

Thüringer Kompetenztest

Landesbericht

Schuljahr 2018/2019

Herausgeber
Projekt kompetenztest.de
unter Leitung von Dr. Christof Nachtigall



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Fakultät für Sozial- und Verhaltenswissenschaften
Lehrstuhl für Methodenlehre und Evaluationsforschung

Inhaltsverzeichnis

1 Testergebnisse auf Landesebene	3
1.1 Teilnahme an den Kompetenztests	3
1.2 Erreichte Gesamtpunktzahlen	4
1.3 Erreichte Gesamtpunktzahlen getrennt nach Schulart	11
1.4 Aufgabenschwierigkeiten	17
2 Schwerpunktthema 2019: Wiederholen von Klassenstufen	62
2.1 Das Konstrukt "Sitzenbleiben"	62
2.2 Forschungsstand über Folgen einer Klassenstufenwiederholung	63
2.3 Neue Untersuchungen	64
2.4 Methoden und Ergebnisse	64
2.5 Fazit	66
3 Rezeption und Evaluation der Tests	68
3.1 Befragung der Lehrkräfte nach der Dateneingabe	68
3.2 Rezeption der Testergebnisse an den Schulen	69
3.2.1 Formen der Rückmeldungen	69
3.2.2 Download-Quoten 2019	70
3.3 Evaluationsbefragung zu den Kompetenztests	73
3.3.1 Vergleich der Einschätzungen von Schulleitern und Fachlehrkräften	75
3.3.2 Vergleich der Einschätzungen nach Schulform	75
3.3.3 Vergleich der Einschätzungen nach Teilnahme am Online-Test versus Pa- piervariante	77
4 Literatur	78
5 Anhang	79

1 Testergebnisse auf Landesebene

In diesem Abschnitt werden die Teilnehmerzahlen und die Ergebnisse der Kompetenztests 2019 auf Landesebene dargestellt. Bei den Ergebnisdarstellungen werden zunächst die Mittelwerte und Häufigkeitsverteilungen des Gesamtergebnisses aller Schüler betrachtet (Abschnitt 1.2). Anschließend erfolgt die Auswertung getrennt nach Schularten (Abschnitt 1.3). In Abschnitt 1.4 befinden sich die Landesergebnisse auf Aufgabenebene.

Für die verschiedenen Kompetenztests werden dabei meist Abkürzungen für die Tests verwendet. Bei MK6 zum Beispiel steht das „M“ für Mathematik, „K6“ für Klassenstufe 6; hierhinter verbirgt sich also der Kompetenztest Mathematik in der Klassenstufe 6. Für die Deutschtests in der Klassenstufe 3 ist noch ein „L“ für den Test Lesen (DK3L) bzw. ein „H“ für den Test Zuhören (DK3H) angehängt.

Aufgrund der sehr geringen Teilnehmerzahlen werden im Fach Französisch keine Auswertungen vorgenommen.

1.1 Teilnahme an den Kompetenztests

Die Kompetenztests 2019 wurden in der Klassenstufe 3 am 15. Mai (Mathematik), am 07. Mai (Deutsch Lesen) und am 09. Mai (Deutsch Zuhören) durchgeführt. In der Klassenstufe 6 erfolgte die Testdurchführung am 27. Februar (Deutsch), 5. März (Mathematik) und 7. März (Englisch). In der Klassenstufe 8 wurde der Kompetenztest in der ersten Fremdsprache (Englisch oder Französisch) am 26. Februar, der Kompetenztest Mathematik am 28. Februar und der Kompetenztest Deutsch am 6. März geschrieben.

Es nahmen insgesamt 48 041 Schüler aus 788 Schulen an den Kompetenztests teil. In Tabelle 1.1 sind die genauen Teilnehmerzahlen, gegliedert nach Klassenstufe und Schulart, aufgeführt.

Tabelle 1.1: Teilnehmerzahlen an den Kompetenztests 2019

Klassenstufe	Schulart	Anzahl Schulen	Anzahl Klassen	Anzahl Schüler
3	Förderschule	7	7	28
	Grundschule	460	916	16 490
6	Förderschule	10	11	54
	Gymnasium	88	241	5 383
	Regelschule	254	516	9 711
8	Förderschule	14	17 ¹	94
	Gymnasium	93	277 ¹	5 996
	Regelschule	252	883 ¹	10 285

¹ In Klassenstufe 8 ist hier die Zahl der Klassen und Kurse angegeben, diese ist nicht direkt vergleichbar mit den Klassenzahlen der anderen Klassenstufen.

Während in den Klassenstufen 3 und 8 die Teilnahme in allen Fächern verpflichtend war, konnten in der Klassenstufe 6 die Tests wahlweise wieder in den Fächern Deutsch, Mathematik und Englisch geschrieben werden. Ein Test war aber auch in dieser Klassenstufe verpflichtend; eine Grenze nach oben wurde dabei nicht gesetzt, sodass auch zwei oder drei Tests geschrieben werden konnten. Dabei galt allerdings die Vorgabe, dass alle Klassen einer Schule in den jeweils gleichen Fächern teilnehmen mussten. Tabelle 1.2 gibt einen detaillierteren Einblick der gewählten Teilnahme in der Klassenstufe 6.

Die Teilnahmequoten ähneln denen des Vorjahres: Der Anteil der Schulen, die in allen Fächern teilnahmen, ist leicht gesunken (von 24 % auf 22 %). Vergleichbar zum Vorjahr wählten 60 % der Schulen „nur“ einen Test. Erneut wurde dabei am häufigsten das Fach Mathematik gewählt.

Tabelle 1.2: Testteilnahme an den verschiedenen Fächerkombinationen in Klassenstufe 6

Testteilnahme an Fächern	Anzahl Schulen	Anzahl Klassen
Nur Deutsch	71	165
Nur Mathematik	114	259
Nur Englisch	27	61
Deutsch und Mathematik	40	82
Deutsch und Englisch	4	7
Mathematik und Englisch	17	42
Deutsch, Mathematik und Englisch	79	152

1.2 Erreichte Gesamtpunktzahlen

Tabelle 1.3 zeigt die erreichten Gesamtpunktzahlen der Kompetenztests 2019 als Mittelwert. Zusätzlich sind die maximal erreichbaren Punkte der jeweiligen Tests angegeben, sowie Streuung und Teilnahmezahl.

Tabelle 1.3: Ergebnisse der Kompetenztests 2019 für die Klassenstufen 3, 6 und 8.

Test	Mittelwert (MW)	Maximalpunktzahl	Streuung (SD)	Schüleranzahl (N)
DK3L	14,02	22	4,45	15 977
DK3H	14,95	22	4,10	15 922
MK3	19,98	34	6,54	15 831
DK6	46,54	71	11,63	7 664
MK6	14,70	27	5,84	10 403
EK6	60,34	100	21,87	4 847
DK8	64,86	103	17,31	14 646
MK8 Heft 1*	27,06	48	8,18	9 870
MK8 Heft 2*	26,52	52	8,53	5 248
EK8	54,00	85	18,48	14 646

* Beim Mathematiktest in Klassenstufe 8 erwies sich das verwendete einheitliche Testheft an einigen Gymnasien als zu leicht. Die Testergebnisse differenzierten bei leistungsstarken Klassen nicht mehr hinreichend deutlich. Seit dem Schuljahr 2017/2018 können Gymnasien zwischen einem normalen und einem schwereren Testheft wählen.

Die folgenden Grafiken geben die erreichte Punktzahl aller Thüringer Schüler für die verschiedenen Tests in Form einer Häufigkeitsverteilung wieder. Die Werte auf der waagerechten Achse geben die erreichte Punktzahl an. Die Höhe der Säulen entspricht der Anzahl von Schülern, welche diese Punktzahl im Test erreicht haben. Der senkrechte Strich in der Verteilung markiert den Landesmittelwert. Die Grafiken zeigen, dass sich die Schülerleistungen in Thüringen über die gesamte Spannweite der erreichbaren Punkte verteilen. Alle Tests beinhalten auch bewusst leichte Aufgaben (siehe Abschnitt 1.4), sodass schwächere Schüler die Möglichkeit haben, Punkte zu erlangen. Für die leistungsstärkeren Schüler bieten die Tests i. d. R. ebenso genügend Differenzierungsmöglichkeiten.

