



Institut für Bildungsevaluation

Check P4 Zug 2025

Kantonsbericht zuhanden der Direktion für Bildung und Kultur, Kanton Zug

IBE – Institut für Bildungsevaluation

Zürich, 29. August 2025

Anschrift

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

Zusammenfassung

Der Check P4 ermöglicht eine externe Standortbestimmung in den Fächern Deutsch, Englisch und Mathematik anhand standardisierter Online-Tests. Im März 2025 wurde der Check P4 in allen 4. Klassen der gemeindlichen Schulen im Kanton Zug zum zweiten Mal seit 2024 durchgeführt. Die Ergebnisse dienen der individuellen Beurteilung sowie der Unterrichts- und Schulentwicklung. Darüber hinaus können sie für das Monitoring des kantonalen Bildungssystems genutzt werden, um beispielsweise Zusammenhänge mit individuellen Merkmalen (Geschlecht, Erstsprache, Alter, Lernziele) zu analysieren oder die Erreichung der Orientierungspunkte im zweiten Zyklus des Lehrplans 21 zu überprüfen. Die wiederholte Durchführung des Check P4 ermöglicht erstmals den Vergleich der Leistungen in den getesteten Kompetenzbereichen über die Jahre hinweg.

Der vorliegende Bericht informiert über die Ergebnisse des Check P4 2025 auf kantonaler Ebene. Die Ergebnisse bestätigen die besondere Stärke der Zuger Schülerinnen und Schüler im Fach Englisch: Rund 80% erreichen oder übertreffen die bis zum Orientierungspunkt im 2. Zyklus vorgesehenen Kompetenzen. In Deutsch und Mathematik liegt dieser Anteil bei rund 60%. Da die Durchführung 2025 zwei Monate früher als im Vorjahr erfolgte, war ein leichter Leistungsrückgang zu erwarten. Tatsächlich liegen die Ergebnisse im Durchschnitt etwas tiefer und die Anteile auf Kompetenzstufe V sind in den meisten Bereichen um 3 bis 6 Prozentpunkte gesunken. Das Ausmass entspricht in etwa dem zu erwartenden Rückgang aufgrund der verkürzten Unterrichtszeit bis zur Check-Durchführung.

Des Weiteren zeigen die Analysen, dass in den Fächern Deutsch und Mathematik die Unterschiede zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache und denjenigen mit Deutsch als Zweitsprache zunahm. Im Fach Deutsch ist die Zunahme dieses Unterschieds markant und entspricht – gemessen am erwarteten Leistungsfortschritt über die Zeit – mehr als einem Semester. Ebenfalls nahm der Vorsprung der Mädchen gegenüber den Knaben im Fach Deutsch zu, wobei dieser Unterschied etwa einem halben Semester entspricht. Im Fach Mathematik nahm der Vorsprung der Knaben gegenüber den Mädchen ab. Es verbleibt allerdings eine Differenz, die annähernd einem Semester entspricht. Ob es sich bei diesen Veränderungen um stabile Trends handelt, wird sich in den kommenden Jahren zeigen, sobald mehr Datenpunkte vorliegen.

Zusätzlich zu den Unterschieden zwischen Gruppen mit verschiedenen individuellen Merkmalen wurden auch die Unterschiede zwischen den Schulklassen und Gemeinden untersucht. In allen drei geprüften Fächern zeigte sich eine grosse Spannweite der Klassenergebnisse, sodass Unterschiede zwischen Klassen bis zu 16% der Varianz in den Schülerergebnissen erklären. Dies unterstreicht die Bedeutung einer vom Klassenkontext unabhängigen Standortbestimmung als Ergänzung zu den stärker klassenbezogenen Schulnoten. Zwischen den Gemeinden wurden hingegen vergleichsweise geringe Unterschiede und eine weitgehende Überlappung der Ergebnisse in allen drei Fächern festgestellt.

Ergänzend zum Check P4 stand den Schülerinnen und Schülern die Aufgabensammlung Mindsteps während des ganzen Schuljahres zur Verfügung. Die Analyse der Nutzerzahlen zeigt, dass die Mehrheit der Schülerinnen und Schüler mit Mindsteps arbeitete. Allerdings nahm dieser Anteil im Vergleich zum Schuljahr 2023/2024 ab und das Nutzerverhalten der Schülerinnen und Schüler sowie der Lehrpersonen veränderte sich. Während die Nutzerzahlen im Bereich *Mind*, in dem Lehrpersonen eigene Aufgabenserien zusammenstellen, deutlich zurückgingen, wurde der Bereich *Steps*, in dem Schülerinnen und Schüler selbstständig arbeiten können, weiterhin häufig genutzt. Wie bereits im Vorjahr geht in allen drei Fächern eine häufige Nutzung von Mindsteps mit besseren Ergebnissen im Check P4 einher. Dies könnte darauf hindeuten, dass Schülerinnen und

Schüler mit Mindsteps einen Lern- und Übungseffekt erzielen. Allerdings ist auch denkbar, dass weitere Einflussfaktoren wie der Unterricht oder Bedingungen an der Schule sowohl die Fähigkeiten der Schülerinnen und Schüler als auch die Nutzung von Mindsteps positiv beeinflussen.

Fazit: Der Check P4 liefert wertvolle Erkenntnisse über den Lernstand der Schülerinnen und Schüler der 4. Klasse im Kanton Zug. Besonders hervorzuheben sind das stabile Leistungsniveau im Vergleich zum vergangenen Jahr, die zunehmenden Unterschiede zwischen Schülerinnen und Schülern mit Deutsch als Erst- und als Zweitsprache sowie das veränderte Nutzerverhalten in Mindsteps. Verbesserungspotenzial zeigt sich bei der einheitlichen Handhabung von angepassten Lernzielen.

Inhalt

1	Einleitung	6
1.1	Ausgangslage	6
1.2	Tests und Methoden	6
1.3	Teilnehmende Schülerinnen und Schüler	8
2	Ergebnisse pro Schulfach im Jahr 2025	10
2.1	Mittelwert und Streuung der Ergebnisse	10
2.2	Ergebnisse nach individuellen Merkmalen	10
2.3	Zusammenhang zwischen den Fächern	13
2.4	Ergebnisse nach Kompetenzstufen	14
2.5	Unterschiede zwischen Klassen	16
2.6	Unterschiede zwischen Gemeinden	18
3	Ergebnisse pro Schulfach im Trend	19
3.1	Trend seit 2024	19
3.2	Trends nach individuellen Merkmalen	20
3.3	Trend in der Erreichung der Kompetenzstufen	23
4	Nutzung von Mindsteps und Ergebnisse im Check P4	24
4.1	Nutzung von Mindsteps	24
4.2	Zusammenhang zwischen Mindsteps und Ergebnissen im Check P4	25

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 der unabhängigen Standortbestimmung für Schülerinnen und Schüler, Lehrpersonen und Schulleitungen. Sie können für die individuelle Beurteilung, die Unterrichts- und Schulentwicklung sowie für das Monitoring des kantonalen Bildungssystems genutzt werden.

Der Check P4 fand im März 2025 in allen 4. Klassen der gemeindlichen Schulen im Kanton Zug statt. Es handelte sich um die erste reguläre Durchführung nach der Pilot-Durchführung im Mai 2024. Gegenüber 2024 erfolgte die Durchführung zwei Monate früher.

Die Ergebnisse der Schülerinnen und Schüler sind seit dem 2. Mai 2025 auf der Onlineplattform *Check Dein Wissen* für die teilnehmenden Lehrpersonen verfügbar. Schulleitungen haben Zugriff auf die aggregierten Klassen- und Schulergebnisse. Damit lagen die Ergebnisse knapp zwei Monate früher vor als bei der Pilot-Durchführung 2024, womit ein zentrales Anliegen der Lehrpersonen aus der Befragung 2024 berücksichtigt wurde.

Der vorliegende Bericht folgt in weiten Teilen der Darstellung von 2024, verzichtet aber auf die damaligen Kapitel zum Vergleich von adaptivem und linearem Testdesign sowie zur Befragung der Lehrpersonen. Neu ermöglicht die wiederholte Durchführung eine längsschnittliche Betrachtung. Kapitel 1.2 beschreibt die standardisierten Online-Tests, Kapitel 1.3 die teilnehmenden Schülerinnen und Schüler. Kapitel 2 berichtet die Ergebnisse nach Fächern, analysiert Unterschiede nach individuellen Merkmalen und Zusammenhänge zwischen den Fächern sowie die erreichten Kompetenzstufen in Bezug auf den Orientierungspunkt des 2. Zyklus (Lehrplan 21). Kapitel 3 vergleicht die Ergebnisse 2025 mit jenen der Pilot-Durchführung 2024. Kapitel 4 betrachtet die Nutzung der digitalen Aufgabensammlung Mindsteps und mögliche Zusammenhänge zu den Ergebnissen (inkl. Vergleich mit dem Jahr 2024).

1.2 Tests und Methoden

Tabelle 1: Fächer, Kompetenzbereiche und Umfang der durchgeführten Tests

Fach	Kompetenzbereiche	Anz. Aufgaben	Testdauer
Deutsch (2 Tests)	Lesen	32	40 min
	Sprache im Fokus	29 – 32 *	20 min
Englisch (2 Tests)	Hören	32	40 min
	Lesen	32	40 min
Mathematik (1 Test)	Zahl und Variable	16	80 min
	Form und Raum	16	
	Grössen, Funktionen, Daten und Zufall	16	

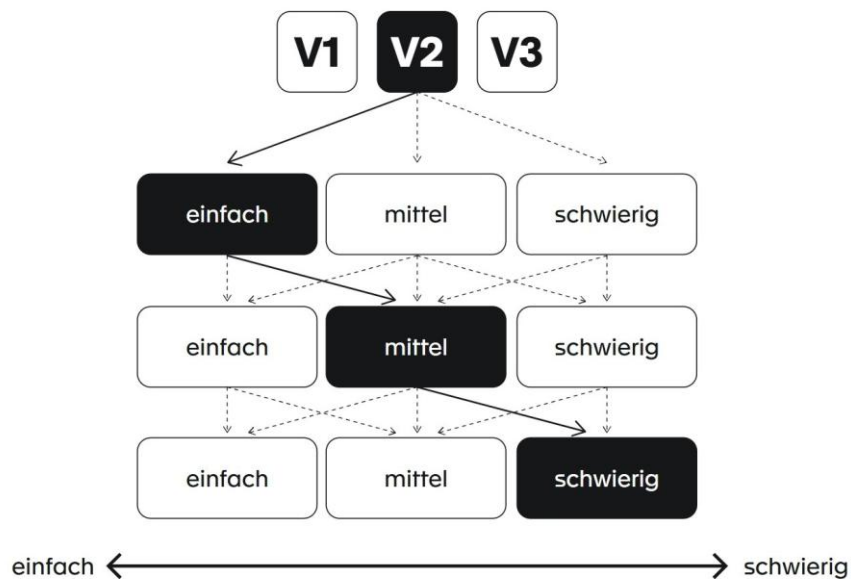
* Der einfachste Aufgabenblock umfasst jeweils eine Aufgabe weniger pro Testteil (7 statt 8 Aufgaben).

Der Check P4 umfasst fünf Online-Tests zu insgesamt sieben Kompetenzbereichen des Lehrplans 21. Die drei Kompetenzbereiche des Fachs Mathematik wurden mit einem

Test geprüft. Für die je zwei Kompetenzbereiche in Deutsch und Englisch wurde je ein eigener Test durchgeführt. Jeder Test umfasste 29 bis 48 Aufgaben, die sich auf vier aufeinanderfolgende Testteile verteilten. Tabelle 1 zeigt die Verteilung der Kompetenzbereiche, die Anzahl Aufgaben und die Testdauer. Die Bearbeitungszeit war pro Testteil begrenzt.

