen in Englisch nahe bei jenen des Vorjahres, in Deutsch und Mathematik sind sie leicht gestiegen. Im Jahr 2025 betrug der Varianzanteil der Klassen 12.3 Prozent in Deutsch, 15.7 Prozent in Englisch und 11.9 Prozent in Mathematik.

Die Unterschiede zwischen den Klassen unterstreichen die Bedeutung einer vom Klassenkontext unabhängigen Standortbestimmung. Der Check P4 ergänzt damit die stärker vom jeweiligen Klassenkontext geprägten Schulnoten um eine standardisierte, klassenübergreifend vergleichbare Leistungsmessung. Offen bleibt, in welchem Ausmass die beobachteten Unterschiede mit individuellen Merkmalen der Schülerinnen und Schüler oder mit Merkmalen der Klassen, der Lehrpersonen und der Schulen zusammenhängen. Zentrale Lernvoraussetzungen wie die soziale Zusammensetzung der Klassen konnten in den Analysen nicht berücksichtigt werden.

Abbildung 4: Ergebnisse nach Klassen

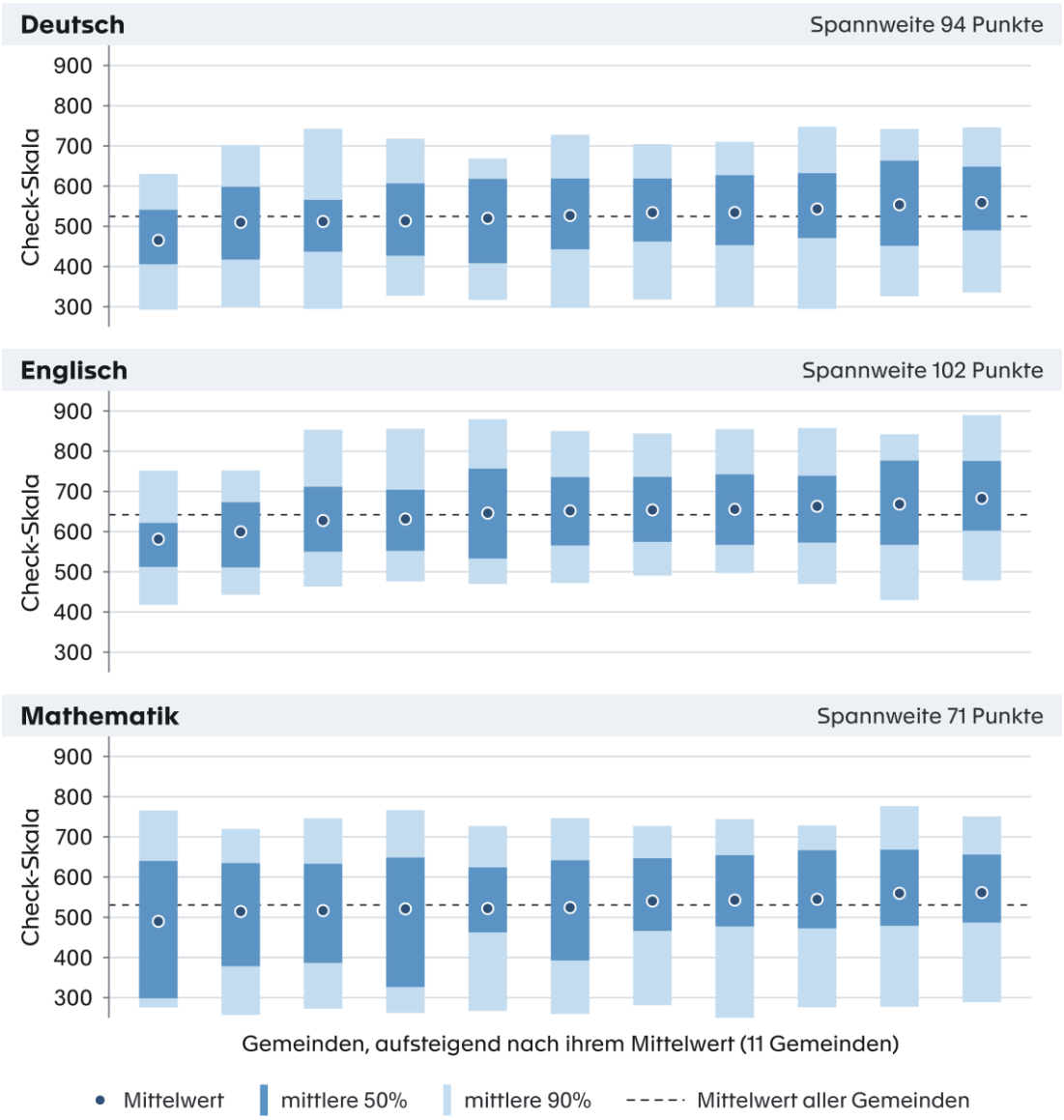


2.6 Unterschiede zwischen Gemeinden

Abbildung 5 zeigt die Ergebnisse der Schülerinnen und Schüler nach Gemeinde. Dabei wurden auch Schülerinnen und Schüler mit angepassten Lernzielen berücksichtigt. Für jede Gemeinde ist der Mittelwert dargestellt. Zusätzlich zeigen die dunkel- und hellblauen Balken den Bereich, in dem die mittleren 50 beziehungsweise 90 Prozent der Ergebnisse innerhalb der Gemeinde liegen.

Die durchschnittlichen Ergebnisse unterscheiden sich zwischen den Gemeinden. Die Differenz zwischen dem niedrigsten und dem höchsten Gemeindemittelwert beträgt in Deutsch 94 Punkte, in Englisch 102 Punkte und in Mathematik 71 Punkte. Gleichzeitig streuen die Ergebnisse innerhalb der einzelnen Gemeinden stark und die dargestellten Bereiche überlappen sich deutlich. Schülerinnen und Schüler mit ähnlichen Ergebnissen finden sich somit auch in Gemeinden mit unterschiedlichen Mittelwerten.

Abbildung 5: Ergebnisse nach Gemeinden



Um die Unterschiede zwischen den Gemeinden zu quantifizieren, wurde analog zum Vergleich der Klassen berechnet, welcher Anteil der Streuung der individuellen Ergebnisse auf Unterschiede zwischen den Gemeinden entfällt. Der Varianzanteil der Gemeinden beträgt 2.0 Prozent in Deutsch, 3.3 Prozent in Englisch und 1.5 Prozent in Mathematik. Damit entfällt in allen drei Fächern nur ein kleiner Teil der Unterschiede in den individuellen Ergebnissen auf Unterschiede zwischen den Gemeinden. Die Unterschiede zwischen den Gemeinden sind deutlich geringer als jene zwischen den Klassen.

Auch im Vorjahr waren die Unterschiede zwischen den Gemeinden gering. Die Varianzanteile betrugen 1.7 Prozent in Deutsch, 4.4 Prozent in Englisch und 1.9 Prozent in Mathematik. Insgesamt zeigt sich damit in beiden Jahren, dass die Gemeindezugehörigkeit nur in geringem Ausmass mit den Ergebnissen im Check P4 zusammenhängt.

Offen bleibt, in welchem Ausmass die beobachteten Unterschiede mit der Zusammensetzung der Schülerschaft oder mit Merkmalen der Schulen und Gemeinden zusammenhängen. Entsprechende Merkmale wurden in den vorliegenden Analysen nicht berücksichtigt.

3 Ergebnisse pro Fach im Trend

Der Check P4 wurde im Kanton Zug 2024 erstmals als Pilot durchgeführt. Seit 2025 findet er jährlich statt. Damit liegen für den Check P4 inzwischen Ergebnisse aus drei Durchführungen vor. Die Ergebnisse aller Checks werden auf derselben Check-Skala ausgewiesen und können damit über die drei Durchführungen hinweg verglichen werden. Eine Punktzahl – beispielsweise 550 Punkte in Deutsch – hat somit in allen Jahren die gleiche Bedeutung. Dadurch lassen sich sämtliche Ergebnisse mit dem Vorjahr vergleichen. Neben der Entwicklung der durchschnittlichen Ergebnisse werden im Folgenden auch Veränderungen nach ausgewählten Merkmalen sowie bei der Verteilung auf die Kompetenzstufen betrachtet.

Dabei ist zu berücksichtigen, dass die Pilotdurchführung 2024 im Mai stattfand, während die regulären Durchführungen seit 2025 im März stattfinden. Die Schülerinnen und Schüler profitierten bei der Pilotdurchführung somit von rund zwei zusätzlichen Unterrichtsmonaten. Dieser Unterschied im Durchführungszeitpunkt ist bei der Interpretation der zeitlichen Entwicklung zu berücksichtigen.