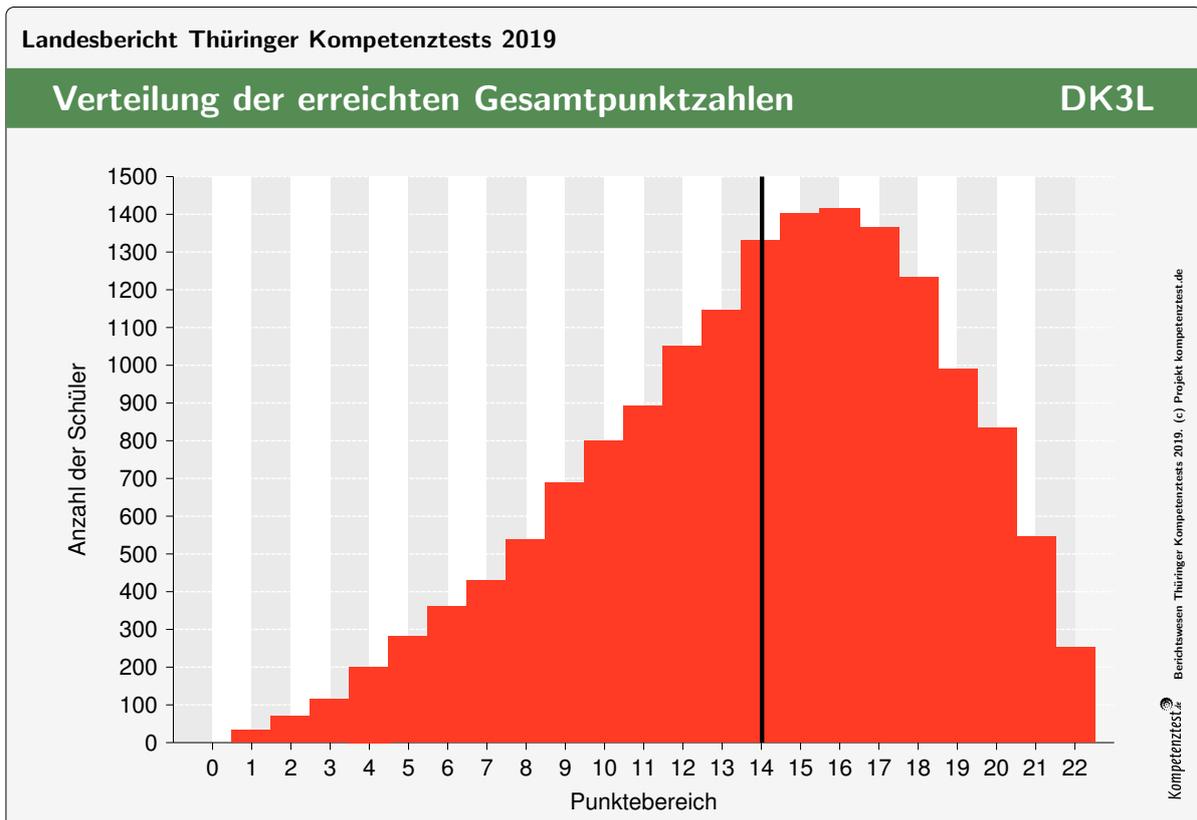


Abbildung 1.1: Verteilung der Gesamtpunktzahlen im Fach Deutsch Lesen, Klassenstufe 3

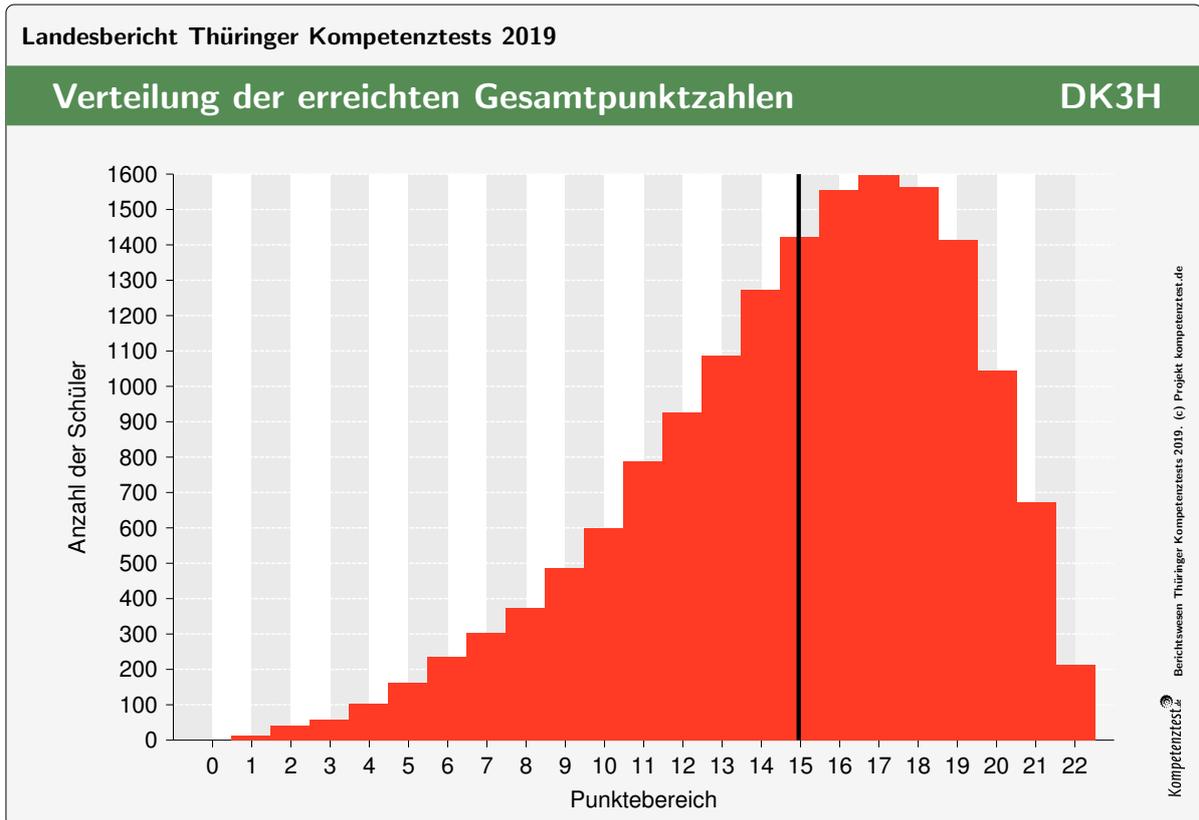


Abbildung 1.2: Verteilung der Gesamtpunktzahlen im Fach Deutsch Zuhören, Klassenstufe 3

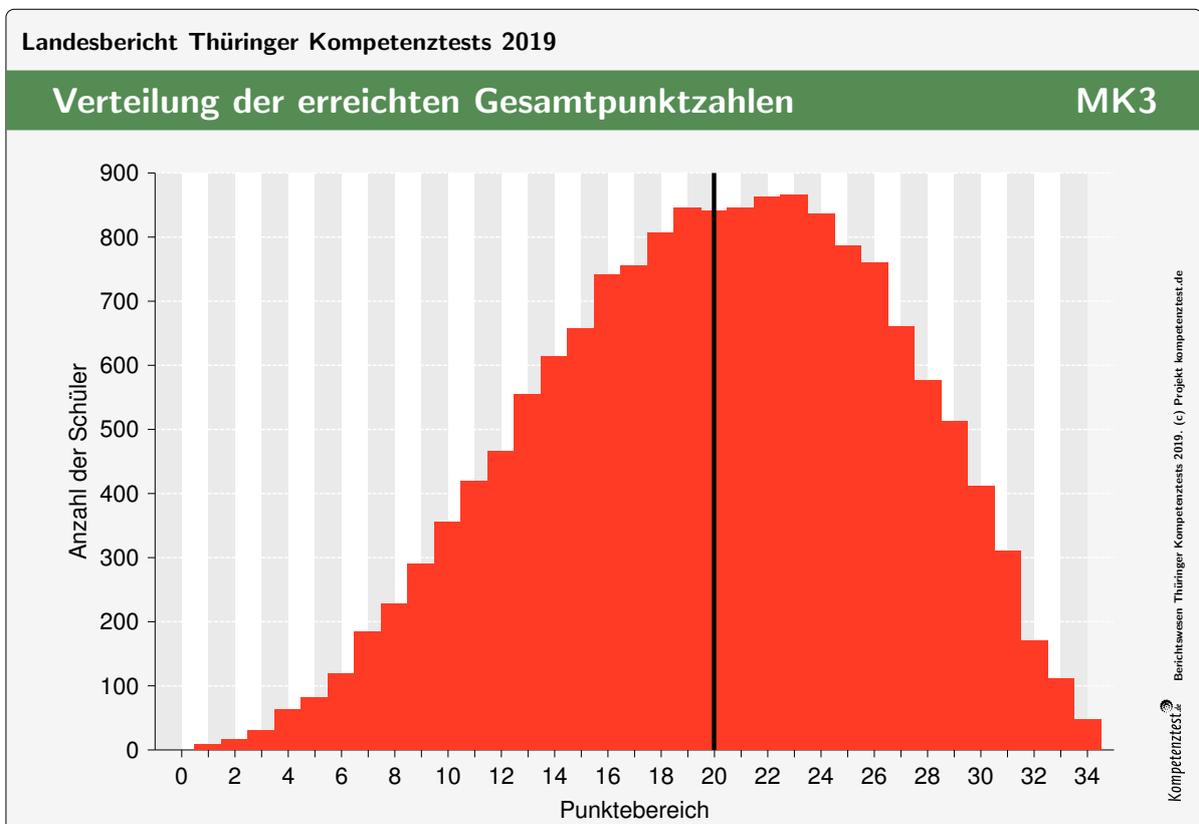


Abbildung 1.3: Verteilung der Gesamtpunktzahlen im Fach Mathematik, Klassenstufe 3