Alle Tests folgten 2025 einem adaptiven Testdesign. Dabei wird die Aufgabenschwierigkeit an die Fähigkeiten der Schülerinnen und Schüler angepasst. Ziel dieses Designs ist es, jeden Schüler und jede Schülerin optimal herauszufordern, sodass er oder sie keiner deutlich über- oder unterfordernden Aufgabe begegnet. Die Anpassung der Aufgabenschwierigkeit bezweckt zudem eine präzisere und verlässlichere Messung der Fähigkeiten der Schülerinnen und Schüler. Das adaptive Testdesign ist in Abbildung 1 visualisiert. Für den ersten Testteil wird jeweils zufällig einer von drei gleich schwierigen Aufgabenblöcken (V1, V2, V3) zugewiesen. In den folgenden Testteilen wählt ein Algorithmus je einen von drei unterschiedlich schwierigen Aufgabenblöcken (einfach, mittel, schwierig) anhand des bisherigen Lösungsverhaltens. Dieses Design stellt sicher, dass alle Schülerinnen und Schüler gleich viele Aufgaben in gleicher Zeit bearbeiten. Lediglich im Test Deutsch Sprache im Fokus umfasst der einfachste Aufgabenblock jeweils eine Aufgabe weniger. Innerhalb eines Aufgabenblocks können Aufgaben übersprungen oder überarbeitet werden.

Abbildung 1: Visualisierung des adaptiven Testdesigns



Anhand der Antworten werden mittels der Item Response Theorie die Fähigkeiten der Schülerinnen und Schüler berechnet 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 (Zyklus 2 und 3 des Lehrplans 21) ab. Die Ergebnisse des Check P4 sind damit grundsätzlich mit Ergebnissen von Checks auf anderen Klassenstufen sowie mit der Durchführung 2024 und künftigen Durchführungen vergleichbar.¹

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

1.3 Teilnehmende Schülerinnen und Schüler

Insgesamt nahmen 1'152 Schülerinnen und Schüler aus 85 Klassen in 30 Schulen am Check P4 teil. Über alle Tests betrachtet, reicht die Anzahl teilnehmender Schülerinnen und Schüler pro Klasse von 3 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 Informationen über die Teilnahme zu erhalten, hat das Amt für gemeindliche Schulen die Anzahl teilnehmender und nicht teilnehmender Schülerinnen und Schüler in allen Gemeinden erfragt und dem Institut für Bildungsevaluation mitgeteilt. Diese Zahlen wurden mit den im Check P4 verfügbaren Ergebnissen abgeglichen.

Gemäss den vom Amt für gemeindliche Schulen erhobenen Daten besuchten zum Zeitpunkt des Check P4 1'183 Schülerinnen und Schüler die 4. Klasse. Somit nahmen 31 Schülerinnen und Schüler nicht teil. Die Teilnahmequote lag bei 97.4%. Diese hohe Ausschöpfung gewährleistet eine hohe Repräsentativität der Ergebnisse.

Einige Schülerinnen und Schüler haben am Check P4 teilgenommen, jedoch einen oder mehrere Tests nicht bearbeitet. Tabelle 2 zeigt die Teilnahmen, Nichtteilnahmen und die Ausschöpfungsquote pro Kompetenzbereich jeweils auf Basis der vom Amt für gemeindliche Schulen erhobenen Zahlen und der im Check P4 ausgewiesenen Ergebnisse.

Tabelle 2: Anzahl teilnehmende Schülerinnen und Schüler sowie Nichtteilnahmen gemäss Amt für gemeindliche Schulen und nach Daten des Check P4

Fach	Kompetenzbereich	Daten Amt			Daten Check P4		
		Teil.	NT	Auss.	Teil.	NT	Auss.
Deutsch	Lesen	1145	38	96.8%	1135	48	95.9%
	Sprache im Fokus	1146	37	96.9%	1116	67	94.3%
Englisch	Hören	1152	31	97.4%	1140	43	96.4%
	Lesen	1151	32	97.3%	1123	60	94.9%
Mathematik		1149	34	97.1%	1129	53	95.5%

Teil.: Anzahl Teilnahmen; NT: Anzahl Nichtteilnahmen

Auss.: Ausschöpfung auf der Basis von $N = 1'183$ Schülerinnen und Schülern

In allen Tests gibt es Abweichungen zwischen den vom Amt für gemeindliche Schulen erhobenen Zahlen und den im Check P4 verfügbaren Ergebnissen. Diese Unterschiede können zumindest teilweise durch unvollständige Teilnahmen erklärt werden: Schliessen Schülerinnen oder Schüler einen Test nicht ab, wird kein Ergebnis berechnet. Solche Fälle gelten im Rahmen des Check P4 als Nichtteilnahmen, wurden von den Schulen aber möglicherweise als Teilnahmen gezählt. Die Ausschöpfung variiert je nach Test zwischen 94.3% (Deutsch Sprache im Fokus) und 96.4% (Englisch Hören).

Das Amt für gemeindliche Schulen hat die Gemeinden zudem zu den Gründen für die Nichtteilnahme befragt. Tabelle 3 fasst die Rückmeldungen zusammen. Am häufigsten wurde eine überdauernde Lernzielanpassung als Grund angegeben (15 Schülerinnen und Schüler in allen Kompetenzbereichen). Aufgrund vorübergehender Lernzielanpassungen nahmen in den beiden Deutsch-Tests insgesamt acht Schülerinnen und Schüler nicht teil, in Mathematik sechs und in Englisch zwei bzw. drei. Zusätzlich wurden pro Kompetenzbereich fünf bis sechs Nichtteilnahmen wegen Krankheit und sieben bis zehn aus anderen Gründen gemeldet.

Tabelle 3: Gründe für die Nichtteilnahme gemäss Erhebung des Amts für gemeindliche Schulen

Fach	Kompetenzbereich	NT	Überd. LZA	Vorüb. LZA	Krank- heit	Andere Gründe
Deutsch	Lesen	38	15	8	6	9
	Sprache im Fokus	37	15	8	6	8
Englisch	Hören	31	15	3	6	7
	Lesen	32	15	2	5	10
Mathematik		34	15	6	6	7

NT: Anzahl Nichtteilnahmen; Überd. LZA: Überdauernde Lernzielanpassung; Vorüb. LZA: Vorübergehende Lernzielanpassung

Für die Anmeldung zum Check P4 wurden die Schülerinnen und Schüler auf der Online-Plattform *Check Dein Wissen* registriert. Dabei erfassten die zuständigen Personen der Schulen auch das Geburtsdatum, das Geschlecht, die Erstsprache (Deutsch als Erstsprache, andere Erstsprache) sowie Fächer, in denen die Schülerinnen oder Schüler angepasste Lernziele haben. Die teilnehmenden Schülerinnen und Schüler sind im Mittel 10.6 Jahre alt (SD = 0.41, Stichdatum 31. März 2025) und der Anteil an Mädchen liegt bei 53.7%. Der Anteil Schülerinnen und Schüler mit Erstsprache Deutsch beträgt 64.7%. In Deutsch haben 90 (7.8%), in Englisch 7 (0.6%) und in Mathematik 79 (6.9%) Schülerinnen und Schüler angepasste Lernziele.

Auch die Check-Daten zeigen, dass angepasste Lernziele häufiger mit Nichtteilnahmen einhergehen. So nahmen am Test in Deutsch Lesen 11% der Schülerinnen und Schüler mit regulären Lernzielen nicht teil bzw. erhielten kein Testergebnis, während es bei Schülerinnen und Schülern mit angepassten Lernzielen 44% waren. Damit ist die Wahrscheinlichkeit einer Nichtteilnahme in diesem Test bei angepassten Lernzielen etwa viermal höher. Dieser Faktor liegt für die Tests Englisch Hören (16), Englisch Lesen (6) und Mathematik (4) noch höher, für Deutsch Sprache im Fokus jedoch tiefer (1.5).

2 Ergebnisse pro Schulfach im Jahr 2025

Die Ergebnisse werden im Folgenden mehrheitlich pro Fach dargestellt. Dazu wurde für jede Schülerin bzw. jeden Schüler der Mittelwert aus den Ergebnissen der Kompetenzbereiche gebildet. Das Ergebnis pro Fach wurde nur berechnet, sofern für die Schülerinnen und Schüler in allen Kompetenzbereichen ein Ergebnis vorliegt.

2.1 Mittelwert und Streuung der Ergebnisse

Die durchschnittlichen Ergebnisse aller Schülerinnen und Schüler sind in Tabelle 4 pro Fach ausgewiesen. Neben dem Mittelwert (M) ist die Standardabweichung (SD) aufgeführt, die die Streuung der Ergebnisse auf der Skala beschreibt. Sie kann vereinfacht als durchschnittliche Abweichung vom Mittelwert interpretiert werden.

Die Ergebnisse in den verschiedenen Fächern sind nicht direkt vergleichbar, da die Tests verschiedener Fächer keine gemeinsame Metrik haben. Allerdings fällt auf, dass der Mittelwert in Englisch im Vergleich zu den Mittelwerten in Deutsch und Mathematik deutlich höher ist. Vergleiche mit Check-Ergebnissen aus anderen Kantonen zeigen, dass dieser Unterschied zwischen den Fachergebnissen in anderen Kantonen so nicht beobachtet wird. Die Leistungen der Schülerinnen und Schüler des Kantons Zug im Fach Englisch sind somit deutlich höher als erwartet. Dieses Muster zeigte sich bereits im vergangenen Jahr und dürfte mit dem vergleichsweise hohen Anteil an Kindern internationaler Fach- und Führungskräfte (Expats) zusammenhängen, die zuhause häufig Englisch sprechen.

Tabelle 4: Mittelwerte (M) und Standardabweichungen (SD) der Ergebnisse in den geprüften Fächern

Fach	M	SD	N
Deutsch	531	126	1114
Englisch	616	116	1118
Mathematik	523	153	1126

2.2 Ergebnisse nach individuellen Merkmalen

In Kombination mit den erfassten Merkmalen der Schülerinnen und Schüler geben die Ergebnisse des Check P4 Aufschluss über Unterschiede oder Ähnlichkeiten in den Leistungen verschiedener Gruppen. Abbildung 2 zeigt die Mittelwerte (dunkelblauer Punkt), die mittleren 50% (blauer Balken) sowie die mittleren 90% (hellblauer Balken) der Ergebnisse pro Fach, jeweils aufgeteilt nach Erstsprache, Geschlecht und Lernzielen. Die Kennwerte zum Gruppenvergleich sind in Tabelle 5 aufgeführt.