Wie bereits für die Gruppenvergleiche (Kapitel 2.2) gilt auch für die Ergebnisse im Trend, dass nicht nur die Differenz in Punkten, sondern auch die standardisierten Mittelwertdifferenzen (Effektstärken *d*) berücksichtigt werden sollten. Zur Einschätzung der Relevanz dient zudem der erwartete mittlere Lernfortschritt pro Semester von rund 47 Punkten in Deutsch, 33 Punkten in Englisch und 50 Punkten in Mathematik.

3.1 Trend seit 2024

Tabelle 7 zeigt die Mittelwerte, Differenzen in Punkten sowie die Effektstärken *d* im Vergleich der jeweils aufeinanderfolgenden Jahre. Die Verteilung der Ergebnisse ist anhand des Mittelwerts sowie der mittleren 50 und 90 Prozent zusätzlich in Abbildung 6 dargestellt.

Tabelle 7: Mittelwerte und Standardabweichungen der Ergebnisse der Jahre 2024 bis 2026 im Vergleich

Jahr	Deutsch		Englisch		Mathematik	
	M	SD	M	SD	M	SD
2024	550	122	622	117	536	141
2025	531	126	645	121	523	153
2026	532	128	649	119	531	153

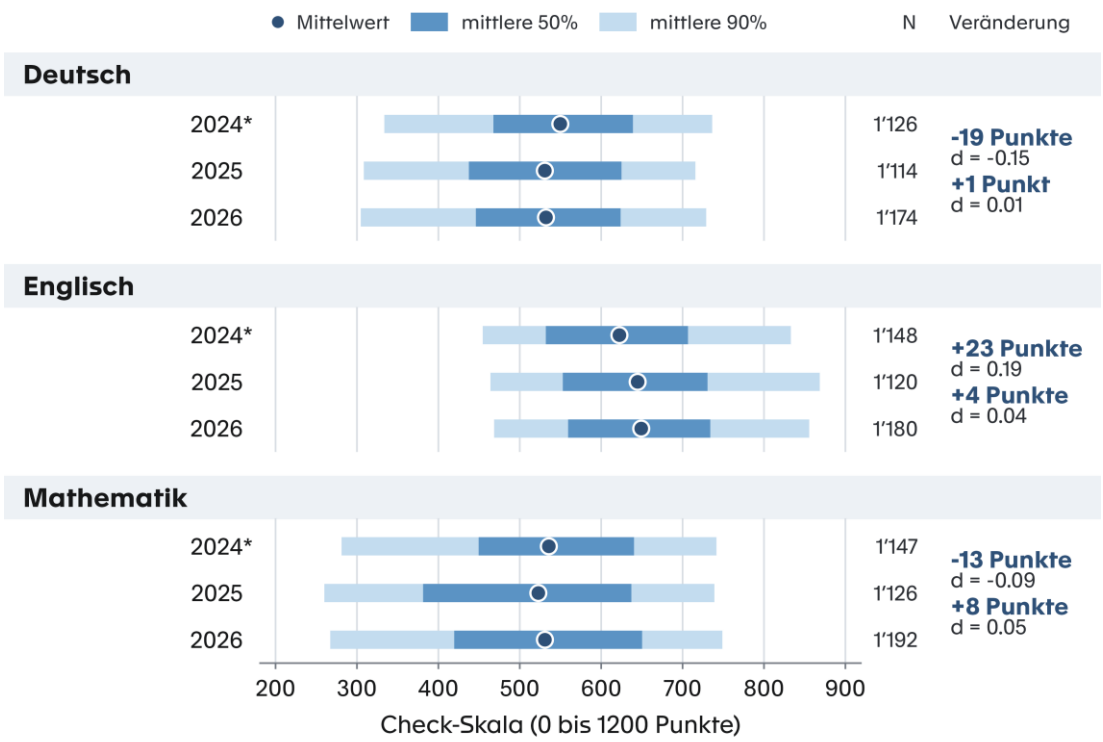
Anmerkung: Die Ergebnisse in Englisch für das Jahr 2025 wurden für den vorliegenden Bericht auf Grundlage der Skalierung von 2026 neu berechnet. Dadurch können sie von den im Kantonsbericht 2025 ausgewiesenen Werten abweichen. Die Werte der übrigen Fächer bleiben unverändert.

Die Mittelwerte sind im Vergleich zum Vorjahr in allen drei Fächern weitgehend stabil geblieben. In Deutsch liegt der Mittelwert 2026 um 1 Punkt höher als 2025 (d = 0.01), in Englisch um 4 Punkte (d = 0.04) und in Mathematik um 8 Punkte (d = 0.05). Die Unterschiede sind damit in allen drei Fächern sehr gering.

Von der Pilotdurchführung 2024 zur ersten regulären Durchführung 2025 lagen die mittleren Ergebnisse in Deutsch und Mathematik tiefer. Eine mögliche Erklärung ist der unterschiedliche Durchführungszeitpunkt: Die Pilotdurchführung 2024 fand im Mai statt, die regulären Durchführungen 2025 und 2026 hingegen im März. Den Schülerinnen und Schülern standen 2024 somit rund zwei zusätzliche Unterrichtsmonate zur Verfügung. Die weitgehend stabilen Ergebnisse von 2025 zu 2026 sprechen dafür, dass der Rückgang von 2024 auf 2025 zumindest teilweise mit dem unterschiedlichen Durchführungszeitpunkt zusammenhängen könnte.

In Englisch zeigt sich ein anderes Muster. Der Mittelwert stieg von 622 Punkten im Jahr 2024 auf 645 Punkte im Jahr 2025 (d = 0.19) und blieb 2026 mit 649 Punkten weitgehend stabil. Der Unterschied zwischen 2024 und 2025 entspricht einem kleinen Effekt.