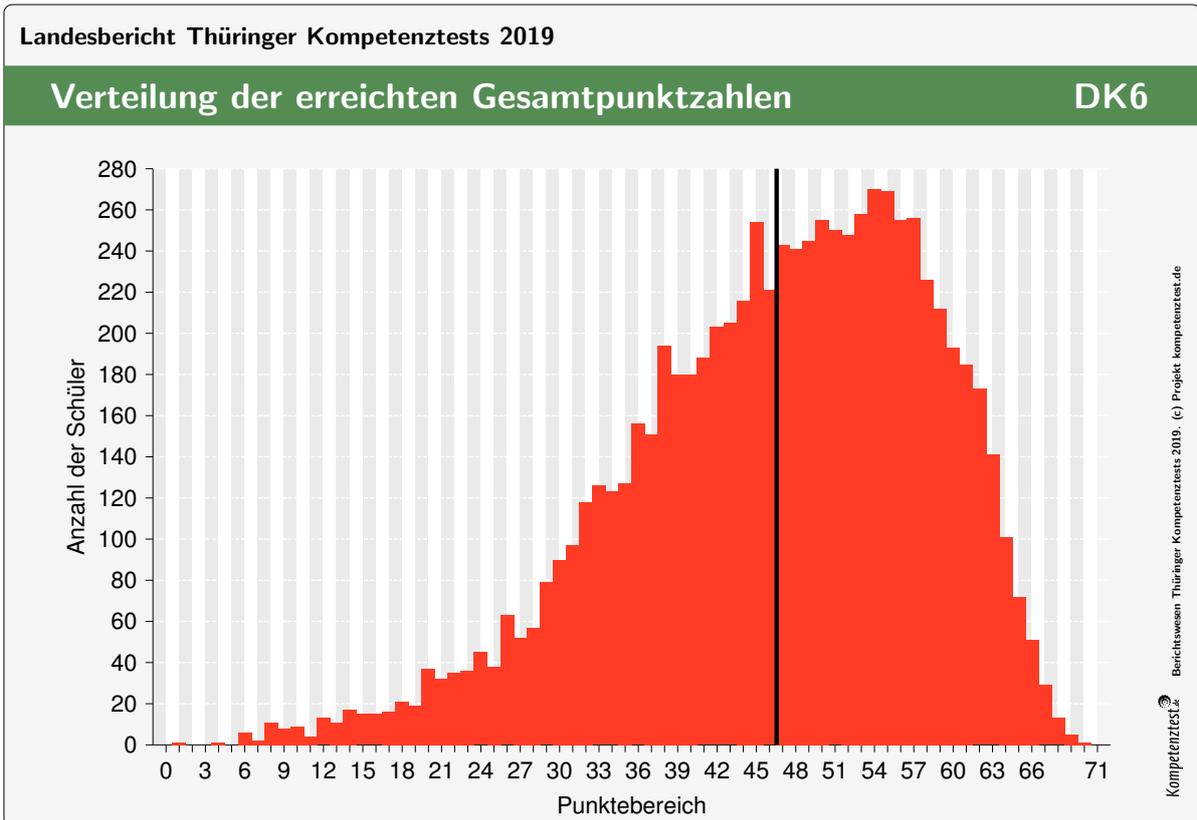


Abbildung 1.4: Verteilung der Gesamtpunktzahlen im Fach Deutsch, Klassenstufe 6

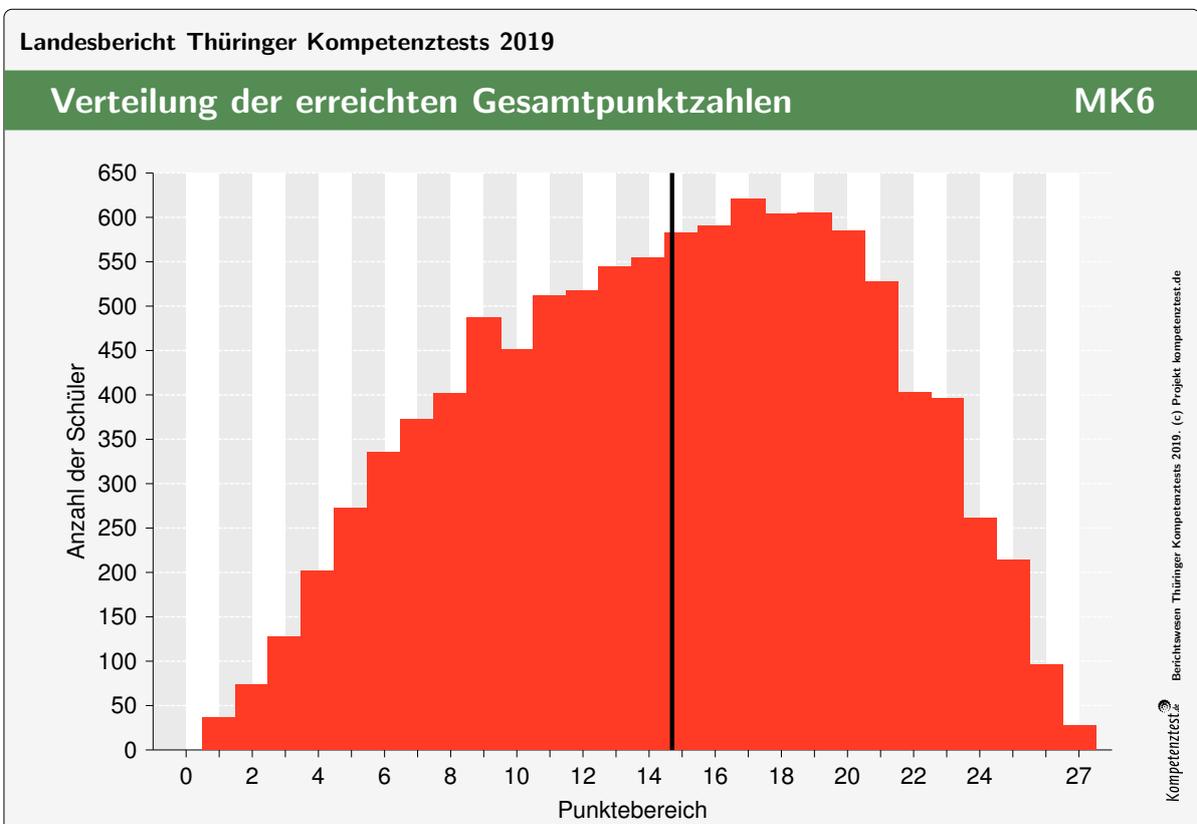


Abbildung 1.5: Verteilung der Gesamtpunktzahlen im Fach Mathematik, Klassenstufe 6