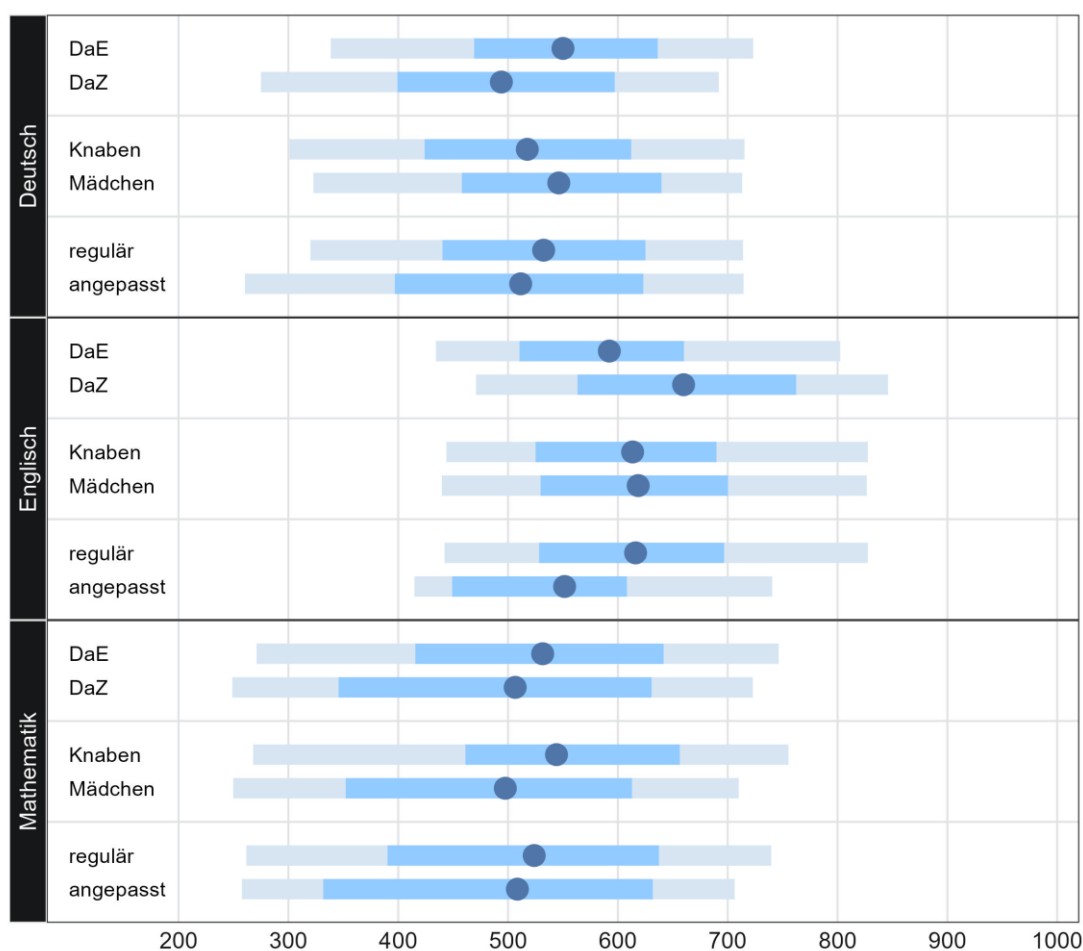
Da die Ergebnisse verschiedener Fächer nicht direkt vergleichbar sind, sollte das Ausmass der Gruppenunterschiede jeweils im Verhältnis zur Streuung der Ergebnisse betrachtet werden. Zu diesem Zweck wird die Effektstärke d verwendet. Sie ergibt sich aus der Differenz der Mittelwerte im Verhältnis zur Streuung der gemessenen Werte innerhalb der beiden Gruppen. Die Effektstärke d drückt also in Standardabweichungen aus, wie gross der Unterschied zwischen zwei Gruppen ist. Nach der Konvention von Cohen wird eine Effektstärke von $d = 0.2$ als klein, von $d = 0.5$ als mittel und von $d = 0.8$ als gross eingestuft.²

² Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.

Zur Einschätzung der praktischen Relevanz kann die Differenz in Punkten zusätzlich mit bekannten Bezugswerten verglichen werden. Erfahrungswerte des Instituts für Bildungsevaluation zeigen, dass Schülerinnen und Schüler auf der Primarstufe pro Semester ein durchschnittlicher Lernfortschritt von rund 47 Punkten in Deutsch, 33 Punkten in Englisch und 50 Punkten in Mathematik erzielen.

Zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache (DaE) und jenen mit Deutsch als Zweitsprache (DaZ) zeigen sich bedeutsame Unterschiede in den Fächern Deutsch und Englisch. In Deutsch beträgt der Vorsprung der Schülerinnen und Schüler mit Erstsprache Deutsch 56 Punkte. Dies entspricht einem kleinen bis mittleren Effekt ($d = 0.46$) und liegt über dem durchschnittlichen Lernfortschritt eines Semesters, weshalb der Unterschied als relevant einzustufen ist. In Englisch hingegen erzielen die Schülerinnen und Schüler mit Deutsch als Zweitsprache durchschnittlich 68 Punkte mehr. Dies entspricht einem mittleren Effekt ($d = 0.55$) und etwa dem erwarteten Fortschritt eines Schuljahres. Dieses Ergebnis könnte darauf zurückzuführen sein, dass Englisch neben Deutsch eine häufige Erstsprache der Schülerinnen und Schüler im Kanton Zug ist.

Abbildung 2: Ergebnisse im Check P4 pro Fach, getrennt nach Geschlecht, nach Erstsprache Deutsch vs. anderer Erstsprache sowie nach Lernzielen



Trotz dieser möglichen Erklärung ist das Ergebnis überraschend. Die Überprüfung der Grundkompetenzen (ÜGK) 2017 stellte für den Kanton Zug fest, dass Schülerinnen und Schüler im 8. Schuljahr, die zuhause nicht Deutsch sprechen, gleich gute oder etwas

tiefere Ergebnisse in Englisch erzielen als jene mit Deutsch als Erstsprache.³ Möglicherweise nimmt dieser Unterschied über die Schullaufbahn bis zum 8. Schuljahr bzw. bis zur 2. Klasse der Sekundarstufe I ab, sodass er dann nicht mehr oder zumindest nicht mehr in diesem Ausmass besteht. Zudem handelt es sich um einen Unterschied auf hohem Leistungsniveau. Die ÜGK unterscheidet lediglich zwischen Erreichen und Nichterreichen der Grundkompetenzen. Auch ein relevanter Unterschied zwischen zwei Gruppen kann in der ÜGK verborgen bleiben, wenn er deutlich oberhalb der Schwelle der Grundkompetenzen liegt.

Im Fach Mathematik beträgt der Vorsprung der Schülerinnen und Schüler mit Deutsch als Erstsprache gegenüber jenen mit Deutsch als Zweitsprache 25 Punkte. Dies entspricht einem kleinen Effekt ($d = 0.16$) und in etwa dem durchschnittlichen Lernfortschritt eines halben Semesters. Die Richtung dieses Unterschieds ist konsistent mit Ergebnissen der ÜGK. Allerdings stellte die ÜGK für das 11. Schuljahr nicht nur einen geringen, sondern einen deutlichen Vorsprung zugunsten der Schülerinnen und Schüler fest, die zuhause Deutsch sprechen⁴. Es könnte sich also auch hier um einen Unterschied handeln, der sich über die Schullaufbahn hinweg verändert, wobei sich dieser Unterschied verstärkt.

Tabelle 5: Ergebnisse pro Fach gruppiert nach individuellen Merkmalen

Fach	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
Erstsprache *	DaZ			DaE		
Deutsch	494	130	382	550	119	732
Englisch	660	121	390	592	107	728
Mathematik	507	158	394	532	150	732
Geschlecht	Mädchen			Knaben		
Deutsch	546	122	516	518	127	598
Englisch	619	117	518	613	116	600
Mathematik	498	149	516	544	154	610
Lernziele	regulär			angepasst		
Deutsch	532	123	1028	512	151	86
Englisch	616	116	1112	552	137	6
Mathematik	524	152	1053	509	164	73

* DaZ: Deutsch als Zweitsprache; DaE: Deutsch als Erstsprache

Zwischen den Ergebnissen von Mädchen und Knaben sind bedeutsame Unterschiede in den Fächern Deutsch und Mathematik zu finden. In Deutsch erzielen Mädchen im Mittel um 28 Punkte höhere Ergebnisse als Knaben, was einem kleinen Effekt ($d = 0.23$) und etwas mehr als dem erwarteten Fortschritt eines halben Semesters entspricht. In Mathematik erzielen hingegen Knaben im Mittel 46 Punkte mehr als Mädchen. Auch dies

³ 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>

⁴ 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>

entspricht einem kleinen Effekt ($d = 0.31$), der jedoch ungefähr einem Semester Lernfortschritt gleichkommt. In Englisch ist hingegen der Unterschied zwischen Mädchen und Knaben mit 6 Punkten ($d = 0.04$) sehr gering und nicht bedeutsam.

Der Leistungsvorsprung von Knaben gegenüber Mädchen in Mathematik wird in der Schweiz regelmässig festgestellt, beispielsweise in der PISA-Erhebung 2022⁵ sowie auch in der ÜGK 2016.⁴ Diese Differenz ist im Kanton Zug auf der Primarstufe aber grösser als in anderen Kantonen (Bildungsraum Nordwestschweiz: ca. 20 Punkte Ende 5. Klasse⁶).

Für Schülerinnen und Schüler mit angepassten Lernzielen ist zu erwarten, dass sie im Mittel tiefere Ergebnisse erzielen als Schülerinnen und Schüler mit regulären Lernzielen, da diese Massnahme bei besonderen Lernherausforderungen eingesetzt wird. Tatsächlich erzielen Schülerinnen und Schüler mit angepassten Lernzielen tiefere Ergebnisse: in Deutsch –20 Punkte, in Englisch –64 Punkte und in Mathematik –15 Punkte. Die Effektstärken sind in Deutsch und Mathematik klein bis vernachlässigbar ($d = 0.17$ bzw. $d = 0.10$), während der Unterschied in Englisch einem mittleren Effekt entspricht ($d = 0.55$). Da die Gruppe mit angepassten Lernzielen im Fach Englisch jedoch nur sechs Schülerinnen und Schüler umfasst, ist dieses Ergebnis nicht verallgemeinerbar.

Auch für Deutsch und Mathematik sind die Vergleiche nach Lernzielanpassung mit Vorsicht zu interpretieren. Wie in Abschnitt 1.3 dargestellt, sind vorübergehende oder überdauernde Lernzielanpassungen ein häufiger Grund für die Nichtteilnahme der Schülerinnen und Schüler an den Tests. Dies deutet darauf hin, dass Schülerinnen und Schüler mit angepassten Lernzielen in den verfügbaren Daten möglicherweise nicht angemessen repräsentiert sind.

2.3 Zusammenhang zwischen den Fächern

Abbildung 3 visualisiert den Zusammenhang der Ergebnisse in Deutsch, Englisch und Mathematik. Die blauen Flächen veranschaulichen, wie häufig verschiedene Kombinationen von Ergebnissen in jeweils zwei Fächern vorkommen. Je dunkler die Schattierung, desto mehr Schülerinnen und Schüler erreichen die jeweilige Kombination von Ergebnissen. Zudem sind die Korrelation und der Determinationskoeffizient aufgeführt.

Die Korrelation (r) beschreibt die Stärke und Richtung des linearen Zusammenhangs zwischen zwei Fächern und kann Werte zwischen -1 (vollständig negativer Zusammenhang) bis $+1$ (perfekter positiver Zusammenhang) annehmen. 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 wie eine Prozentzahl interpretiert werden. Beispielsweise bedeutet der Determinationskoeffizient von $R^2 = 0.22$ für den Zusammenhang zwischen den Ergebnissen in Deutsch und Englisch, dass die Ergebnisse in Englisch zu 22% durch die Ergebnisse in Deutsch vorhersagbar sind – und umgekehrt.