Abbildung 6: Ergebnisse der Jahre 2024 bis 2026 im Vergleich

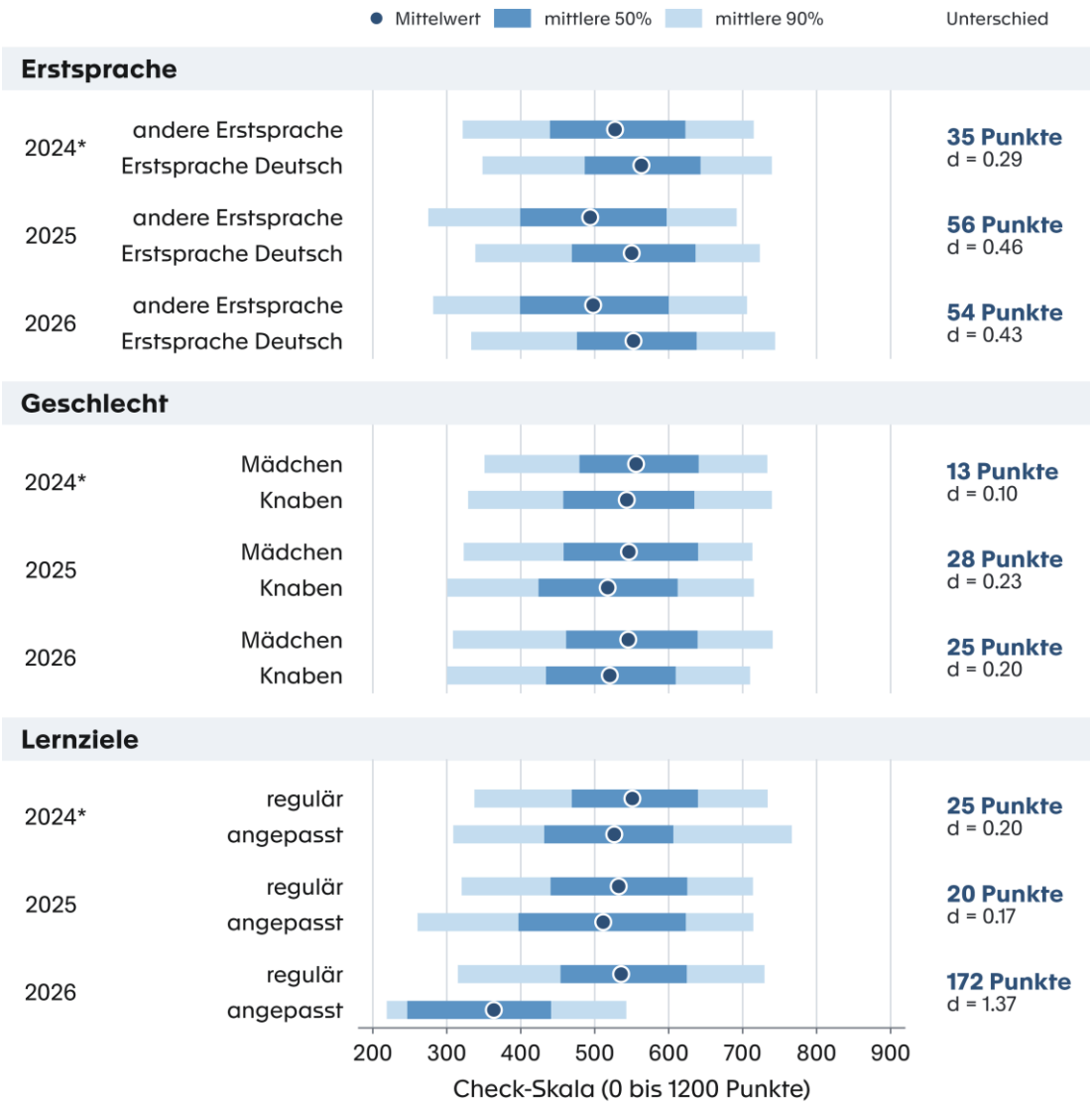


Die Balken zeigen die mittleren 50 bzw. 90 Prozent der Ergebnisse. Veränderung: Differenz der Mittelwerte gegenüber dem Vorjahr in Punkten auf der Check-Skala; d = Effektstärke nach Cohen (0.2 klein, 0.5 mittel, 0.8 gross). * 2024: Pilotdurchführung im Mai, zwei Monate später als ab 2025.

3.2 Trends nach individuellen Merkmalen

Im Folgenden wird untersucht, wie sich die Ergebnisse verschiedener Gruppen von Schülerinnen und Schülern seit 2024 entwickelt haben. Im Zentrum steht dabei die Frage, ob sich die in Kapitel 2.2 beschriebenen Unterschiede nach Erstsprache, Geschlecht und angepassten Lernzielen über die drei Durchführungen verändert haben. Die Ergebnisse werden getrennt nach Fach dargestellt und anhand der Mittelwerte sowie der Unterschiede und Effektstärken zwischen den Gruppen verglichen (Abbildungen 7 bis 9). Zur Einordnung der Unterschiede kann zusätzlich der erwartete Lernfortschritt herangezogen werden. Auf der Check-Skala entsprechen in Deutsch rund 47 Punkten, in Englisch 33 Punkten und in Mathematik 50 Punkten dem durchschnittlich erwarteten Lernfortschritt eines Semesters.

Abbildung 7: Ergebnisse in Deutsch nach individuellen Merkmalen und Jahr



Die Balken zeigen die mittleren 50 bzw. 90 Prozent der Ergebnisse. Unterschied: Differenz der Mittelwerte beider Gruppen in Punkten auf der Check-Skala; d = Effektstärke nach Cohen (0.2 klein, 0.5 mittel, 0.8 gross). * 2024: Pilotdurchführung im Mai, zwei Monate später als ab 2025.

Im Fach Deutsch (Abbildung 7) zeigt sich bei der Erstsprache eine ähnliche Entwicklung wie im Vorjahr. Schülerinnen und Schüler mit Deutsch als Erstsprache erzielen in allen drei Durchführungen höhere Ergebnisse als Schülerinnen und Schüler mit einer anderen Erstsprache. Der Unterschied nahm von 35 Punkten im Jahr 2024 auf 56 Punkte im Jahr 2025 zu und liegt 2026 mit 54 Punkten (d = 0.43) auf einem vergleichbaren Niveau wie im Vorjahr. Dies entspricht etwas mehr als dem erwarteten Lernfortschritt eines Semesters. Die im letzten Jahr beobachtete Zunahme des Unterschieds setzt sich damit 2026 nicht fort.

Auch Mädchen erzielen in Deutsch in allen drei Jahren höhere Ergebnisse als Knaben. Der Unterschied nahm von 13 Punkten im Jahr 2024 auf 28 Punkte im Jahr 2025 zu. Im Jahr 2026 beträgt er 25 Punkte (d = 0.20) und ist damit gegenüber dem Vorjahr weitgehend stabil geblieben. Auch hier setzt sich die von 2024 auf 2025 beobachtete Zunahme somit nicht weiter fort.

Deutlich anders fällt die Entwicklung bei den angepassten Lernzielen aus. Schülerinnen und Schüler mit angepassten Lernzielen erzielten 2024 und 2025 im Mittel nur etwas tiefere Ergebnisse als Schülerinnen und Schüler mit regulären Lernzielen. Die Differenz betrug 25 beziehungsweise 20 Punkte. Im Jahr 2026 ist der Unterschied mit 172 Punkten (d = 1.37) dagegen sehr gross.

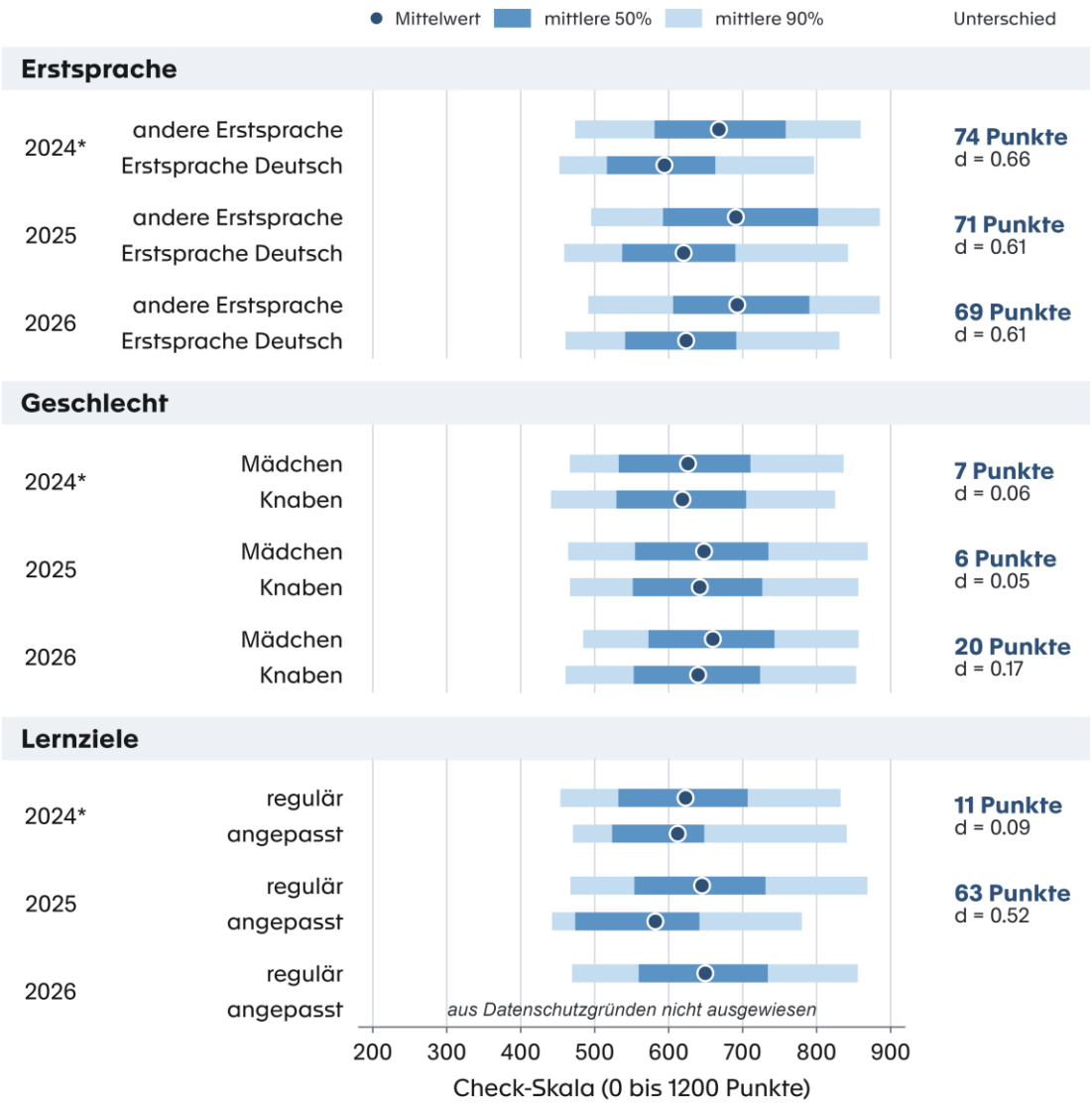
Bei der Interpretation dieser Veränderung ist zu berücksichtigen, dass die Anzahl der Schülerinnen und Schüler mit registrierten angepassten Lernzielen 2026 deutlich geringer ist als in den Vorjahren. Zudem nahmen Schülerinnen und Schüler mit angepassten Lernzielen häufiger nicht am Check P4 teil oder schlossen den Test nicht ab (vgl. Kapitel 1.3). Die starke Zunahme des Gruppenunterschieds lässt sich daher nicht eindeutig als Leistungsentwicklung der Schülerinnen und Schüler mit angepassten Lernzielen interpretieren. Vielmehr dürfte sie zumindest teilweise mit der veränderten Zusammensetzung der ausgewerteten Gruppe zusammenhängen.