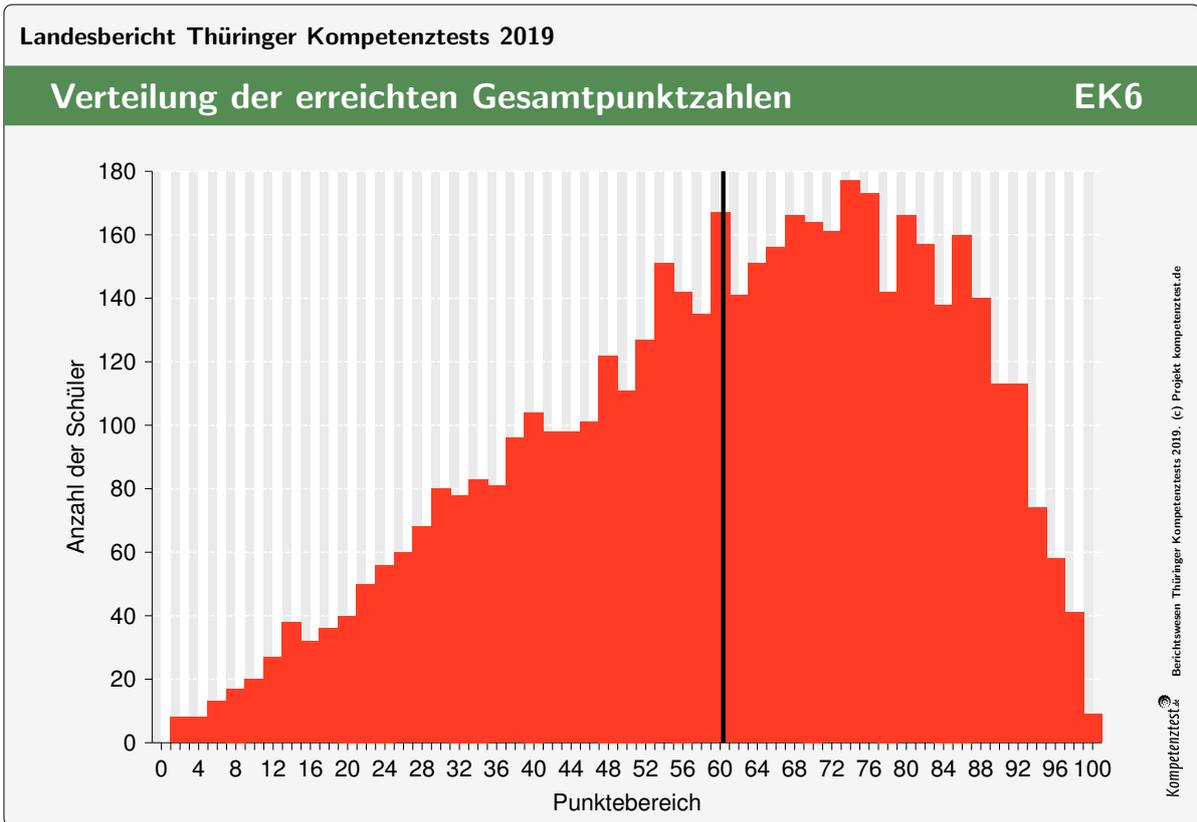


Abbildung 1.6: Verteilung der Gesamtpunktzahlen im Fach Englisch, Klassenstufe 6

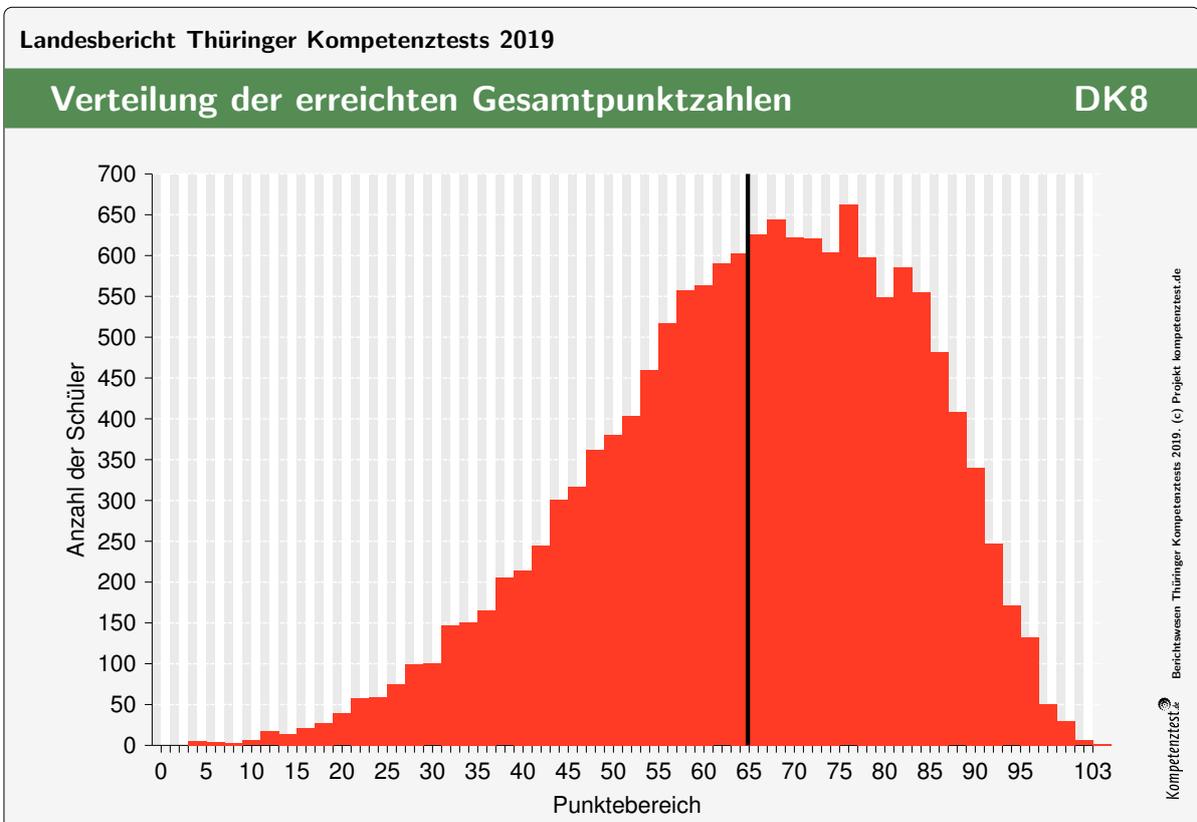


Abbildung 1.7: Verteilung der Gesamtpunktzahlen im Fach Deutsch, Klassenstufe 8

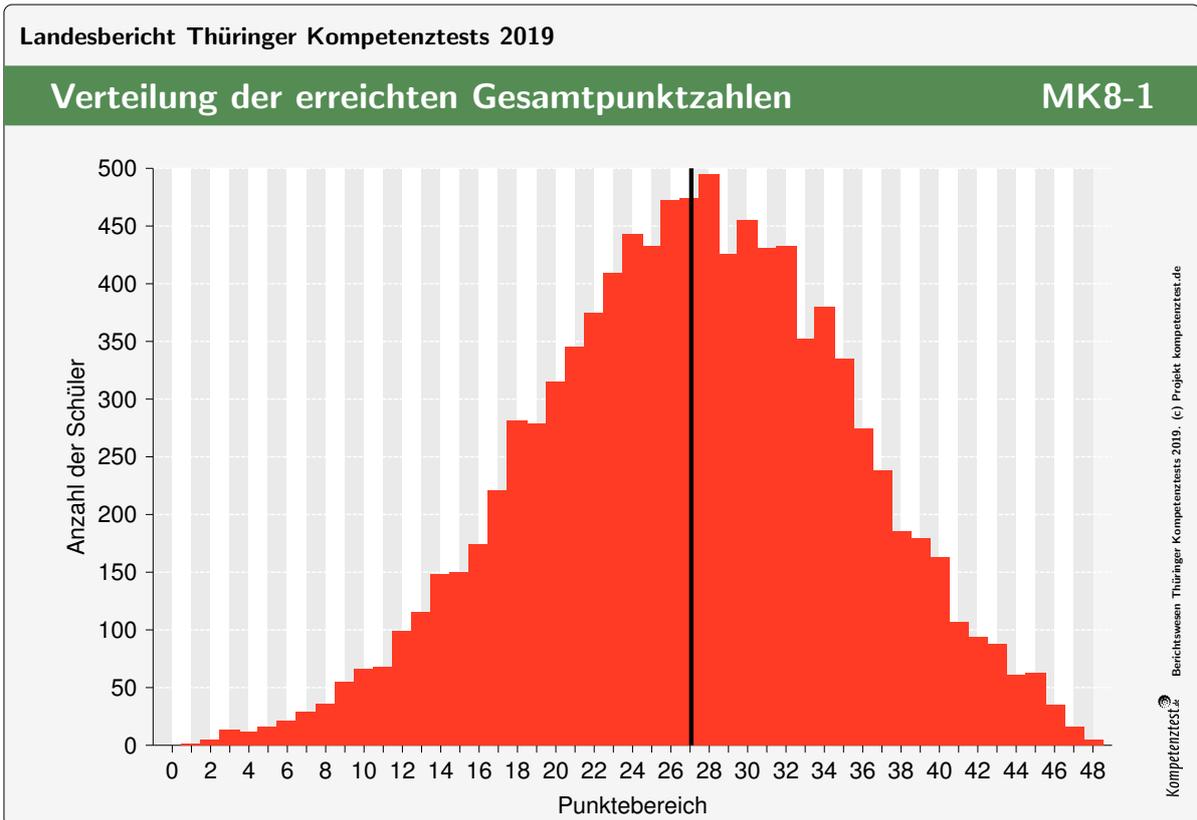


Abbildung 1.8: Verteilung der Gesamtpunktzahlen im Fach Mathematik, Heft 1, Klassenstufe 8

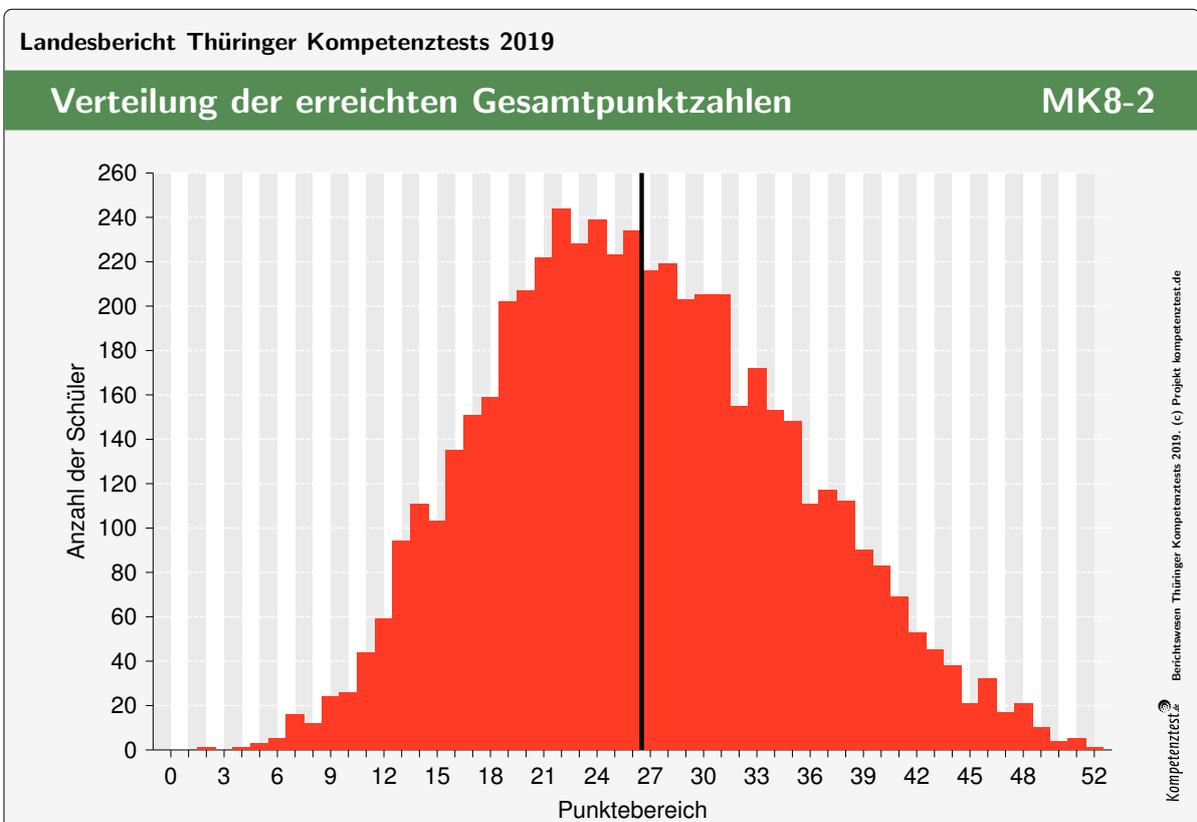


Abbildung 1.9: Verteilung der Gesamtpunktzahlen im Fach Mathematik, Heft 2, Klassenstufe 8