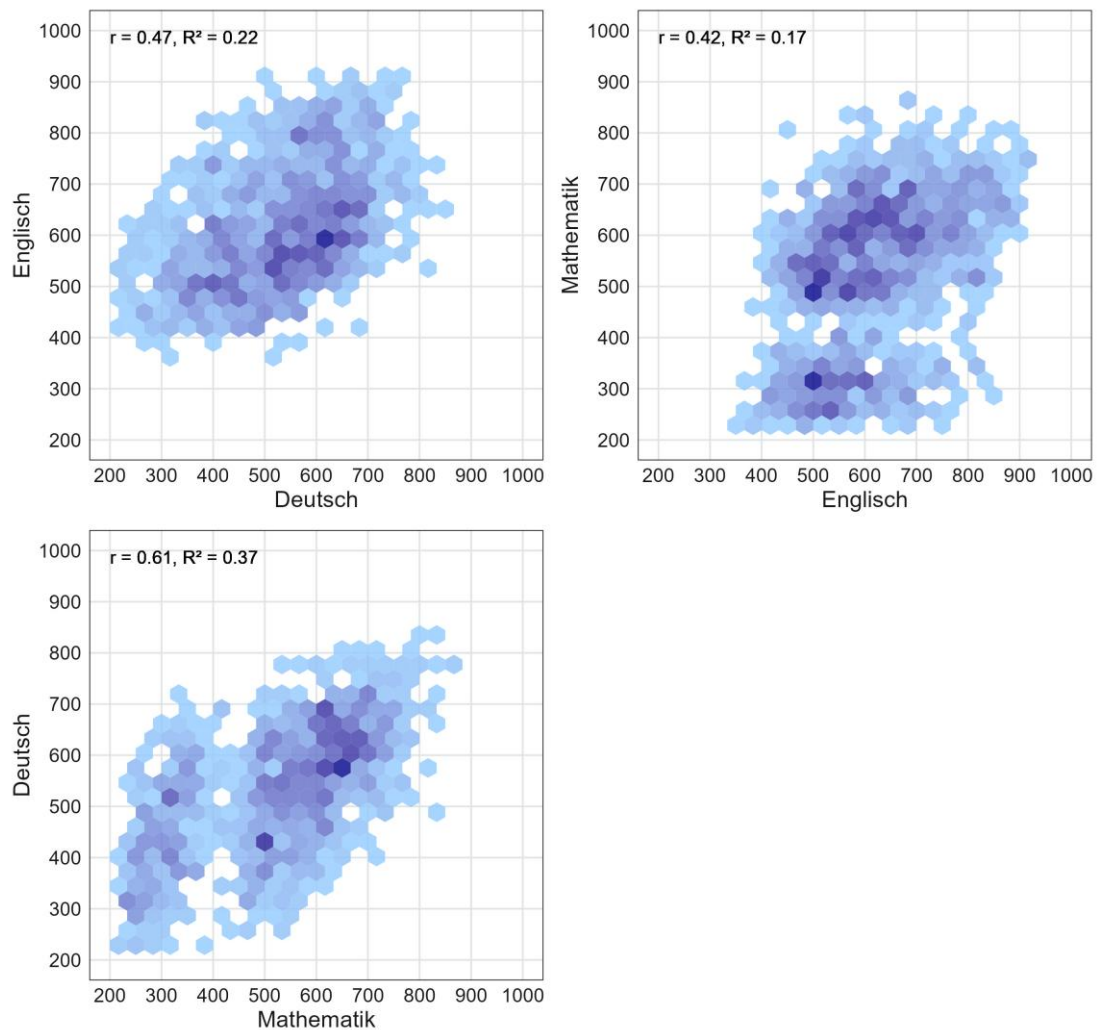
Der stärkste Zusammenhang zeigt sich zwischen den Ergebnissen in Deutsch und Mathematik ($R^2 = 0.37$). Deutsch ist nicht nur die Unterrichtssprache, sondern auch die Sprache, in der die Mathematikaufgaben formuliert sind, was diesen vergleichsweise engen Zusammenhang erklären dürfte. Mathematische Probleme stellen sich meist in einem sprachlichen Kontext, weshalb gute Sprachkompetenzen für das Lösen von Mathematikaufgaben ein Vorteil sind.

⁵ 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>

⁶ IBE (2023). *Check P5 2023: Ergebnisbericht für den Bildungsraum Nordwestschweiz*. https://www.bildungsraum-nw.ch/schwerpunkte/volksschule/ergebnisberichte_checks/copy3_of_ergebnisberichte2020/kantonsbericht-check-p5-2023.pdf

Die Zusammenhänge zwischen Deutsch und Englisch sowie zwischen Englisch und Mathematik sind wesentlich schwächer als der Zusammenhang zwischen Deutsch und Mathematik. Auch hier könnten mögliche Erklärungsansätze in generellen sprachlichen Fähigkeiten liegen. Die Zusammenhänge deuten zudem darauf hin, dass die Ergebnisse in einem schulischen Leistungstest neben den Fähigkeiten der Schülerinnen und Schüler auch von anderen Einflussfaktoren abhängen, die von dem zu erfassenden Kompetenzbereich nicht immer vollkommen getrennt werden können.

Abbildung 3: Zusammenhang zwischen Ergebnissen der drei geprüften Fächer



2.4 Ergebnisse nach Kompetenzstufen

Zur inhaltlichen Interpretation lassen sich alle Ergebnisse aus *Check Dein Wissen* einer Kompetenzstufe zuordnen. Die Kompetenzstufen reichen beim Check P4 von II bis IX und umschreiben, welche Aufgaben eine Schülerin oder ein Schüler in der Regel sicher lösen kann, beziehungsweise welche Kompetenzen sie oder er beherrscht. Aufgaben der niedrigeren Kompetenzstufen werden ebenfalls sicher gelöst, während Aufgaben der höheren Kompetenzstufen noch nicht oder nur teilweise bewältigt werden können.

Die im Lehrplan 21 definierten Mindestansprüche und Orientierungspunkte können in den Kompetenzstufen verortet werden. Für die 4. Klasse ist der Orientierungspunkt im 2. Zyklus von Interesse. Er beschreibt die Kompetenzen, die bis zum Ende der 4. Klasse

behandelt werden sollen. Schülerinnen und Schüler, die die bis zum Orientierungspunkt behandelten Themen beherrschen, erreichen im Check P4 die Kompetenzstufe V.

Die ebenfalls im Lehrplan 21 festgelegten Grundansprüche definieren Kompetenzstufen, die die Schülerinnen und Schüler bis zum Ende des 2. Zyklus erreichen sollten. 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 Frühling und somit vor Ende der 4. Klasse stattfand. Deshalb ist es möglich, dass einzelne Themen, die bis zum Orientierungspunkt vorgesehen sind, zum Zeitpunkt des Check P4 noch nicht behandelt wurden.

Tabelle 6 zeigt die Verteilung der Schülerinnen und Schüler über die Kompetenzstufen II bis IX in Prozent. Eine grafische Darstellung der Anteile der Schülerinnen und Schüler, welche mindestens die Kompetenzstufe V erreichten, ist in Kapitel 3.3 zu finden.

Tabelle 6: Verteilung der Kompetenzstufen in den geprüften Kompetenzbereichen

Fach	Kompetenzbereich	Kompetenzstufen (%)							
		II	III	IV	V*	VI	VII	VIII	IX
Deutsch	Lesen	7.3	15.1	14.6	30.9	24.4	5.8	1.9	
	Sprache im Fokus	3.8	13.4	22.7	24.1	27.3	7.3	1.5	
Englisch	Hören		1.2	15.1	32.0	29.6	15.4	5.0	1.7
	Lesen	0.1	2.4	19.0	25.9	25.3	17.6	6.5	3.2
Mathematik	Zahl und Variable	16.6	9.5	13.8	24.5	26.6	7.2	1.9	
	Form und Raum	6.4	14.0	20.2	21.6	23.0	10.1	4.3	0.4
	Grössen, Funktionen, Daten und Zufall	16.3	11.3	9.7	30.3	19.4	10.0	2.7	0.4

* Die Kompetenzstufe V entspricht dem Erreichen des Orientierungspunktes im Zyklus 2 nach Lehrplan 21

In *Deutsch Lesen* erreichen rund 63% der Schülerinnen und Schüler mindestens die Kompetenzstufe V. Diese umfasst das Verstehen verschiedener einfacher Texte, deren Themen und Wortschatz dem Weltwissen entsprechen (z. B. kurze Sachtexte, kindgerechte Reportagen, Ausschnitte aus Kinderbüchern). Schülerinnen und Schüler auf der Kompetenzstufe V beherrschen auch Kompetenzen der vorhergehenden Stufen und können beispielsweise einfachen diskontinuierlichen Texten (z. B. Einkaufsliste) explizite Informationen entnehmen (Kompetenzstufe III). Sie beherrschen aber noch nicht oder erst teilweise die Kompetenzen der höheren Kompetenzstufen, wie das Entnehmen von Hauptaussage und Detailinformationen aus anspruchsvolleren Texten (z. B. Kurzgeschichte oder Sachtext mit Diagrammen; Kompetenzstufe VII).

In *Sprache im Fokus* erreichen oder übertreffen rund 60% der Schülerinnen und Schüler die Kompetenzstufe V. Diese beinhaltet unter anderem das Bestimmen typischer Nomen, Verben und Adjektive, das Anwenden grundlegender Rechtschreibregeln und das Setzen von Kommas in Aufzählungen. Ebenfalls beherrscht werden Kompetenzen der vorhergehenden Stufen wie beispielsweise die Steigerung regelmässiger Adjektive (Kompetenzstufe IV) und das korrekte Schreiben von Wörtern wie «Tropfen» oder «Krokodil» (Kompetenzstufe III). Kompetenzen höherer Stufen, etwa das sichere Bestimmen von Zeiformen (Kompetenzstufe VI) oder das Setzen von Kommas zwischen Verbgruppen (Kompetenzstufe VII), werden hingegen noch nicht durchgängig beherrscht.

Im Fach Englisch wird die Kompetenzstufe V von besonders vielen Schülerinnen und Schülern erreicht oder überschritten: rund 84% in *Englisch Hören* und rund 79% in *Englisch Lesen*. In *Englisch Hören* können Schülerinnen und Schüler auf diese Kompetenzstufe kurze Gespräche über vertraute Themen verstehen. Zudem können sie grundlegende Informationen aus einfachen, kurzen Hörtexten entnehmen, wenn langsam und deutlich gesprochen wird. Die Schülerinnen und Schüler können unter anderem auch einfache Fragen zur eigenen oder zu anderen Personen verstehen (Kompetenzstufe IV). Sie können aber noch nicht oder nur teilweise längeren Gesprächen und Hörtexten folgen bzw. vertraute Themen darin erkennen (Kompetenzstufe VIII). In *Englisch Lesen* beinhaltet die Kompetenzstufe V, dass Schülerinnen und Schüler einfache, kurze Texte und Informationsmaterialien (z. B. Comic, SMS, Rezept) verstehen bzw. grundlegende Aussagen daraus entnehmen. Sie verstehen aber noch nicht oder erst teilweise authentische Texte (z. B. Reportage) und persönliche Texte (z. B. Brief, Blog), selbst wenn diese klar strukturiert sind (Kompetenzstufe VIII).

Die Kompetenzstufe V wird in den Kompetenzbereichen der Mathematik ähnlich oft erreicht. In *Zahl und Variable* sind es rund 60% der Schülerinnen und Schüler. Diese beherrschen unter anderem das schriftliche Addieren und Subtrahieren, können für Addition und Subtraktion Rechenwege notieren und Zahlen bis 1 Million ordnen. Zudem können sie bereits Umkehroperationen nutzen, um Aufgaben in Teilschritten zu lösen (Kompetenzstufe IV) und Lösungen für Ungleichungen bestimmen (Kompetenzstufe III). Sie können aber beispielsweise noch nicht oder nicht sicher Zahlen mit vier Wertziffern addieren und subtrahieren (z. B. $183 + 2401$; Kompetenzstufe VI) oder Dezimalzahlen ordnen, addieren und subtrahieren (Kompetenzstufe VII).

In *Form und Raum* erreichen rund 59% der Schülerinnen und Schüler die Kompetenzstufe V. Sie können unter anderem Symmetrieachsen erkennen und Spiegelbilder von Figuren skizzieren. Ebenfalls können sie bereits in einem Koordinatensystem Punkte beschreiben oder Figuren anhand von Koordinaten einzeichnen (Kompetenzstufe IV). Sie können allerdings noch nicht oder nicht sicher mehrere Symmetrieachsen in einer Figur erkennen (Kompetenzstufe VI) oder die Fläche von Rechtecken berechnen (Kompetenzstufe VII).