Im Fach Englisch (Abbildung 8) zeigt sich über alle drei Durchführungen ein deutlicher Unterschied nach Erstsprache zugunsten der Schülerinnen und Schüler mit einer anderen Erstsprache. Der Unterschied beträgt 74 Punkte im Jahr 2024, 71 Punkte im Jahr 2025 und 69 Punkte im Jahr 2026 (d = 0.61). Er bleibt damit über die drei Durchführungen weitgehend stabil. Das bereits in den Vorjahren beobachtete Muster bestätigt sich somit auch 2026. Der Unterschied entspricht etwas mehr als dem erwarteten Lernfortschritt eines Schuljahres.

Mädchen erzielen in allen drei Jahren höhere Ergebnisse als Knaben, wobei der Unterschied 2024 und 2025 gering ausfiel. Er beträgt 7 Punkte im Jahr 2024 und 6 Punkte im Jahr 2025. Im Jahr 2026 liegt der Unterschied bei 20 Punkten (d = 0.17). Der Unterschied zwischen den Geschlechtern hat damit gegenüber den beiden Vorjahren zugenommen, bleibt gemessen an der Effektstärke jedoch gering. Ob sich daraus eine längerfristige Entwicklung ergibt, lässt sich auf Grundlage der drei bisherigen Durchführungen noch nicht beurteilen.

Für Schülerinnen und Schüler mit angepassten Lernzielen lässt sich die Entwicklung für die Durchführung 2026 nicht sinnvoll beurteilen. Bereits 2025 lagen lediglich sechs Ergebnisse vor, 2026 sind es noch drei. Aufgrund der sehr geringen Fallzahlen wird für 2026 kein Mittelwert für Schülerinnen und Schüler mit angepassten Lernzielen ausgewiesen und es wird auf einen Vergleich mit der aktuellen Durchführung verzichtet.

Abbildung 8: Ergebnisse in Englisch nach individuellen Merkmalen und Jahr

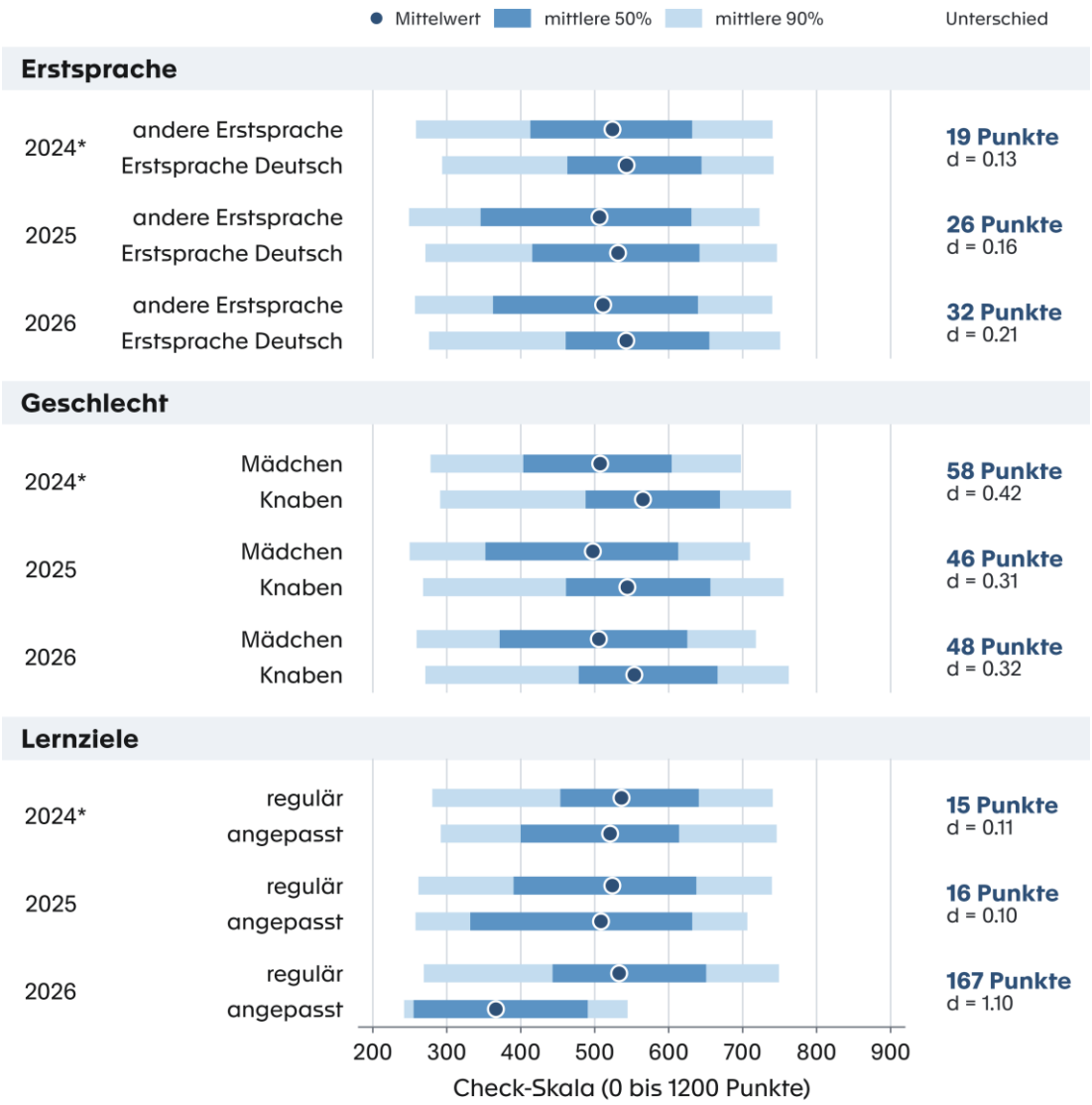


Die Balken zeigen die mittleren 50 bzw. 90 Prozent der Ergebnisse. Unterschied: Differenz der Mittelwerte beider Gruppen in Punkten auf der Check-Skala; d = Effektstärke nach Cohen (0.2 klein, 0.5 mittel, 0.8 gross). * 2024: Pilotdurchführung im Mai, zwei Monate später als ab 2025.

Im Fach Mathematik (vgl. Abbildung 9) erzielen Schülerinnen und Schüler mit Deutsch als Erstsprache in allen drei Durchführungen höhere Ergebnisse als Schülerinnen und Schüler mit einer anderen Erstsprache. Der Unterschied beträgt 19 Punkte im Jahr 2024, 26 Punkte im Jahr 2025 und 32 Punkte im Jahr 2026 (d = 0.21). Damit hat der Unterschied über die drei Durchführungen tendenziell zugenommen, bleibt gemessen an der Effektstärke jedoch klein. Auf Grundlage der drei bisherigen Durchführungen lässt sich noch nicht beurteilen, ob sich diese Entwicklung fortsetzt.

Knaben erzielen in Mathematik in allen drei Durchführungen höhere Ergebnisse als Mädchen. Der Unterschied beträgt 58 Punkte im Jahr 2024, 46 Punkte im Jahr 2025 und 48 Punkte im Jahr 2026 (d = 0.32). Der Geschlechterunterschied ist damit seit 2025 weitgehend stabil und entspricht 2026 ungefähr dem erwarteten Lernfortschritt eines Semesters.

Abbildung 9: Ergebnisse in Mathematik nach individuellen Merkmalen und Jahr



Die Balken zeigen die mittleren 50 bzw. 90 Prozent der Ergebnisse. Unterschied: Differenz der Mittelwerte beider Gruppen in Punkten auf der Check-Skala; d = Effektstärke nach Cohen (0.2 klein, 0.5 mittel, 0.8 gross). * 2024: Pilotdurchführung im Mai, zwei Monate später als ab 2025.

Bei den angepassten Lernzielen zeigt sich ähnlich wie in Deutsch 2026 eine deutliche Veränderung. In den Jahren 2024 und 2025 erzielten Schülerinnen und Schüler mit angepassten Lernzielen im Mittel 15 beziehungsweise 16 Punkte tiefere Ergebnisse als Schülerinnen und Schüler mit regulären Lernzielen. Im Jahr 2026 beträgt der Unterschied dagegen 167 Punkte (d = 1.10). Gleichzeitig liegen 2026 nur noch 15 Ergebnisse von Schülerinnen und Schülern mit angepassten Lernzielen vor, gegenüber 43 im Jahr 2024 und 73 im Jahr 2025. Wie in Deutsch lässt sich die starke Zunahme des Gruppenunterschieds deshalb nicht eindeutig als Leistungsentwicklung der Schülerinnen und Schüler mit angepassten Lernzielen interpretieren. Vielmehr dürfte sie zumindest teilweise mit der veränderten Zusammensetzung der ausgewerteten Gruppe zusammenhängen.