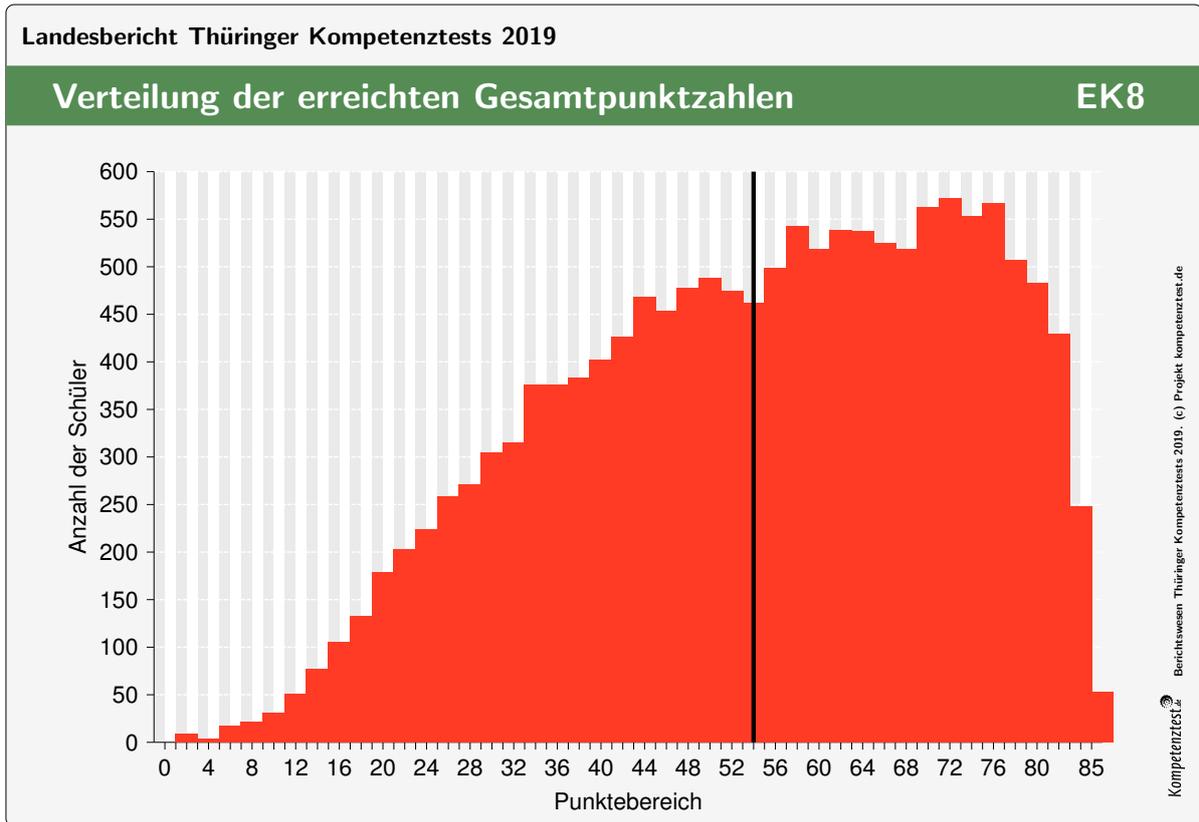


Abbildung 1.10: Verteilung der Gesamtpunktzahlen im Fach Englisch, Klassenstufe 8

1.3 Erreichte Gesamtpunktzahlen getrennt nach Schulart

In diesem Abschnitt wird die Häufigkeitsverteilung der erreichten Gesamtpunktzahlen hinsichtlich der verschiedenen Schularten differenziert. Die Darstellung erfolgt durch so genannte Perzentilbänder, wie sie auch in den großen internationalen Studien (wie z. B. PISA) üblich sind. Die erreichten Punktzahlen der Schüler werden dabei durch einen waagerechten Balken gekennzeichnet. Dessen Abschnitte, die so genannten Perzentile, geben jeweils den Punktwert wieder, der von einem bestimmten Anteil leistungsähnlicher Schüler maximal erreicht wurde.

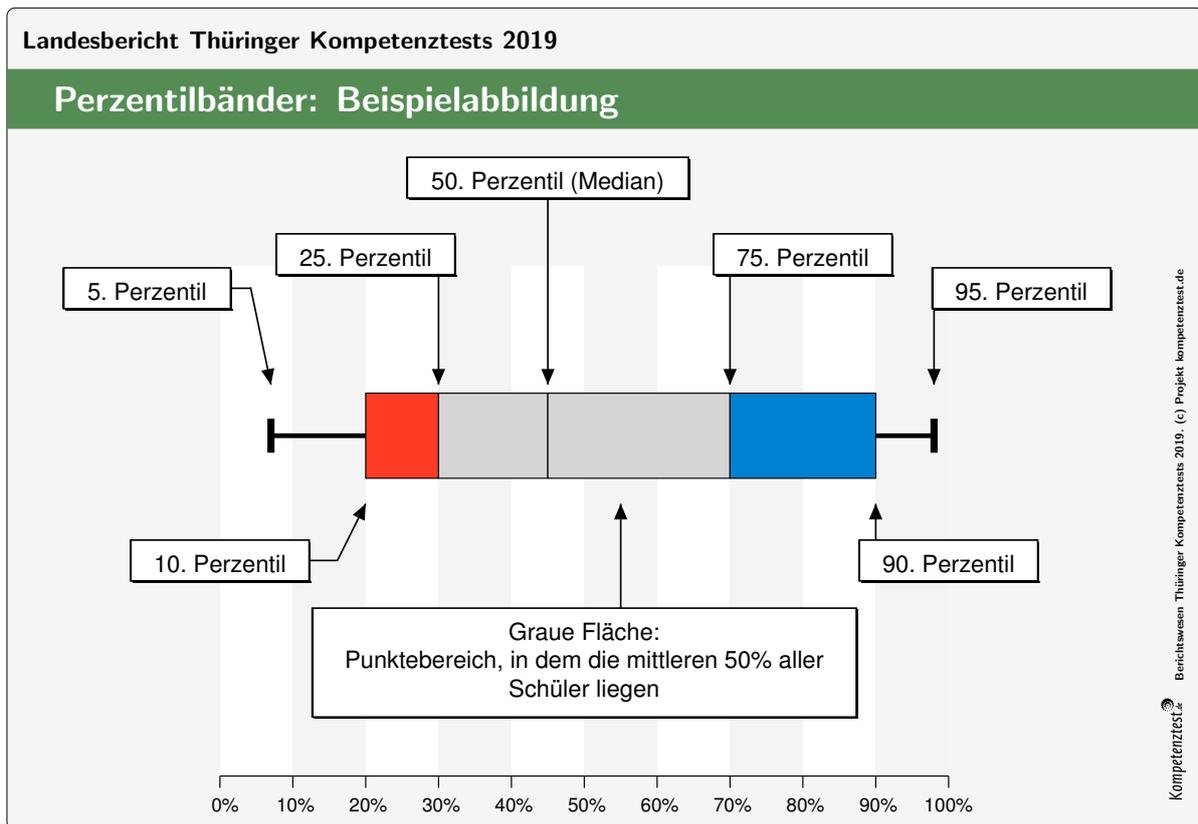


Abbildung 1.11: Beispiel für die Darstellung der erreichten Punktzahl in Prozent in Form eines Perzentilbandes

Abbildung 1.11 zeigt dies an einem Beispiel. Wiedergegeben ist das 5., 10., 25., 50., 75., 90. und 95. Perzentil. Der Grafik ist beispielsweise zu entnehmen, dass das leistungsschwächste Viertel der Schüler (25. Perzentil) maximal 30 % der Punkte erreicht hat.

Insgesamt zeigen sich deutliche Unterschiede zwischen den verschiedenen Schularten im Hinblick auf die erreichten Schülerleistungen. In der Klassenstufe 6 zum Beispiel liegen die 25. Perzentile der Gymnasiasten in allen Fächern in etwa auf Höhe des 75. Perzentils der Regelschüler. Das bedeutet, dass ein Gymnasiast, der bezüglich seiner Leistung stärker ist als 25 % aller Gymnasiasten, eine höhere Leistung als 75 % aller Regelschüler erbringt (siehe Abbildungen 1.14 bis 1.16).

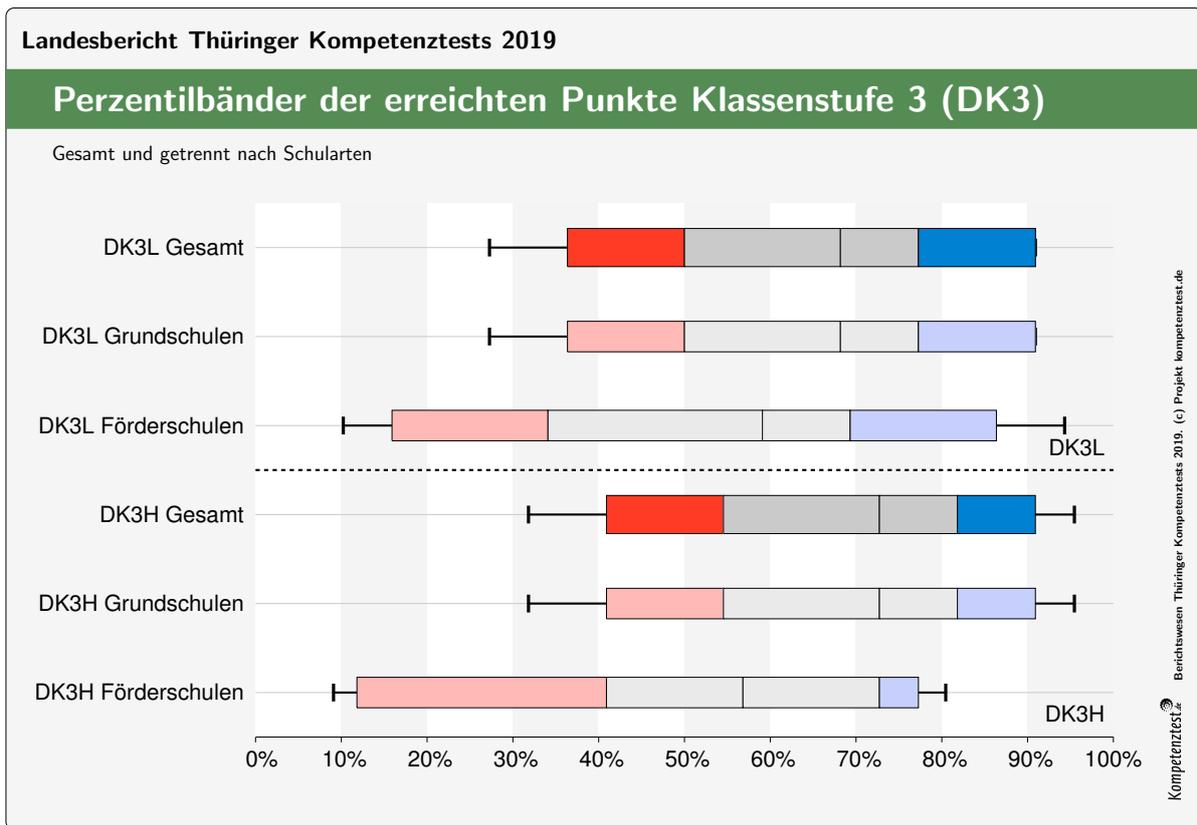


Abbildung 1.12: Ergebnisse bei den Deutschtests Lesen und Zuhören in der Klassenstufe 3, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