Im Kompetenzbereich *Grössen, Funktionen, Daten und Zufall* wird die Kompetenzstufe V von rund 63% der Schülerinnen und Schüler erreicht. Diese Schülerinnen und Schüler können Grössen in andere Masseinheiten umwandeln, damit einfache Rechnungen durchführen und einfache Textaufgaben mit Grössen lösen. Ebenfalls beherrschen sie bereits die Grundoperationen mit Geldbeträgen (Kompetenzstufe IV) und können lineare und nichtlineare Zahlenfolgen fortsetzen (Kompetenzstufe IV). Noch nicht oder nicht sicher können Schülerinnen und Schüler auf der Kompetenzstufe V proportionale Beziehungen anwenden (z. B. Fr. / kg oder l / km; Kompetenzstufe VI) oder Funktionswerte aus Graphen auslesen (Kompetenzstufe VII).

Zusammenfassend ist festzuhalten, dass in Deutsch und Mathematik etwa 60% der Schülerinnen und Schüler die Kompetenzen bis zum Orientierungspunkt des 2. Zyklus beherrschen oder übertreffen. In Englisch liegt dieser Anteil sogar bei über 80% der Schülerinnen und Schüler. 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 gibt einen Überblick über die Unterschiede zwischen den teilnehmenden Klassen. Die Klassenergebnisse entsprechen dem Mittelwert der Ergebnisse der Schülerinnen und Schüler mit regulären Lernzielen. Für Klassen mit weniger als fünf Ergebnissen wurde kein Mittelwert berechnet. Dadurch wurden vier kleine Klassen (in Deutsch fünf

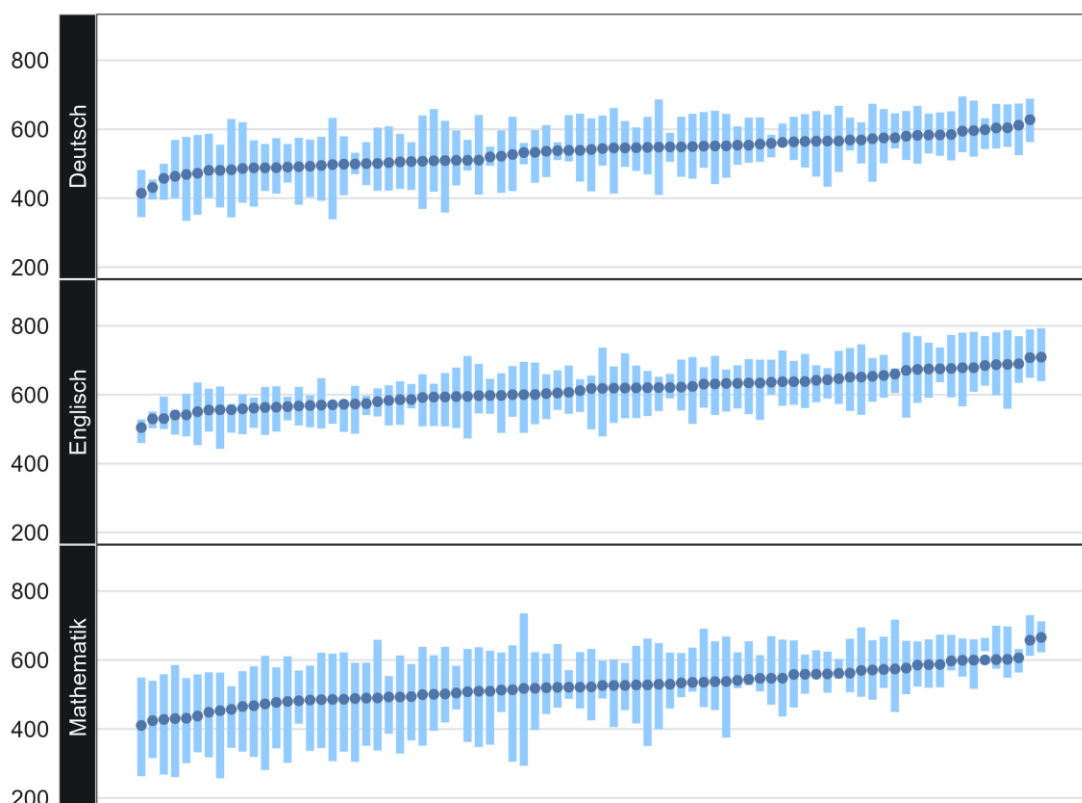
Klassen) von der Auswertung ausgeschlossen. In den verbleibenden 81 (bzw. 80) Klassen liegen jeweils Ergebnisse von 5 bis 21 Schülerinnen und Schülern vor. Die durchschnittliche Klassengrösse beträgt 13.8 Schülerinnen und Schüler.

Die Unterschiede zwischen den Klassen fallen deutlich aus. Die Differenz zwischen dem niedrigsten und dem höchsten Klassenmittelwert beträgt in Deutsch 213 Punkte, in Englisch 204 Punkte und in Mathematik 255 Punkte. Zur Quantifizierung dieser Unterschiede wurde der Varianzanteil berechnet, der auf die Klassen entfällt. Ein Varianzanteil von 0% würde bedeuten, dass die Ergebnisse vollständig unabhängig von der Klassenzugehörigkeit sind, ein Wert von 100% würde hingegen bedeuten, dass die Ergebnisse ausschliesslich durch die Klassenzugehörigkeit bestimmt werden.

Der Varianzanteil der Klassen beträgt in Deutsch 12.3%, in Englisch 15.7% und in Mathematik 11.9%. Die Ergebnisse in Englisch variieren damit etwas stärker zwischen den Klassen als in Deutsch und Mathematik. Die Werte liegen insgesamt unter jenen des Vorjahres (2024: Deutsch 20.1%, Englisch 16.0%, Mathematik 18.0%). Mögliche Gründe für diese Angleichung könnten unter anderem in der Einführung des Check P4 liegen.

Offen bleibt allerdings, in welchem Ausmass die Unterschiede zwischen den Klassen auf individuelle Merkmale der Schülerinnen und Schüler oder auf Faktoren der Klasse, der Lehrpersonen bzw. der Schule zurückzuführen sind, da zentrale Lernvoraussetzungen wie die soziale Zusammensetzung der Klasse in den Analysen nicht berücksichtigt werden konnten.

Abbildung 4: Mittelwert und mittlere 50% der Klassenergebnisse, aufsteigend sortiert



Die ausgeprägten Unterschiede zwischen den Klassen verdeutlichen, weshalb eine unabhängige Standortbestimmung als Ergänzung zur klassenbezogenen Beurteilung anhand von Noten hilfreich ist. Die Ergebnisse des Check P4 ermöglichen eine objektive

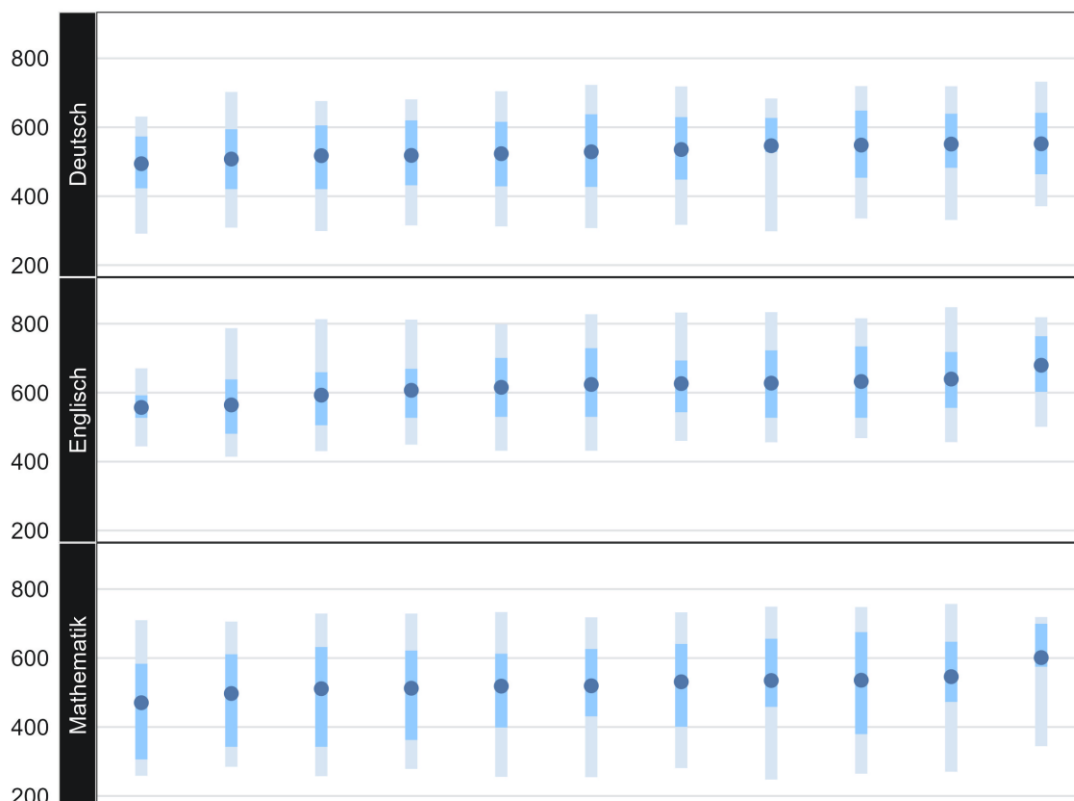
Einschätzung, die nicht vom spezifischen Klassenkontext abhängt, und bieten damit eine zusätzliche Grundlage für eine faire und ganzheitliche Beurteilung der Schülerinnen und Schüler.

2.6 Unterschiede zwischen Gemeinden

Abbildung 5 zeigt die mittleren Ergebnisse sowie die Streuung der Ergebnisse (mittlere 50% und 90%) pro Gemeinde. Dabei wurden auch die Ergebnisse von Schülerinnen und Schülern mit angepassten Lernzielen berücksichtigt. Pro Gemeinde sind Ergebnisse von 18 bis 234 Schülerinnen und Schülern verfügbar (Mittelwert 105). Die Leistungsunterschiede zwischen den Gemeinden sind deutlich geringer als jene zwischen den Schulen. Die Differenz zwischen dem niedrigsten und dem höchsten Gemeindemittelwert beträgt in Deutsch 58 Punkte, in Englisch 122 Punkte und in Mathematik 131 Punkte. Der vergleichsweise grosse Unterschied in Englisch und Mathematik lässt sich vor allem auf eine einzelne Gemeinde zurückführen, die in diesen beiden Fächern – nicht jedoch in Deutsch – auffallend hohe Mittelwerte erzielt.

Insgesamt überlappen sich die Ergebnisse verschiedener Gemeinden stark. Ähnlich wie beim Vergleich der Klassenmittelwerte (vgl. Kapitel 2.5) lässt sich auch hier der Varianzanteil bestimmen, der auf die Gemeinden zurückzuführen ist. Dieser beträgt in Deutsch 1.7%, in Englisch 4.3% und in Mathematik 1.9%. Damit unterscheiden sich die Ergebnisse in Englisch etwas stärker zwischen den Gemeinden als in Deutsch und Mathematik. In allen drei Fächern ist jedoch insgesamt nur ein geringer Einfluss der Gemeinden erkennbar. Im Vergleich zum Vorjahr ist der Einfluss der Gemeinden auf die Ergebnisse in Deutsch stabil geblieben (2024: 1.9%), in Englisch angestiegen (2024: 2.4%) und in Mathematik zurückgegangen (2024: 2.6%).

Abbildung 5: Mittelwert, mittlere 50% (blau) und mittlere 90% (hellblau) der Ergebnisse pro Gemeinde im Kanton Zug, aufsteigend sortiert



3 Ergebnisse pro Schulfach im Trend

Der Check P4 wurde im Jahr 2025 im Kanton Zug zum zweiten Mal durchgeführt, erstmals als reguläre Durchführung. Die Ergebnisse sind über die beiden Jahre hinweg pro Kompetenzbereich und Fach direkt vergleichbar. Eine Punktzahl – beispielsweise 520 Punkte in Deutsch – hat somit in beiden Jahren die gleiche Bedeutung. Dadurch lassen sich sämtliche Ergebnisse mit dem Vorjahr vergleichen.