Zusammenfassend zeigen sich bei den Unterschieden nach Erstsprache und Geschlecht überwiegend stabile Muster. In Deutsch hatten die Unterschiede nach Erstsprache und Geschlecht von 2024 auf 2025 zugenommen; von 2025 auf 2026 blieben sie weitgehend stabil. In Englisch ist der deutliche Unterschied nach Erstsprache über alle drei Durchführungen weitgehend konstant. Der Unterschied zwischen Mädchen und Knaben ist 2026 zwar grösser als in den beiden Vorjahren, bleibt gemessen an der Effektstärke jedoch gering. In Mathematik hat der Unterschied nach Erstsprache über die drei Durchführungen tendenziell zugenommen, während der Geschlechterunterschied seit 2025 weitgehend stabil ist.

Die Ergebnisse nach angepassten Lernzielen sind dagegen nur eingeschränkt über die drei Durchführungen vergleichbar. In Deutsch und Mathematik fallen die Gruppenunterschiede 2026 deutlich grösser aus als in den Vorjahren. Gleichzeitig ist die Zahl der Schülerinnen und Schüler mit angepassten Lernzielen, die in die Auswertungen eingehen, deutlich geringer. In Englisch sind die Fallzahlen zu klein für einen aussagekräftigen Vergleich. Die Veränderungen können daher nicht eindeutig als Leistungsentwicklung interpretiert werden und dürften zumindest teilweise mit einer veränderten Zusammensetzung der ausgewerteten Gruppen zusammenhängen.

3.3 Trends nach Kompetenzstufen

Abbildung 10 zeigt die Entwicklung des Anteils der Schülerinnen und Schüler, die mindestens Kompetenzstufe V erreichen. Wie in Kapitel 2.4 beschrieben, beherrschen Schülerinnen und Schüler ab Kompetenzstufe V in der Regel die Kompetenzen, die bis zum Orientierungspunkt in der Mitte des 2. Zyklus vorgesehen sind.

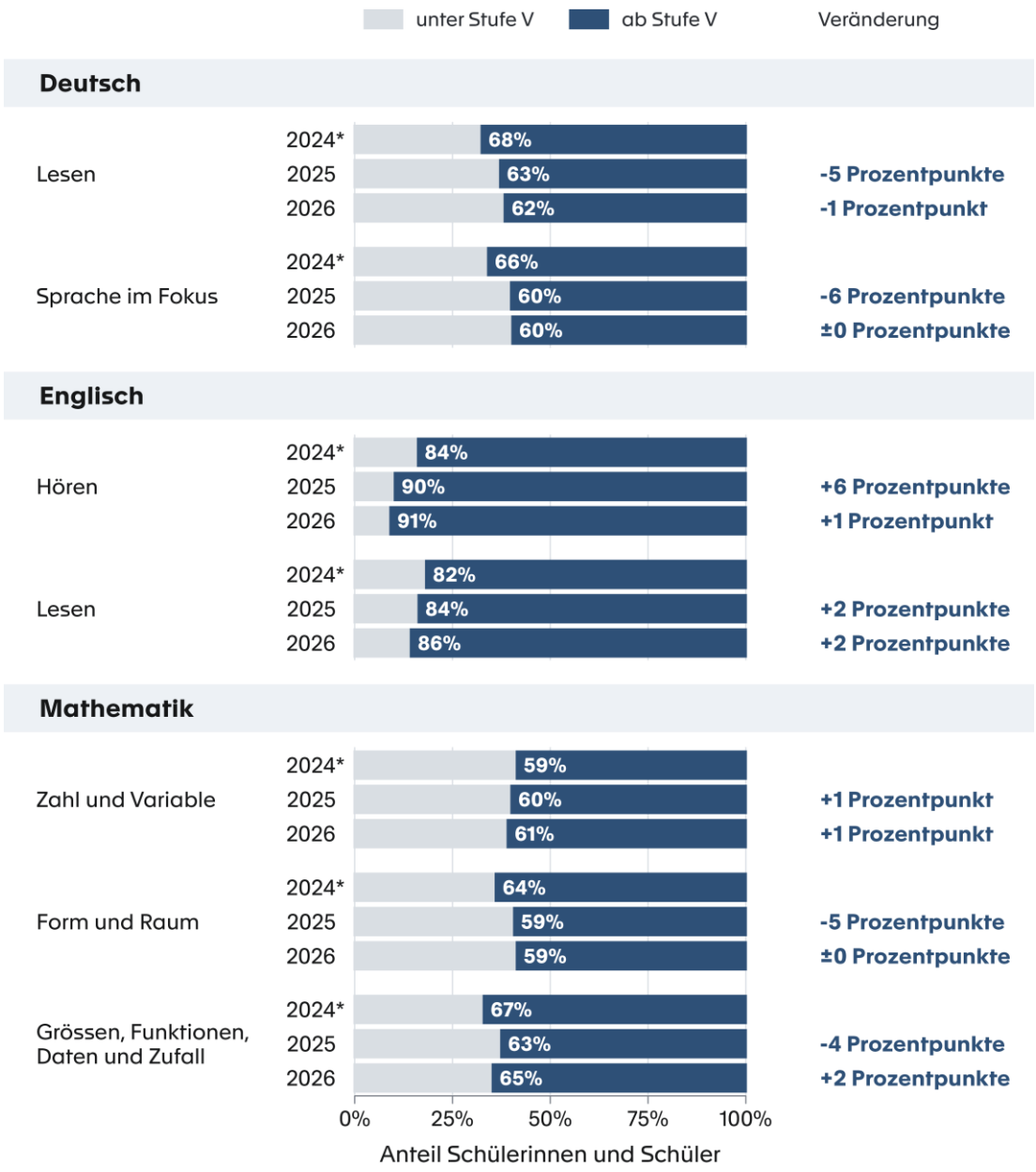
In Deutsch ging der Anteil von 2024 auf 2025 in beiden Kompetenzbereichen zurück: um 4.7 Prozentpunkte im Lesen und um 5.9 Prozentpunkte in Sprache im Fokus. Von 2025 auf 2026 fielen die Veränderungen deutlich geringer aus. Im Lesen sank der Anteil nochmals um 1.2 Prozentpunkte, in Sprache im Fokus um 0.3 Prozentpunkte. Der deutlichere Rückgang zwischen 2024 und 2025 setzte sich damit 2026 nicht fort.

In Englisch stieg der Anteil in beiden Kompetenzbereichen. Von 2024 auf 2025 betrug die Zunahme 5.9 Prozentpunkte im Hören und 1.9 Prozentpunkte im Lesen. Von 2025 auf 2026 nahm der Anteil nochmals um 1.1 Prozentpunkte im Hören und 2.0 Prozentpunkte im Lesen zu. Damit zeigt sich über beide Zeiträume hinweg eine positive Entwicklung, wobei der stärkere Anstieg im Hören bereits zwischen 2024 und 2025 erfolgte.

In Mathematik unterscheiden sich die Entwicklungen zwischen den Kompetenzbereichen. In Zahl und Variable stieg der Anteil von 2024 auf 2025 um 1.3 Prozentpunkte und von 2025 auf 2026 um weitere 1.0 Prozentpunkte. In Form und Raum ging der Anteil zunächst um 4.7 Prozentpunkte zurück und blieb anschliessend mit einer Veränderung von –0.6 Prozentpunkten weitgehend stabil. In Grössen, Funktionen, Daten und Zufall sank der Anteil von 2024 auf 2025 um 4.5 Prozentpunkte, stieg von 2025 auf 2026 jedoch wieder um 2.2 Prozentpunkte.

Zusammenfassend zeigen sich zwischen 2024 und 2025 stärkere Veränderungen als zwischen 2025 und 2026. In Deutsch gingen die Anteile zunächst deutlich zurück und blieben anschliessend weitgehend stabil. In Englisch nahmen sie in beiden Zeiträumen zu. In Mathematik zeigt sich kein einheitliches Muster.

Abbildung 10: Anteil der Schülerinnen und Schüler mit mindestens Kompetenzstufe V nach Kompetenzbereich und Jahr



Stufe V entspricht dem Orientierungspunkt im 2. Zyklus nach Lehrplan 21. Veränderung: Differenz des Anteils «ab Stufe V» gegenüber dem Vorjahr in Prozentpunkten. * 2024: Pilotdurchführung im Mai, zwei Monate später als ab 2025.