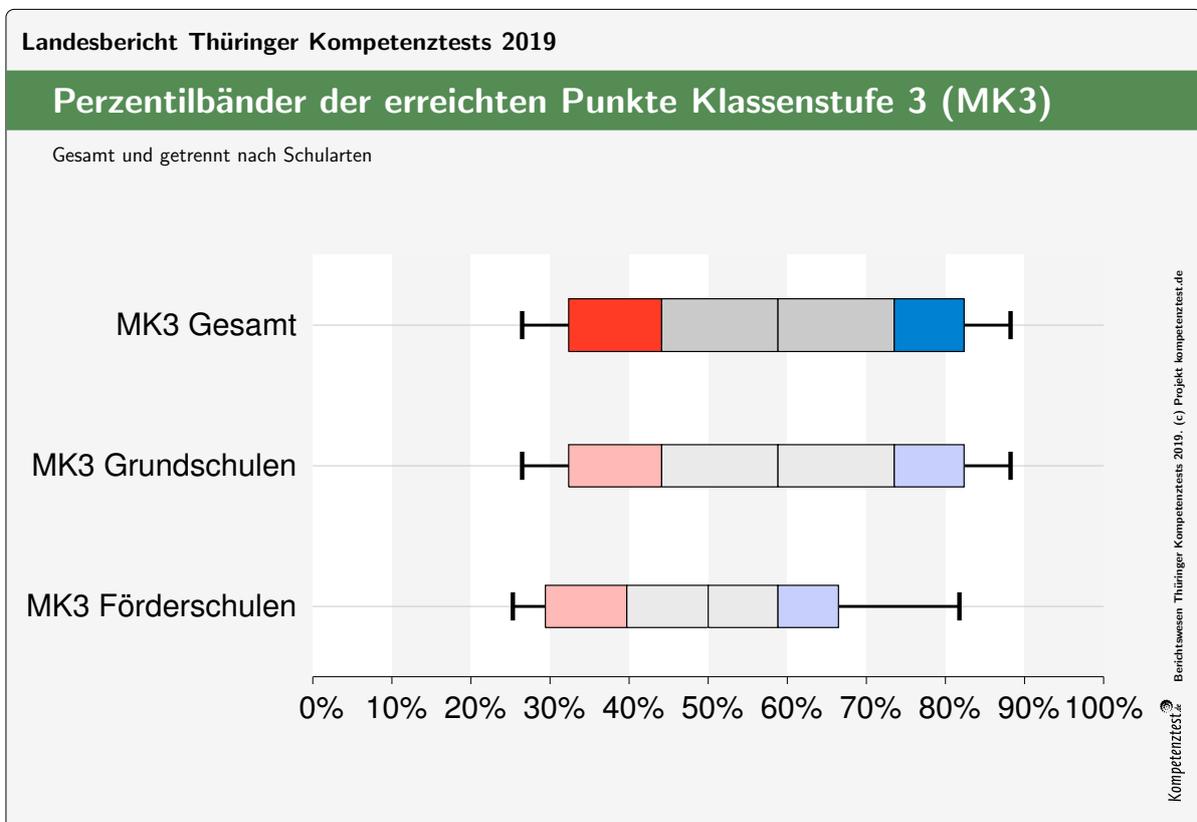


Abbildung 1.13: Ergebnisse beim Mathematiktest in der Klassenstufe 3, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

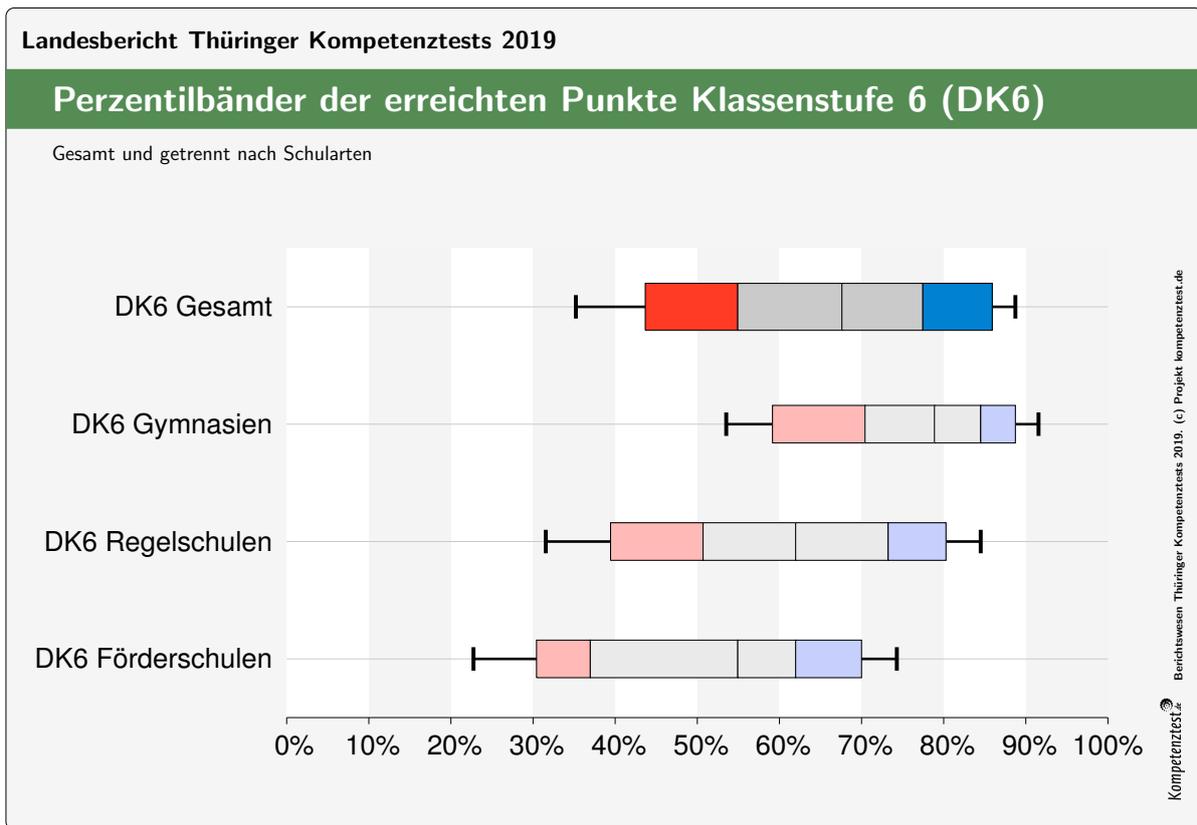


Abbildung 1.14: Ergebnisse beim Deutschtest in der Klassenstufe 6, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

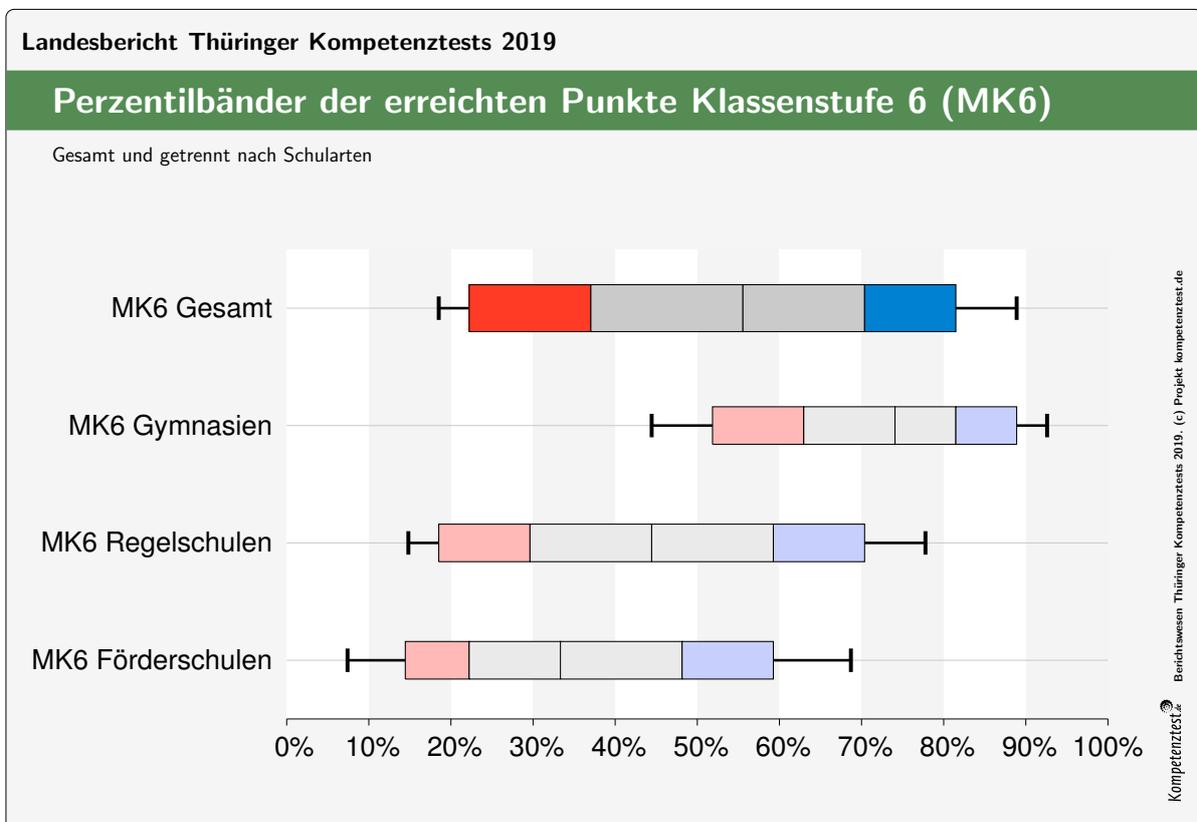


Abbildung 1.15: Ergebnisse beim Mathematiktest in der Klassenstufe 6, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