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 Jahre 2024 und 2025. Die Verteilung der Ergebnisse (Mittelwert, mittlere 50% und 90%) ist zusätzlich in Abbildung 6 dargestellt.

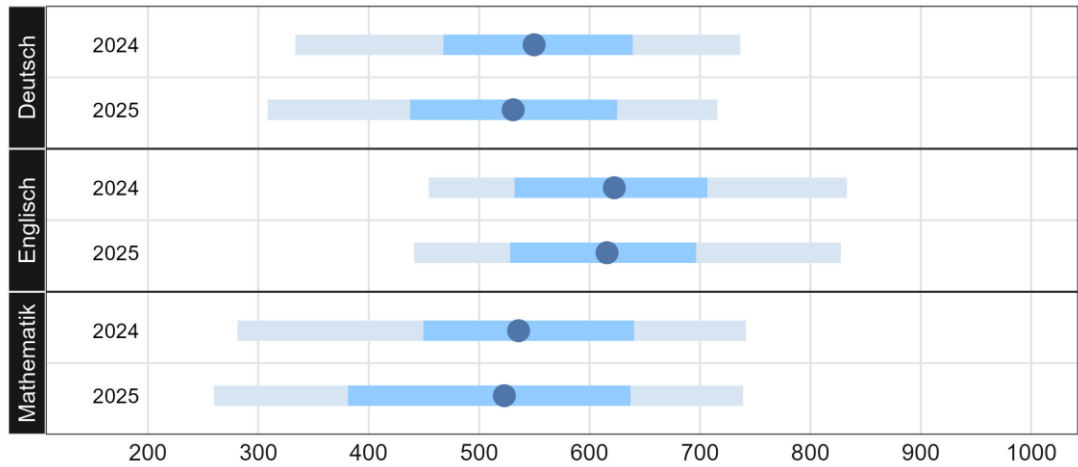
In allen drei Fächern ist ein leichter Rückgang des Mittelwerts zu beobachten. Die Effektstärken liegen zwischen $d = 0.09$ und $d = 0.15$ und entsprechen damit sehr geringen bis vernachlässigbaren Effekten. Auch im Vergleich zum erwarteten Fortschritt pro Semester sind die Unterschiede eher klein. In Deutsch entspricht die Differenz mit -19 Punkten weniger als der Hälfte des Fortschritts eines Semesters. In Englisch (-7 Punkte) und Mathematik (-13 Punkte) entsprechen die Unterschiede sogar weniger als einem Viertel des erwarteten Fortschritts im jeweiligen Fach.

Tabelle 7: Ergebnisse der Jahre 2024 und 2025 im Vergleich

Fach	Jahr	M	SD	N	Differenz	Cohen's d
Deutsch	2024	550	122	1126		
	2025	531	126	1114	-19	-0.15
Englisch	2024	622	117	1148		
	2025	616	116	1118	-7	-0.06
Mathematik	2024	536	141	1147		
	2025	523	153	1126	-13	-0.09

Da der Check P4 im Jahr 2025 zwei Monate früher stattfand als im Jahr 2024 (März statt Mai), ist ein leichter Rückgang der Punktzahlen im Umfang von etwa einem Drittel eines Semesters zu erwarten. Vor diesem Hintergrund kann geschlossen werden, dass das mittlere Fähigkeitsniveau der Schülerinnen und Schüler in allen drei Fächern gegenüber dem Vorjahr stabil geblieben ist.

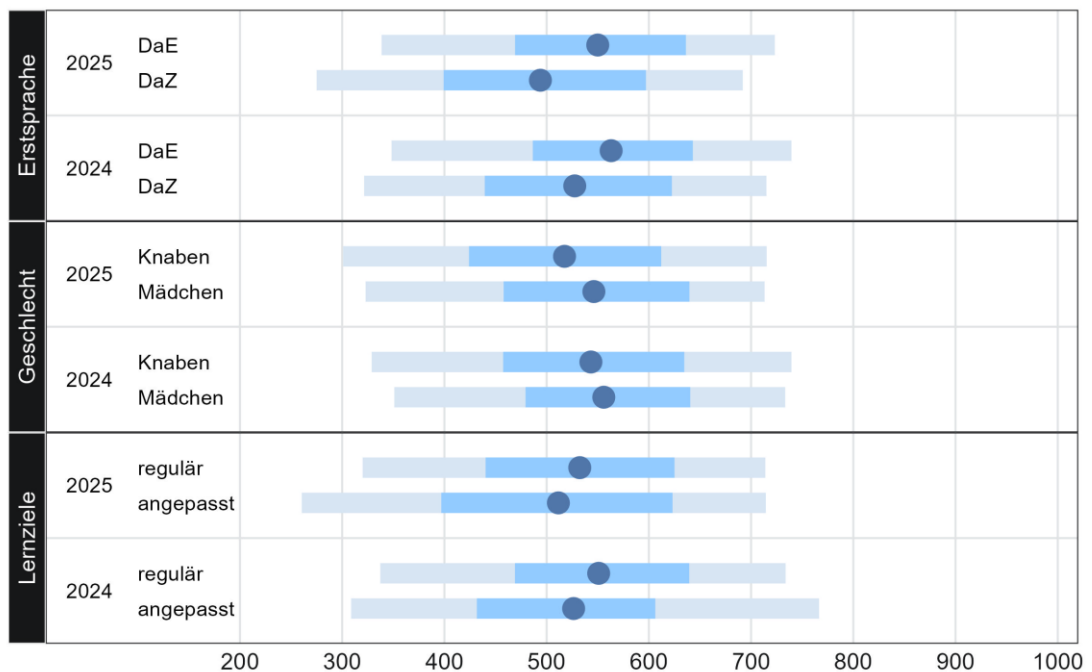
Abbildung 6: Mittelwert, mittlere 50% (blau) und mittlere 90% (hellblau) der Ergebnisse nach Jahr



3.2 Trends nach individuellen Merkmalen

Die Unterschiede zwischen Schülerinnen und Schülern mit unterschiedlicher Erstsprache, Geschlecht oder Lernzielanpassung werden im Folgenden pro Jahr betrachtet. Die Abbildungen 7 bis 9 zeigen die Ergebnisse für die Fächer Deutsch, Englisch und Mathematik. Damit lässt sich beurteilen, ob sich die verschiedenen Gruppen über die Jahre stärker unterscheiden oder ob sie sich einander annähern. Bislang sind keine generellen Aufwärts- oder Abwärtstrends erkennbar (vgl. Kapitel 3.1). Sollten sich künftig Veränderungen zeigen, könnte geprüft werden, ob diese auf bestimmte Merkmalsgruppen der Schülerinnen und Schüler zurückzuführen sind.

Abbildung 7: Mittelwert, mittlere 50% (blau) und mittlere 90% (hellblau) in Deutsch nach Jahr, aufgeteilt nach individuellen Merkmalen



In Deutsch (Abbildung 7) ist eine leichte Zunahme des Unterschieds zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache und als Zweitsprache zu erkennen. Die

Differenz nahm von 35 Punkten ($d = 0.29$) im Jahr 2024 auf 56 Punkte ($d = 0.46$) im Jahr 2025 zu. Während die Schülerinnen und Schüler mit anderer Erstsprache im Jahr 2024 einen Rückstand von etwa drei Vierteln eines Semesters aufwiesen, entsprach dieser Rückstand 2025 bereits mehr als einem Semester.

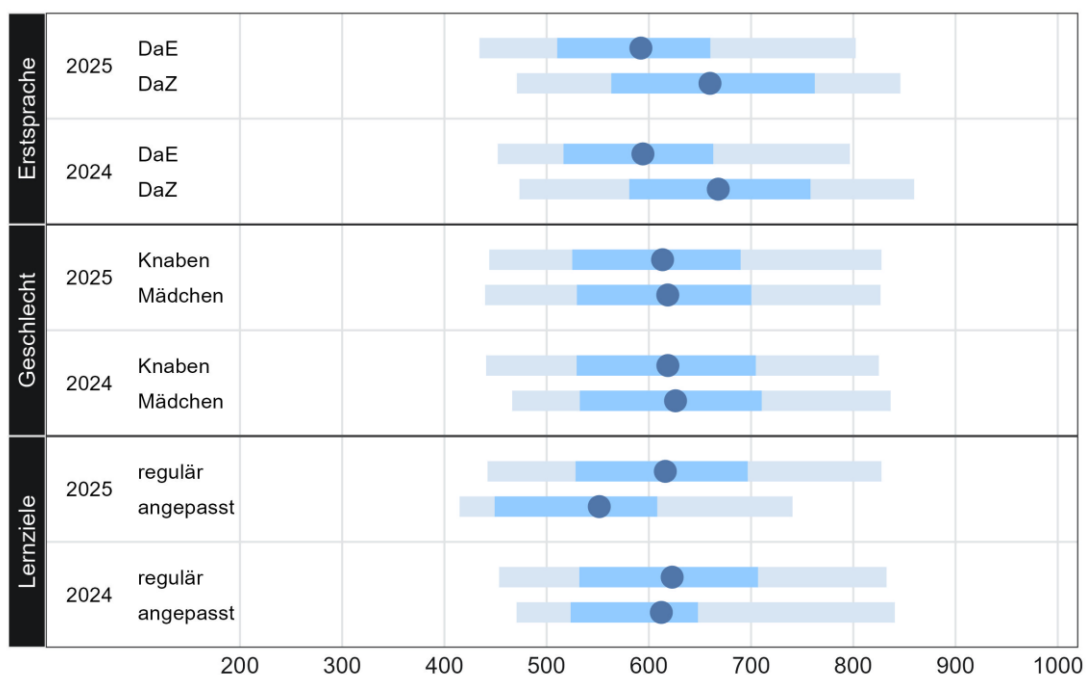
Auch der Unterschied zwischen den Geschlechtern nahm in Deutsch zu. Der Vorsprung der Mädchen gegenüber den Knaben stieg von 13 auf 28 Punkte an. Während der Effekt im Jahr 2024 noch vernachlässigbar war ($d = 0.10$), entspricht die Differenz im Jahr 2025 einem kleinen Effekt ($d = 0.23$). Die Differenz übertrifft damit den Fortschritt eines halben Semesters und die Geschlechterdifferenz im Fach Deutsch kann 2025 als relevant beurteilt werden. Unklar bleibt allerdings aufgrund des unterschiedlichen Durchführungszeitpunktes, ob dieser Anstieg auf eine Leistungsabnahme der Knaben oder auf eine Zunahme der Ergebnisse bei den Mädchen zurückzuführen ist. Diese Frage lässt sich erst mit den Ergebnissen im Jahr 2026 klären.

Unterschiede zwischen Schülerinnen und Schülern mit regulären und angepassten Lernzielen sind dagegen wegen der geringen Fallzahlen und der hohen Zahl fehlender Werte in dieser Gruppe nur eingeschränkt aussagekräftig. Für das Fach Deutsch lässt sich aber festhalten, dass der Unterschied zwischen diesen Gruppen im Jahr 2025 mit 25 Punkten ($d = 0.20$) ähnlich ausfällt wie im Vorjahr mit 20 Punkten ($d = 0.17$).

Im Fach Englisch (Abbildung 8) zeigt sich erneut ein deutlicher Vorteil zugunsten der Schülerinnen und Schüler mit einer anderen Erstsprache als Deutsch. Die Differenz von 74 Punkten ($d = 0.66$) im Jahr 2025 liegt auf vergleichbarem Niveau wie im Vorjahr (68 Punkte; $d = 0.60$). Der Unterschied zwischen Mädchen und Knaben ist hingegen mit 6 Punkten ($d = 0.04$) weiterhin sehr gering und vernachlässigbar.