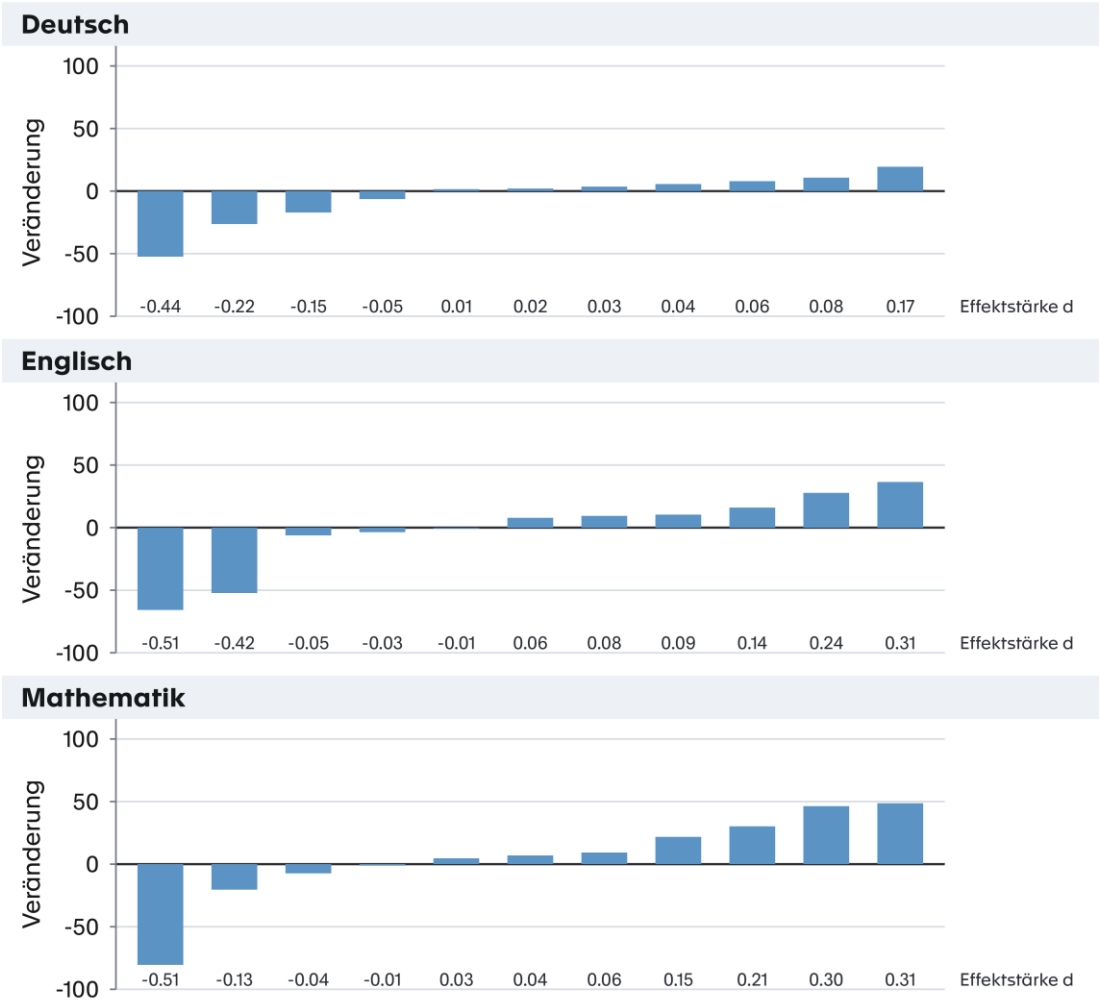
3.4 Veränderung der Gemeindemittelwerte

Abbildung 11 zeigt für jede Gemeinde die Veränderung des Mittelwerts von 2025 zu 2026. Die Gemeinden sind nach der Höhe der Veränderung geordnet. Zusätzlich ist für jede Veränderung die Effektstärke *d* angegeben.

In allen drei Fächern gibt es sowohl Gemeinden mit höheren als auch mit tieferen Ergebnissen als im Vorjahr. In Deutsch reichen die Veränderungen von rund –52 bis +19

Punkten, in Englisch von rund –66 bis +36 Punkten und in Mathematik von rund –81 bis +49 Punkten. Die Gemeindemittelwerte bleiben jedoch grösstenteils stabil. Die Effektstärken liegen überwiegend unter $d = 0.2$ und sind damit als klein einzustufen. Einzelne Gemeinden weisen grössere Veränderungen auf; die stärksten Veränderungen in Englisch und Mathematik erreichen mit rund $d = 0.5$ einen mittleren Effekt.

Abbildung 11: Veränderung der Gemeindemittelwerte von 2025 zu 2026



Veränderung: Differenz der Gemeinde-Mittelwerte zwischen 2025 und 2026 in Punkten auf der Check-Skala; d = Effektstärke nach Cohen (0.2 klein, 0.5 mittel, 0.8 gross).

Die überwiegend geringen Veränderungen der Gemeindemittelwerte unterstützen den Befund, dass nur ein geringer Anteil der Leistungsunterschiede auf Unterschiede zwischen Gemeinden entfällt (vgl. Kapitel 2.6). Im Vergleich zu den Unterschieden zwischen Klassen ist die Gemeindezugehörigkeit damit von untergeordneter Bedeutung. Die Veränderungen der Gemeindemittelwerte sind zudem nicht als Hinweis auf unterschiedliche Schulqualität zu verstehen. Bei der Interpretation ist zu beachten, dass Gemeindemittelwerte auf einer vergleichsweise kleinen Anzahl Schülerinnen und Schüler beruhen können. Je kleiner eine Gemeinde ist, desto stärker können einzelne Ergebnisse oder einzelne Klassen den Mittelwert beeinflussen. Grössere

Veränderungen von einem Jahr zum nächsten weisen deshalb nicht zwingend auf eine systematische Veränderung der Leistungen in einer Gemeinde hin.

4 Nutzung von Mindsteps und Ergebnisse im Check P4

Mindsteps ist eine digitale Aufgabensammlung mit über 60'000 kompetenzorientierten Aufgaben in Deutsch, Englisch, Französisch, Mathematik sowie Natur und Technik. Sie ermöglicht Schülerinnen und Schülern, gezielt zu üben und zu lernen, und unterstützt Lehrpersonen bei der individuellen Förderung. Die Aufgaben beziehen sich auf die Kompetenzbereiche des Lehrplans 21 und sind auf derselben Skala verortet wie die Aufgaben der Checks.

Mindsteps wurde auf der Grundlage der zentralen Prinzipien von «Lernen sichtbar machen» des neuseeländischen Bildungsforschers John Hattie entwickelt.[9] Mindsteps ergänzt die standardisierten Checks, unterscheidet sich von diesen jedoch in seiner Funktion. Während die Checks nach einheitlichen Vorgaben durchgeführt werden und eine unabhängige Standortbestimmung ermöglichen, dient Mindsteps dem Üben und Lernen im Unterricht. Schülerinnen und Schüler können sowohl von Lehrpersonen zusammengestellte Aufgabenserien im Bereich Mind als auch selbstständig ausgewählte Aufgabenserien im Bereich Steps bearbeiten. Die Aufgaben können im Unterricht oder zu Hause gelöst werden.

Mind und Steps umfassen Aufgaben zu allen Kompetenzen, die auch im Check P4 relevant sind. Darüber hinaus bietet Mindsteps den Schülerinnen und Schülern die Möglichkeit, sich mit den Aufgabenformaten und der Bearbeitung am Computer vertraut zu machen.

Kapitel 4 untersucht zunächst, wie häufig Mindsteps im Schuljahr 2025/26 genutzt wurde und wie sich die Nutzung gegenüber den beiden Vorjahren entwickelt hat. Anschliessend wird untersucht, in welchem Zusammenhang die Nutzung von Mindsteps mit den Ergebnissen im Check P4 steht. Dabei ist zu berücksichtigen, dass aus einem solchen Zusammenhang nicht auf eine ursächliche Wirkung von Mindsteps auf die Leistungen geschlossen werden kann. Sowohl die Nutzung von Mindsteps als auch die Ergebnisse im Check P4 können mit weiteren Merkmalen der Schülerinnen und Schüler, des Unterrichts oder der Schulen zusammenhängen.