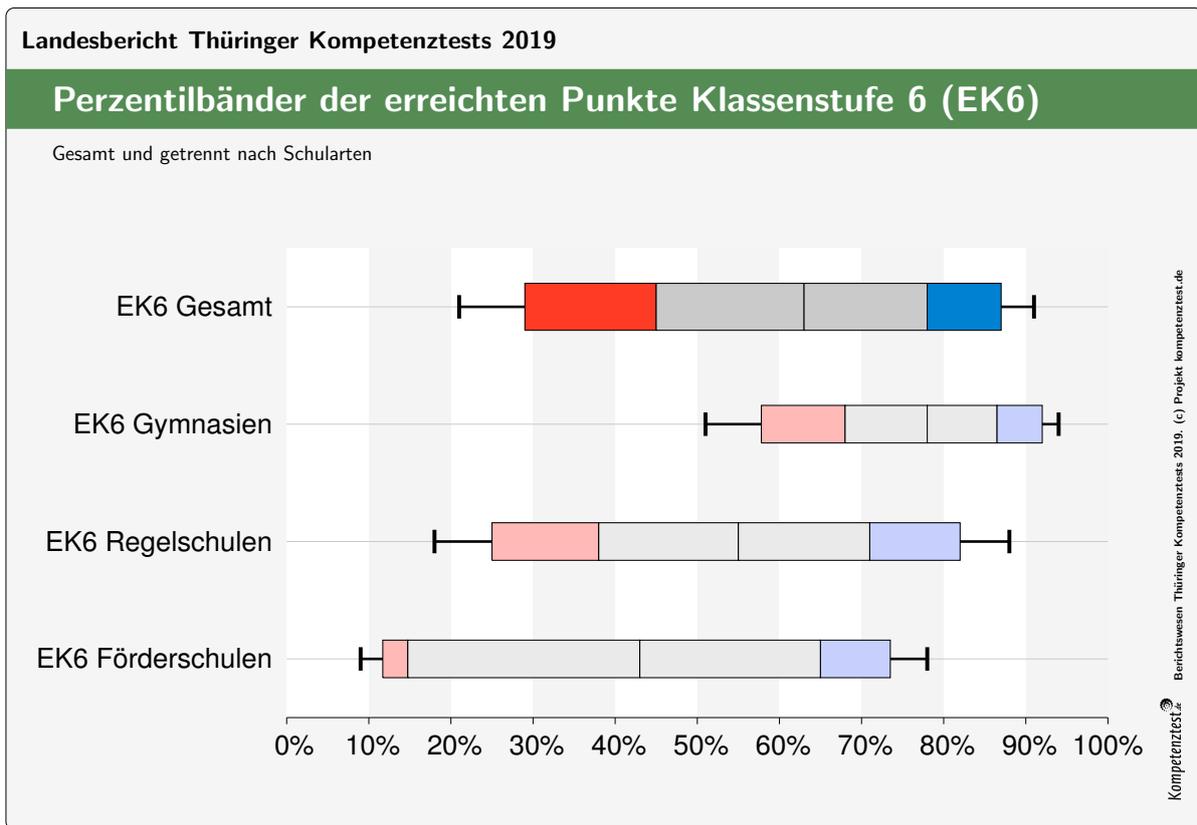


Abbildung 1.16: Ergebnisse beim Englischtest in der Klassenstufe 6, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

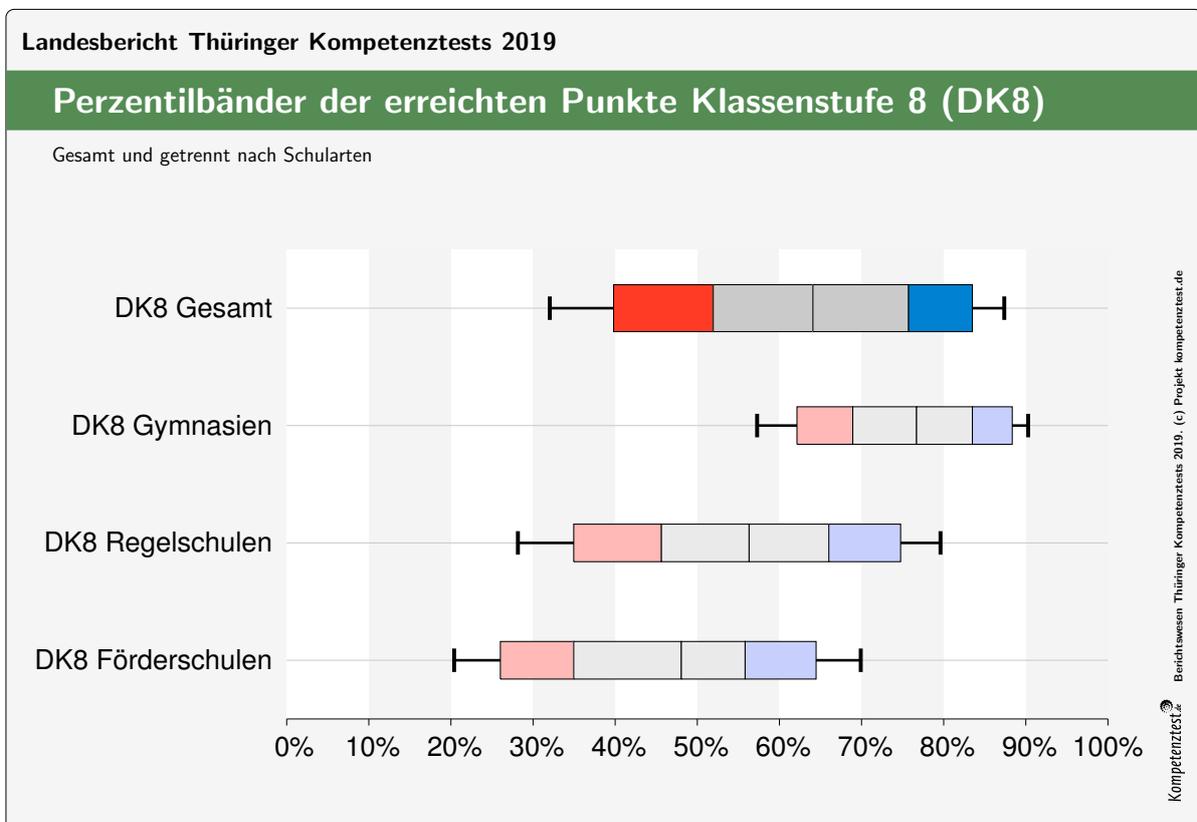


Abbildung 1.17: Ergebnisse beim Deutschtest in der Klassenstufe 8, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

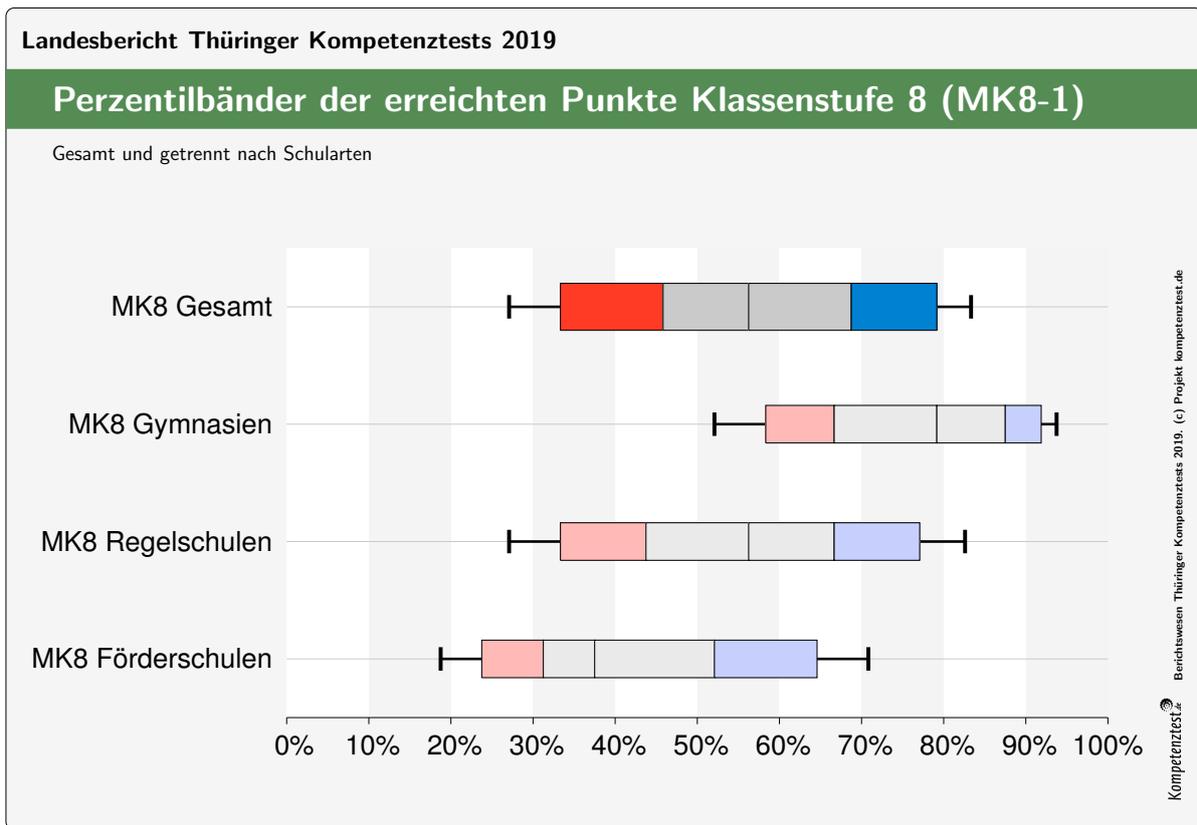


Abbildung 1.18: Ergebnisse beim Mathematiktest (Heft 1) in der Klassenstufe 8, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