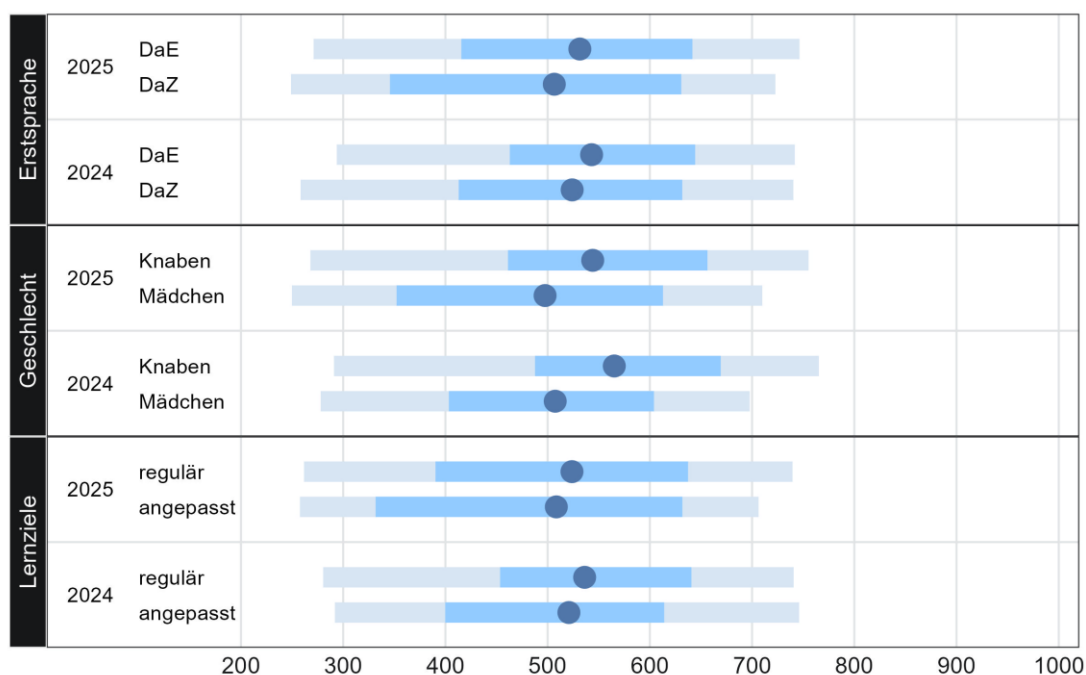
In Englisch ist die Differenz zwischen den Schülerinnen und Schülern mit regulären und jenen mit angepassten Lernzielen von 11 Punkten ($d = 0.09$) auf 64 Punkte ($d = 0.55$) angestiegen. Dieses Ergebnis ist jedoch kaum belastbar, da lediglich sieben Schülerinnen und Schüler mit angepassten Lernzielen in Englisch eine Punktzahl erreichten.

Abbildung 8: Mittelwert, mittlere 50% (blau) und mittlere 90% (hellblau) in Englisch nach Jahr, aufgeteilt nach individuellen Merkmalen



Im Fach Mathematik (Abbildung 9) ist eine geringe Zunahme des Vorsprungs von Schülerinnen und Schülern mit Deutsch als Erstsprache im Vergleich zu jenen mit Deutsch als Zweitsprache zu erkennen. Er betrug 19 Punkte ($d = 0.13$) im Jahr 2024 und 25 Punkte ($d = 0.16$) im Jahr 2025. Beide Differenzen entsprechen nur kleinen Effekten und deuten noch nicht auf ein Auseinanderdriften der Sprachgruppen hin.

Abbildung 9: Mittelwert, mittlere 50% (blau) und mittlere 90% (hellblau) in Mathematik nach Jahr, aufgeteilt nach individuellen Merkmalen



Der Vorsprung der Knaben gegenüber den Mädchen ging leicht zurück: von 58 Punkten ($d = 0.42$) im Jahr 2024 auf 46 Punkte ($d = 0.31$) im Jahr 2025. In beiden Jahren entspricht die Differenz etwa dem erwarteten Fortschritt während eines Semesters. Ob es sich um einen Trend handelt, bei dem sich die Ergebnisse der Geschlechtergruppen langsam annähern, oder lediglich um eine unbedeutende Schwankung, wird sich in den Ergebnissen der kommenden Jahre zeigen.

Der Unterschied zwischen Schülerinnen und Schülern mit regulären und angepassten Lernzielen blieb in Mathematik über beide Jahre hinweg stabil bei 15 Punkten, mit einer nahezu unveränderten Effektstärke ($d = 0.11$ im Jahr 2024; $d = 0.10$ im Jahr 2025). Beide Werte sind sowohl im Hinblick auf die Effektgrösse als auch auf den erwarteten Lernfortschritt pro Semester als sehr gering einzustufen.

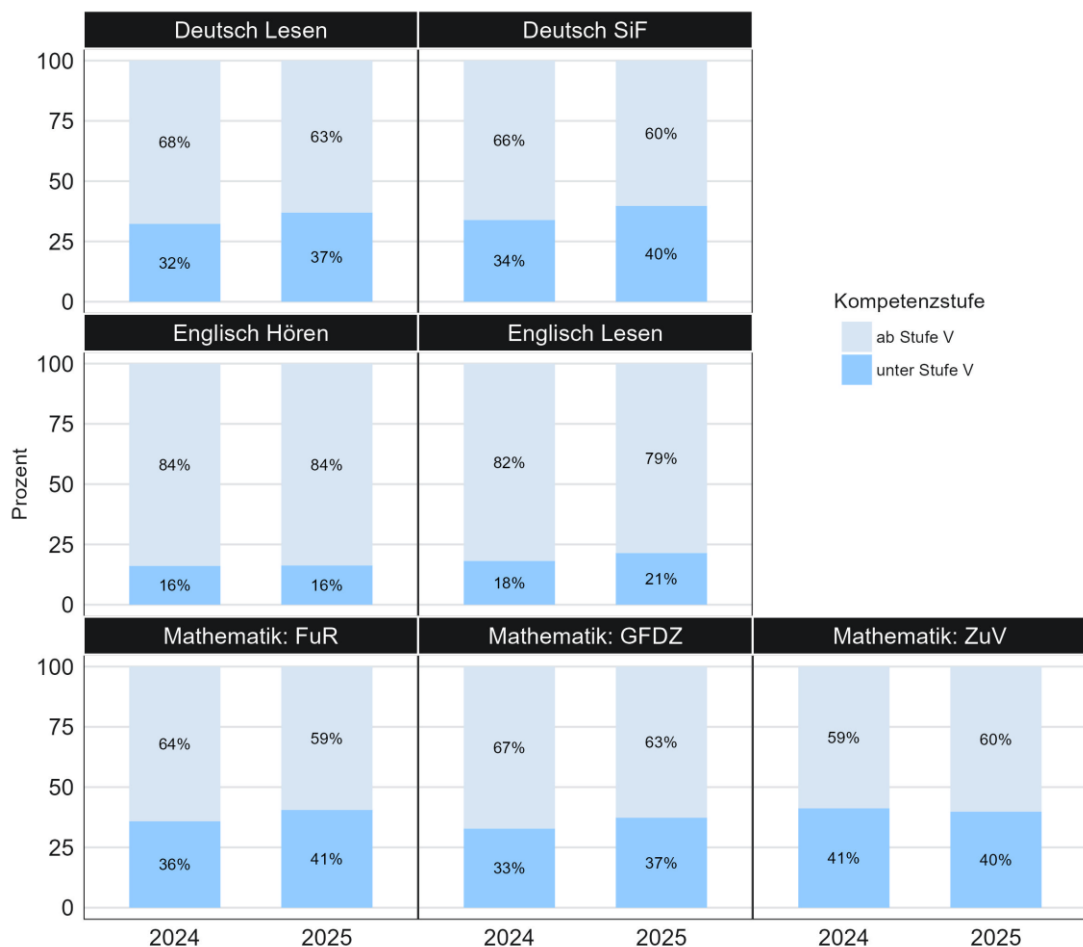
Zusammenfassend zeigen die Trends nach individuellen Merkmalen, dass sich die Unterschiede zwischen Schülerinnen und Schülern mit Deutsch als Erst- oder Zweitsprache sowie zwischen den Geschlechtern in Deutsch leicht vergrößert haben. Auch in Mathematik nahm der Vorsprung der Schülerinnen und Schüler mit Deutsch als Erstsprache leicht zu, während der Vorsprung der Knaben etwas abnahm. In Englisch blieben die Differenzen über beide Jahre hinweg stabil. Die Ergebnisse zu den Schülerinnen und Schülern mit regulären beziehungsweise angepassten Lernzielen sind aufgrund der geringen Fallzahlen nur eingeschränkt interpretierbar. Insgesamt deuten die Ergebnisse auf weitgehend stabile Gruppenunterschiede hin.

3.3 Trend in der Erreichung der Kompetenzstufen

In Kapitel 2.5 wurde dargestellt, wie häufig die Schülerinnen und Schüler gemäss den Ergebnissen des Check P4 verschiedene Kompetenzstufen erreichen. Auch diese Ergebnisse können über mehrere Jahre verglichen werden. Um die Übersicht zu erhöhen, wird hier auf das Erreichen der Kompetenzstufe V fokussiert, die im Lehrplan 21 als Orientierungspunkt im 2. Zyklus definiert ist. Der Orientierungspunkt markiert alle Themen, die bis zum Ende der 4. Klasse behandelt werden sollen. Da der Check P4 im Frühling der 4. Klasse durchgeführt wird, besteht nicht der Anspruch, dass alle Schülerinnen und Schüler die Kompetenzstufe V erreichen. Die Ergebnisse geben jedoch Auskunft darüber, wie weit sie sich dem Orientierungspunkt zum Zeitpunkt des Checks angenähert haben.

Abbildung 10 zeigt den Anteil der Schülerinnen und Schüler pro Kompetenzbereich, die im aktuellen und im vergangenen Jahr die Kompetenzstufe V oder höher erreichten. In den meisten Kompetenzbereichen – mit Ausnahme von Englisch Hören sowie Mathematik Zahl und Variable – ist der Anteil der Schülerinnen und Schüler auf diesem Niveau um 3 bis 6 Prozent zurückgegangen. Dieser Rückgang war zu erwarten, da der Check P4 im Jahr 2025 zwei Monate früher stattfand als im Vorjahr und damit weniger Unterrichtszeit zur Verfügung stand.

Abbildung 10: Anteile der Schülerinnen und Schüler, die pro Kompetenzbereich Kompetenzstufe V oder höher erreichen, pro Jahr



4 Nutzung von Mindsteps und Ergebnisse im Check P4

Mindsteps ist eine Aufgabensammlung mit über 60'000 kompetenzorientierten Aufgaben, die dazu dient, Potenziale und Lernentwicklung von Schülerinnen und Schülern sichtbar zu machen. *Lernen sichtbar machen* ist die Quintessenz aus über 800 Metaanalysen und 50'000 Studien zu wirksamen Faktoren für den schulischen Lernerfolg von John Hattie⁷. Mindsteps wurde auf der Grundlage der zentralen Prinzipien dieser Studien entwickelt.

Mit Mindsteps können Lehrpersonen den Lernfortschritt in den Kompetenzbereichen der Fächer Deutsch, Englisch und Mathematik verfolgen, das Erreichen der Kompetenzstufen des Lehrplans 21 überprüfen oder Lernkontrollen zu aktuellen Themen im Unterricht erstellen. Mindsteps ist eine Ergänzung zu den standardisierten und unabhängigen Checks und wird im Unterricht entsprechend den Check-Ergebnissen eingesetzt. Das Format der Aufgaben ist bei Mindsteps und bei den Checks identisch, die Art der Auswahl der Aufgaben und die Durchführung unterscheiden sich hingegen. Während die Checks nach einheitlichen und standardisierten Regeln zusammengestellt und durchgeführt werden und dadurch zu einer unabhängigen Standortbestimmung führen, dient Mindsteps dem formativen Lernen und Üben im Unterricht.