4.1 Nutzung von Mindsteps

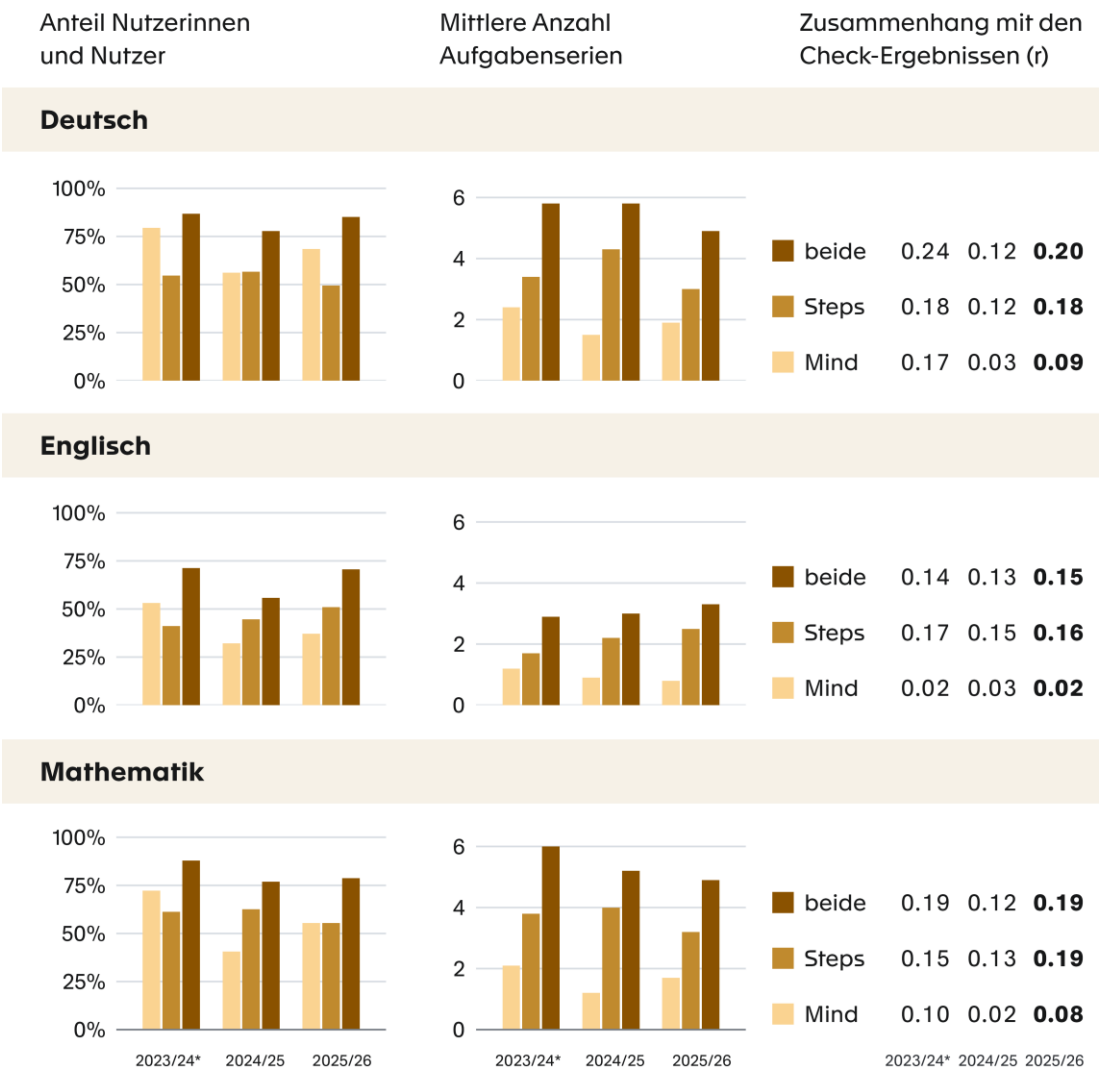
Für alle Schülerinnen und Schüler der 4. Klassen im Kanton Zug wurde die Anzahl der bearbeiteten Aufgabenserien in den im Check P4 geprüften Fächern ab Beginn des Schuljahres bis zum Ende des Durchführungszeitraums des Check P4 ausgezählt. Dabei wurden nur Aufgabenserien berücksichtigt, in denen die Schülerinnen und Schüler mindestens die Hälfte der Aufgaben beantworteten und sich dafür pro Aufgabe mindestens fünf Sekunden Zeit nahmen. Mit diesem Kriterium werden insbesondere Aufgabenserien ausgeschlossen, die lediglich angesehen oder durchgeklickt wurden,

[9] Hattie, John A. C. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement.* London & New York: Routledge. https://doi.org/10.4324/9780203887332

ohne dass die Aufgaben tatsächlich bearbeitet wurden. Dadurch wurden rund 9 Prozent der durchgeführten Aufgabenserien ausgeschlossen, überwiegend im Bereich Steps.

Abbildung 12 zeigt für jedes Fach und Schuljahr den Anteil der Schülerinnen und Schüler, die Mindsteps im betrachteten Zeitraum mindestens einmal nutzten, sowie die durchschnittliche Anzahl bearbeiteter Aufgabenserien. Dabei ist zu berücksichtigen, dass der Check P4 im Jahr 2024 rund zwei Monate später stattfand als 2025 und 2026. Der betrachtete Nutzungszeitraum war im Schuljahr 2023/24 entsprechend länger.

Abbildung 12: Nutzung von Mindsteps und Zusammenhang mit den Ergebnissen im Check P4 nach Fach und Schuljahr



Mind-Bereich: von der Lehrperson vorbereitete Aufgabenserien. Steps-Bereich: Aufgabenserien, welche die Schülerinnen und Schüler selbständig lösen. Anteil Nutzerinnen und Nutzer: mindestens eine Aufgabenserie im betreffenden Fach. r: Korrelation zwischen der logarithmierten Anzahl Aufgabenserien und dem Check-Ergebnis; hervorgehoben ist das aktuelle Schuljahr. * 2023/24: Der Check P4 fand zwei Monate später statt, der Nutzungszeitraum war entsprechend länger.

Im Schuljahr 2025/26 nutzte die Mehrheit der Schülerinnen und Schüler Mindsteps in allen drei Fächern mindestens einmal. Gegenüber dem Vorjahr stieg der Anteil in Deutsch von 78 auf 85 Prozent, in Englisch von 56 auf 71 Prozent und in Mathematik von 77 auf 79 Prozent. Damit nahm die Reichweite von Mindsteps insbesondere in Englisch und Deutsch zu. Die durchschnittliche Anzahl bearbeiteter Aufgabenserien ging dagegen in Deutsch von 5.8 auf 4.9 und in Mathematik von 5.2 auf 4.9 leicht zurück. In Englisch stieg sie von 3.0 auf 3.3 Aufgabenserien.

Bei der getrennten Betrachtung von Mind und Steps zeigt sich ein unterschiedliches Muster. Die Nutzung von Mind nahm gegenüber dem Vorjahr in allen drei Fächern zu, gemessen am Anteil der Schülerinnen und Schüler mit mindestens einer bearbeiteten Aufgabenserie. Besonders deutlich ist die Zunahme in Mathematik von 41 auf 56 Prozent und in Deutsch von 56 auf 68 Prozent. In Englisch stieg der Anteil von 32 auf 37 Prozent. Auch die durchschnittliche Anzahl bearbeiteter Mind-Aufgabenserien nahm in Deutsch und Mathematik zu, während sie in Englisch mit 0.8 gegenüber 0.9 Aufgabenserien weitgehend stabil blieb.

Bei Steps zeigt sich eine andere Entwicklung. In Deutsch und Mathematik gingen sowohl der Anteil der Schülerinnen und Schüler mit mindestens einer bearbeiteten Aufgabenserie als auch die durchschnittliche Anzahl bearbeiteter Aufgabenserien gegenüber dem Vorjahr zurück. In Englisch nahm die Steps-Nutzung hingegen zu: Der Anteil stieg von 45 auf 51 Prozent, die durchschnittliche Anzahl bearbeiteter Aufgabenserien von 2.2 auf 2.5.

Insgesamt wurde Mindsteps im Schuljahr 2025/26 von einem grösseren Anteil der Schülerinnen und Schüler genutzt als im Vorjahr. Gleichzeitig verschob sich die Nutzung in Deutsch und Mathematik stärker in Richtung des von den Lehrpersonen gesteuerten Bereichs Mind, während die Nutzung von Steps zurückging. In Englisch nahmen dagegen sowohl die Nutzung von Mind als auch jene von Steps zu. Aus den Nutzungsdaten allein lässt sich nicht ableiten, welche Gründe diesen Veränderungen zugrunde liegen.

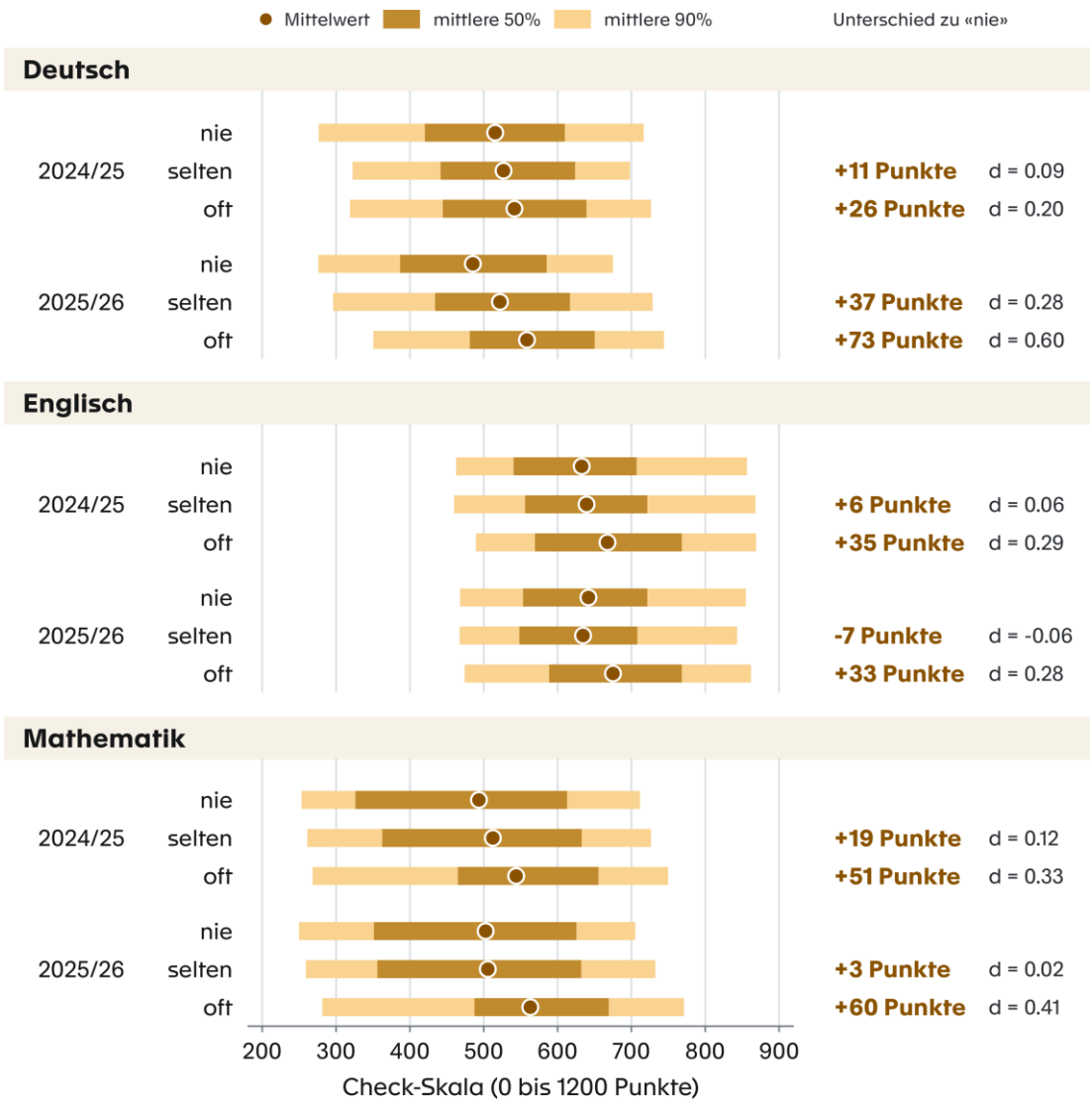
4.2 Zusammenhang der Mindsteps-Nutzung und der Ergebnisse im Check P4

Zwischen der Anzahl bearbeiteter Aufgabenserien in Mindsteps und den Ergebnissen im Check P4 besteht in allen drei Fächern und über alle drei Durchführungen hinweg ein positiver Zusammenhang (Abbildung 12). Schülerinnen und Schüler, die mehr Aufgabenserien bearbeiteten, erzielten im Durchschnitt höhere Ergebnisse im Check P4. Die Zusammenhänge fallen jedoch insgesamt gering aus. In Deutsch ging die Korrelation von $r = 0.24$ im Jahr 2024 auf $r = 0.12$ im Jahr 2025 zurück und stieg 2026 wieder auf $r = 0.20$. In Mathematik zeigt sich mit $r = 0.19$, $r = 0.12$ und $r = 0.19$ eine ähnliche Entwicklung. In Englisch blieb der Zusammenhang mit $r = 0.14$, $r = 0.13$ und $r = 0.15$ über die drei Durchführungen weitgehend stabil.

Um den Zusammenhang anschaulicher darzustellen, wurden die Schülerinnen und Schüler anhand ihrer Nutzung von Mindsteps in drei Gruppen eingeteilt: keine Nutzung, seltene Nutzung mit einer bis drei bearbeiteten Aufgabenserien und häufige Nutzung mit mehr als drei bearbeiteten Aufgabenserien. Abbildung 13 zeigt die Ergebnisse dieser drei Gruppen für die beiden Schuljahre 2024/25 und 2025/26.

In Deutsch erzielten Schülerinnen und Schüler mit häufiger Mindsteps-Nutzung in beiden Schuljahren die höchsten Ergebnisse. Im Schuljahr 2025/26 beträgt der Mittelwert dieser Gruppe 558 Punkte, gegenüber 522 Punkten bei seltener Nutzung und 485 Punkten bei keiner Nutzung. Der Unterschied zwischen häufiger und keiner Nutzung beträgt damit 73 Punkte. Er fällt im Schuljahr 2025/26 deutlich grösser aus als im Schuljahr 2024/25.

Abbildung 13: Ergebnisse im Check P4 nach Nutzung von Mindsteps im aktuellen und vergangenen Schuljahr



Nutzung von Mindsteps pro Fach: nie = keine Aufgabenserie, selten = 1 bis 3, oft = mehr als 3 Aufgabenserien. Die Balken zeigen die mittleren 50 bzw. 90 Prozent der Ergebnisse. Unterschied: Differenz der Mittelwerte zur Gruppe «nie» in Punkten auf der Check-Skala, d = Effektstärke nach Cohen (0.2 klein, 0.5 mittel, 0.8 gross).

Auch in Englisch erzielten Schülerinnen und Schüler mit häufiger Mindsteps-Nutzung in beiden Schuljahren die höchsten Ergebnisse. Im Schuljahr 2025/26 liegt ihr Mittelwert bei 675 Punkten, gegenüber 635 Punkten bei seltener und 642 Punkten bei keiner Nutzung. Der Unterschied zwischen häufiger und keiner Nutzung beträgt damit 33 Punkte

(35 Punkte im Schuljahr 2024/25). Zwischen keiner und seltener Nutzung zeigt sich hingegen kein konsistentes Muster. Die beiden Gruppen liegen in beiden Schuljahren nahe beieinander.

In Mathematik erzielten Schülerinnen und Schüler mit häufiger Nutzung ebenfalls in beiden Schuljahren die höchsten Ergebnisse. Im Schuljahr 2025/26 liegt ihr Mittelwert bei 563 Punkten, gegenüber 506 Punkten bei seltener und 503 Punkten bei keiner Nutzung. Der Unterschied zwischen häufiger und keiner Nutzung beträgt damit 60 Punkte und fällt grösser aus als im Schuljahr 2024/25. Zwischen keiner und seltener Nutzung besteht im Schuljahr 2025/26 dagegen kaum ein Unterschied.

Zusammenfassend zeigt sich über alle drei Durchführungen und Fächer hinweg ein konsistentes Muster: Schülerinnen und Schüler, die Mindsteps häufig nutzen, erzielen im Durchschnitt höhere Ergebnisse im Check P4 als Schülerinnen und Schüler ohne Mindsteps-Nutzung. Betrachtet man die Bereiche Mind und Steps getrennt, fällt der Zusammenhang mit den Check-Ergebnissen 2026 in allen drei Fächern für Steps etwas stärker aus. Dieser Unterschied ist jedoch schwierig zu interpretieren, da nicht bekannt ist, in welchem Ausmass die Nutzung von Steps selbstständig erfolgt oder durch die Lehrpersonen angeregt wird.

Aus den beobachteten Zusammenhängen lässt sich zudem nicht auf eine ursächliche Wirkung der Mindsteps-Nutzung auf die Leistungen schliessen. Es ist beispielsweise möglich, dass eine häufigere Nutzung von Mindsteps mit einem höheren Engagement der Schülerinnen und Schüler oder der Lehrpersonen zusammenhängt. Solche und weitere Merkmale können sowohl die Nutzung von Mindsteps als auch die Ergebnisse im Check P4 beeinflussen.