Schülerinnen und Schüler haben einen eigenen Zugang zu Mindsteps und können sowohl im Unterricht als auch zuhause Aufgabenserien lösen. Dabei wird zwischen zwei Bereichen unterschieden: Im Bereich *Mind* finden Schülerinnen und Schüler Aufgabenserien, die ihre Lehrpersonen für sie vorbereitet haben. Im Bereich *Steps* können die Schülerinnen und Schüler selbstständig Aufgabenserien lösen, die dort zum Üben und Lernen bereitstehen. Die Lehrpersonen haben Einsicht in die Ergebnisse beider Bereiche.

Mindsteps kann im Unterricht oder zuhause genutzt werden. Sowohl der *Mind*- als auch der *Steps*-Bereich umfassen Aufgaben zu allen Kompetenzen, die auch im Check P4 relevant sind. Darüber hinaus bietet Mindsteps den Schülerinnen und Schülern die Möglichkeit, sich mit dem Aufgabenformat und der Testsituation am Computer vertraut zu machen. Vor diesem Hintergrund stellen sich zwei zentrale Fragen: Wie häufig wurde Mindsteps zur Vorbereitung auf den Check P4 eingesetzt, und in welchem Zusammenhang steht die Nutzung von Mindsteps mit den Ergebnissen im Check P4?

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 Kompetenzbereichen ab Beginn des Schuljahres (1. August 2024) 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 5 Sekunden Zeit nahmen. Hintergrund für dieses Filterkriterium ist, dass Schülerinnen und Schüler auf Mindsteps auch Aufgabenserien ansehen und «durchklicken» können, ohne die Aufgaben tatsächlich zu lösen. Durch dieses Ausschlusskriterium wurden ca. 9% der durchgeführten Aufgabenserien ausgeschlossen, überwiegend im Bereich *Steps*.

Tabelle 8 zeigt pro Fach und Schuljahr, welcher Anteil der Schülerinnen und Schüler der 4. Klassen Mindsteps im genannten Zeitraum mindestens einmal nutzte und wie viele Aufgabenserien durchschnittlich bearbeitet wurden. Dabei ist zu berücksichtigen, dass

⁷ Hattie, John A. C. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating on achievement*. London & New York: Routledge. <https://doi.org/10.4324/9780203887332>

der Check P4 im Jahr 2025 rund zwei Monate früher stattfand als im Jahr 2024, sodass der berücksichtigte Zeitraum im Schuljahr 2024/2025 kürzer ausfiel.

Tabelle 8: Nutzung von Mindsteps im Kanton Zug und Zusammenhang mit Ergebnissen im Check P4

	<i>Mind</i> -Bereich			<i>Steps</i> -Bereich			Beide Bereiche		
	Nutz.	<i>M</i> Anz.	<i>r</i>	Nutz.	<i>M</i> Anz.	<i>r</i>	Nutz.	<i>M</i> Anz.	<i>r</i>
2023/2024									
Deutsch	79.4%	2.4	0.17	54.6%	3.4	0.18	86.8%	5.8	0.24
Englisch	53.1%	1.2	0.02	41.2%	1.7	0.17	71.3%	2.9	0.14
Mathematik	72.3%	2.1	0.10	61.2%	3.8	0.15	87.9%	6.0	0.19
2024/2025									
Deutsch	56.2%	1.5	0.03	56.7%	4.3	0.12	77.8%	5.8	0.12
Englisch	32.1%	0.9	0.03	44.7%	2.2	0.15	55.8%	3.0	0.13
Mathematik	40.6%	1.2	0.02	62.6%	4.0	0.13	76.9%	5.2	0.12

Nutz.: Anteil Schülerinnen und Schüler, die mindestens eine Aufgabenserie bearbeiteten.

*M*Anz.: Mittlere Anzahl Aufgabenserien pro Schülerin/Schüler

r: Korrelation zwischen Anzahl Aufgabenserien (logarithmiert) und Ergebnis im Check P4

Wie die untere Hälfte der Tabelle 8 zeigt, nutzte auch im Schuljahr 2024/2025 die Mehrheit der Schülerinnen und Schüler Mindsteps mindestens einmal in einem der drei Fächer. In Deutsch (78%) und Mathematik (77%) liegt der Anteil ähnlich hoch, ist aber gegenüber dem Vorjahr um jeweils etwa 10% gesunken. In Englisch ist der Anteil mit 56% tiefer und auch der Rückgang mit rund 15% stärker.

Die Zahl der bearbeiteten Aufgabenserien variiert stark und reicht von 0 (in allen drei Fächern) bis zu maximal 102 in Deutsch, 47 in Mathematik und 43 in Englisch. Die meisten Schülerinnen und Schüler bearbeiteten zwischen drei und sechs Serien pro Fach. Ein Rückgang gegenüber dem Vorjahr ist nur im Fach Mathematik erkennbar, in dem durchschnittlich 0.8 Aufgabenserien weniger gelöst wurden.

Betrachtet man die Bereiche *Mind* und *Steps* separat, so fällt auf, dass der Rückgang der Nutzung ausschliesslich auf den *Mind*-Bereich zurückzuführen ist. Im *Steps*-Bereich stiegen sowohl der Anteil der Nutzerinnen und Nutzer als auch die durchschnittliche Anzahl bearbeiteter Aufgabenserien in allen Fächern an.

Unklar bleibt, in welchem Ausmass die Schülerinnen und Schüler im *Steps*-Bereich aus eigenem Antrieb aktiv waren oder von ihren Lehrpersonen dazu angeleitet wurden. Daraus ergeben sich zwei mögliche Erklärungen für die beobachteten Veränderungen: Einerseits könnte Mindsteps im Vergleich zum Vorjahr bei den Lehrpersonen an Bedeutung verloren haben (Rückgang bei *Mind*), während es bei den Schülerinnen und Schülern oder auch deren Eltern weiterhin beliebt blieb (Zuwachs bei *Steps*). Andererseits könnte sich auch das Nutzungsverhalten der Lehrpersonen verändert haben – etwa indem sie ihre Klassen vermehrt zur individuellen Arbeit im *Steps*-Bereich anregten, statt selbst Aufgabenserien im *Mind*-Bereich vorzubereiten.

4.2 Zusammenhang zwischen Mindsteps und Ergebnissen im Check P4

Tabelle 8 zeigt neben Kennwerten zur Nutzung von Mindsteps auch die Korrelation zwischen der Nutzung und den Ergebnissen im Check P4. Wie erwartet, sind die Zusam-

menhänge in allen drei Fächern positiv: Schülerinnen und Schüler, die Mindsteps häufiger nutzten, erzielten im Check P4 im Mittel etwas bessere Ergebnisse als jene, die selten oder nie Aufgabenserien bearbeiteten. Die Korrelationen liegen im aktuellen Jahr mit $r = 0.12$ bis $r = 0.13$ im niedrigen Bereich und sind etwas geringer als im Vorjahr. Im Fach Deutsch hat der Zusammenhang im Vergleich zum Schuljahr 2023/2024 am stärksten abgenommen.

Abbildung 11: Mittelwerte (orange), mittlere 50% (gelb) und mittlere 90% (hellgelb) der Ergebnisse im Check P4 nach Nutzung von Mindsteps im aktuellen und vergangenen Schuljahr



oft: mehr als 3 Aufgabenserien im jeweiligen Fach
 selten: 1 – 3 Aufgabenserien im jeweiligen Fach
 nie: keine Aufgabenserie im jeweiligen Fach

Um die Relevanz dieses Zusammenhangs besser einschätzen zu können, zeigt Abbildung 11 die Ergebnisse im Check P4 für drei Nutzergruppen: Schülerinnen und Schüler, die Mindsteps nie nutzten (keine Aufgabenserie bearbeitet), jene, die Mindsteps selten nutzten (1 – 3 Aufgabenserien) und solche, die Mindsteps oft nutzten (mehr als 3 Aufgabenserien). Für die Gruppierung wurden jeweils nur Aufgabenserien berücksichtigt, die mindestens zur Hälfte und nicht übereilt beantwortet wurden.

In allen drei Fächern ist zu erkennen, dass Schülerinnen und Schüler, die Mindsteps im Schuljahr 2024/2025 oft nutzten, im Mittel etwas bessere Ergebnisse im Check P4 erzielten als die beiden anderen Gruppen. Gegenüber jenen, die Mindsteps selten nutzten beträgt dieser Vorsprung in Deutsch 15 Punkte ($d = 0.12$), in Englisch 27 Punkte ($d = 0.23$)

und in Mathematik 32 Punkte ($d = 0.21$). Diese Differenzen entsprechen überwiegend kleinen Effekten. In Englisch und Mathematik liegen sie ungefähr im Bereich des erwarteten Lernfortschritts eines halben Semesters.

Vergleicht man die Ergebnisse von Schülerinnen und Schülern, die Mindsteps häufig nutzten, mit jenen, die Mindsteps nie nutzten, betragen die Unterschiede 26 Punkte in Deutsch ($d = 0.20$), 33 Punkte in Englisch ($d = 0.29$) und 50 Punkte in Mathematik ($d = 0.33$). In Englisch und Mathematik entspricht dies in etwa dem durchschnittlichen Lernfortschritt eines Semesters auf der Primarstufe. Die Ergebnisse derjenigen, die Mindsteps nie nutzten, unterscheiden sich hingegen kaum von jenen, die Mindsteps selten nutzten (Deutsch: 11 Punkte, $d = 0.09$; Englisch: 6 Punkte, $d = 0.06$; Mathematik: 19 Punkte, $d = 0.12$).

Im Vergleich zum Vorjahr ist der Zusammenhang in Deutsch und Mathematik schwächer geworden, insbesondere in Deutsch ($r = 0.12$ gegenüber $r = 0.24$ im Vorjahr). Dieses Muster deckt sich mit den veränderten Nutzungsgewohnheiten der Schülerinnen und Schüler: Während häufiges Arbeiten mit mehr als drei Aufgabenserien weiterhin mit besseren Ergebnissen verbunden ist, zeigen sich bei seltener Nutzung (1–3 Serien) kaum Vorteile. Werden die Bereiche *Mind* und *Steps* getrennt betrachtet, so weisen die *Mind*-Daten im aktuellen Schuljahr keine nennenswerten Zusammenhänge mehr mit den Check-Ergebnissen auf ($r = 0.02$ bis $r = 0.03$), was für Englisch bereits im Vorjahr der Fall war. Dagegen haben die Zusammenhänge im *Steps*-Bereich nur leicht abgenommen ($r = 0.12$ – 0.15 im Schuljahr 2024/2025 gegenüber $r = 0.15$ – 0.18 im Schuljahr 2023/2024).

Insgesamt zeigt sich, dass eine häufige Nutzung von Mindsteps – insbesondere im Bereich *Steps* – auch 2025 mit besseren Ergebnissen im Check P4 einhergeht. Die Effekte fallen allerdings überwiegend klein aus und liegen im Vergleich zum Vorjahr etwas tiefer. Bei der Interpretation ist jedoch Vorsicht geboten. Ob die beobachteten Zusammenhänge auf einen direkten kausalen Einfluss der Mindsteps-Nutzung zurückzuführen sind, bleibt offen. Ebenso ist es denkbar, dass Drittvariablen wie beispielsweise das Engagement oder die Haltung der Lehrpersonen sowohl die Nutzung von Mindsteps als auch die Ergebnisse im Check P4 beeinflussen.