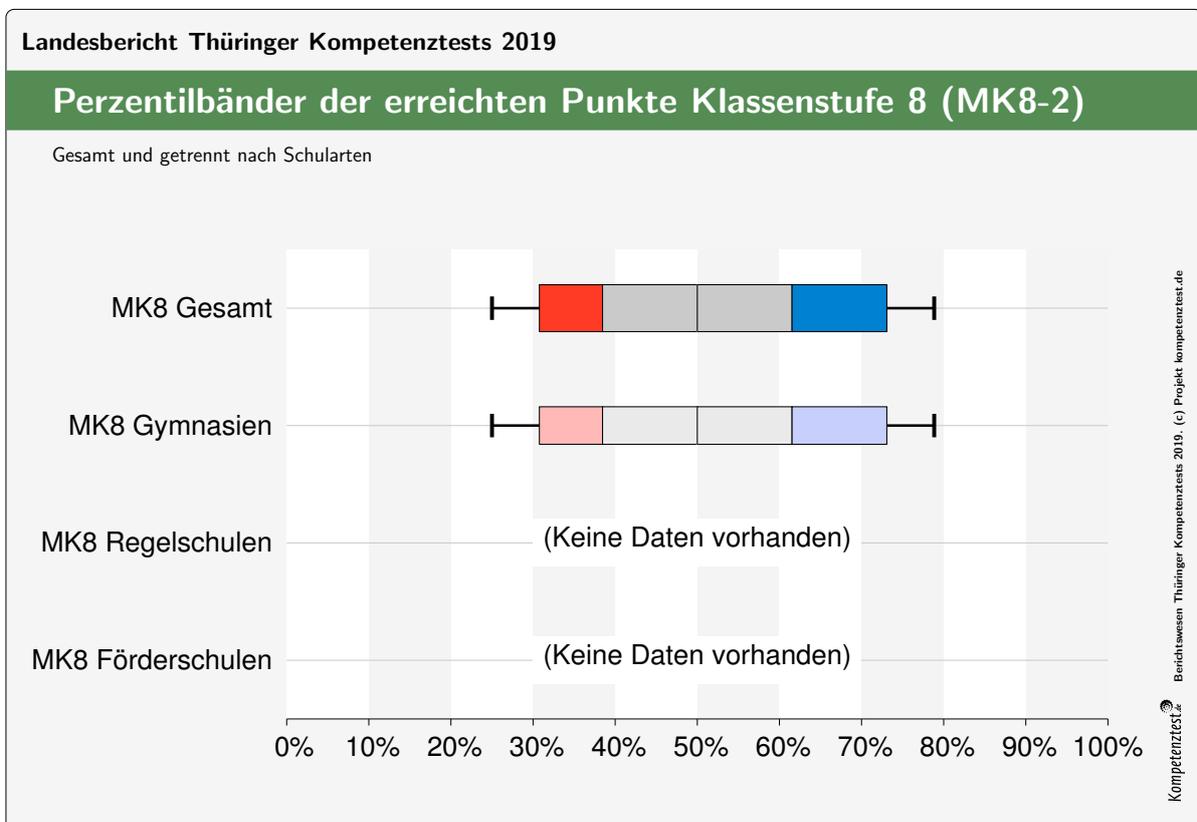


Abbildung 1.19: Ergebnisse beim Mathematiktest (Heft 2) in der Klassenstufe 8, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

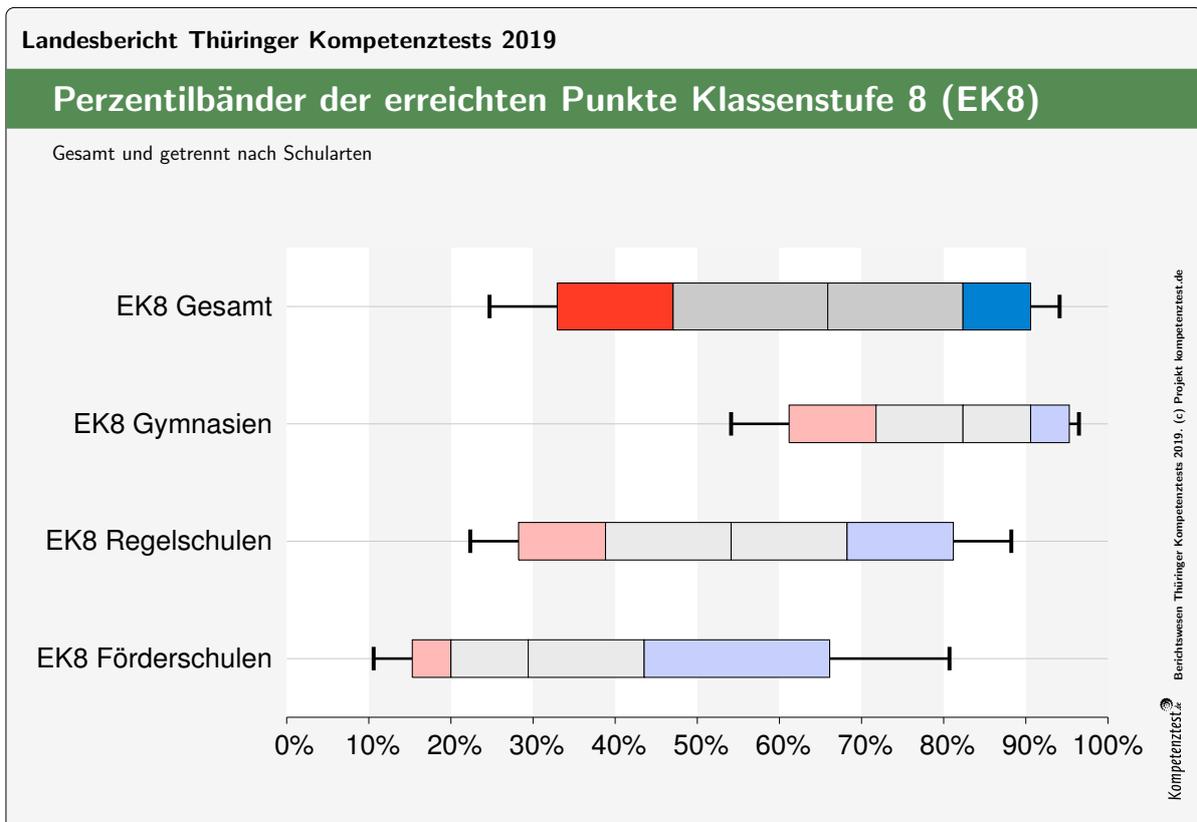


Abbildung 1.20: Ergebnisse beim Englischtest in der Klassenstufe 8, getrennt nach Schulart. Die Abbildung gibt die erreichte Punktzahl in Prozent wieder.

1.4 Aufgabenschwierigkeiten

Neben der Analyse der Gesamtergebnisse sind die Ergebnisse der einzelnen Aufgaben, welche in dem nun folgenden Abschnitt dargestellt werden, von besonderem Interesse. Hier zeigt sich sehr konkret, in welchen Bereichen die Stärken und Schwächen der Schüler liegen.

Wie wird die Schwierigkeit einer Aufgabe bestimmt? Die Schwierigkeit berechnet sich wie folgt: Die im Durchschnitt von den Schülern erreichte Punktzahl wird ins Verhältnis zur maximal erreichbaren Punktzahl gesetzt. Das Ergebnis wird in Prozent angegeben. Sind zum Beispiel bei einer Aufgabe maximal 5 Punkte möglich und die Schüler erreichen im Durchschnitt 3 Punkte, dann ist die Schwierigkeit dieser Aufgabe $\frac{3}{5} = 0,6$. Anders ausgedrückt, erreichen die Schüler im Mittel 60 % der maximalen Punktzahl. Bei einem Prozentwert nahe 0, ist die Aufgabe als sehr schwierig zu interpretieren. Liegt der Prozentwert jedoch nahe 100, so kann (fast) jeder Schüler dieser Schulart die Aufgabe vollständig lösen; die Aufgabe ist folglich für die Schüler als sehr leicht einzustufen.

Die ab Seite 19 folgenden Abbildungen geben die Aufgabenschwierigkeit bei den verschiedenen Teilaufgaben der Tests wieder. Dargestellt ist für jede Aufgabe der Landesdurchschnitt der erreichten Punkte in Prozent. Die Angaben sind außerdem nach der Schulart aufgeschlüsselt. Hier zeigen sich die zu erwartenden Unterschiede zwischen den Schularten². In den Abbildungen sind die Aufgaben zusätzlich bestimmten Kompetenzstufen bzw. Anforderungsbereichen zugeordnet. Die Erläuterungen dazu sind in Tabelle 1.4 wiedergegeben. Zu den Abbildungen im Fach Englisch gibt es keine Erläuterungen; hier sind die Aufgaben den Niveaustufen des Gemeinsamen Europäischen Referenzrahmens (GER) zugeordnet.

Im Kompetenztest Englisch der Klassenstufe 6 waren erneut Aufgaben zum Kompetenzbereich Schreiben enthalten. Die Ergebnisse in diesen Aufgaben können nicht ohne Weiteres als Schwierigkeiten interpretiert werden. Daher wurden gesonderte Abbildungen erstellt, welche auf die verschiedenen Levels des GER zurückgehen. In den Abbildungen 1.38 und 1.39 werden die Unterkriterien des Bereichs Schreiben dargestellt. Der linke Balken zeigt jeweils den Anteil aller teilnehmenden Schüler, welche die entsprechenden Kriterien mindestens auf dem betreffenden Level gelöst haben. Alle weiteren Balken geben diesen Wert schulartspezifisch wieder.

² Aufgrund der geringen Teilnehmerzahl können die Ergebnisse von Förderschulen in Klassenstufe 6 nicht anonymisiert dargestellt werden. Daher sind für diese Schülergruppe keine Balken in den Abbildungen für Klassenstufe 6 enthalten.

Tabelle 1.4: Erläuterung der Teilbereiche, denen die Aufgaben in den Abbildungen ab Seite 19 zugeordnet sind.

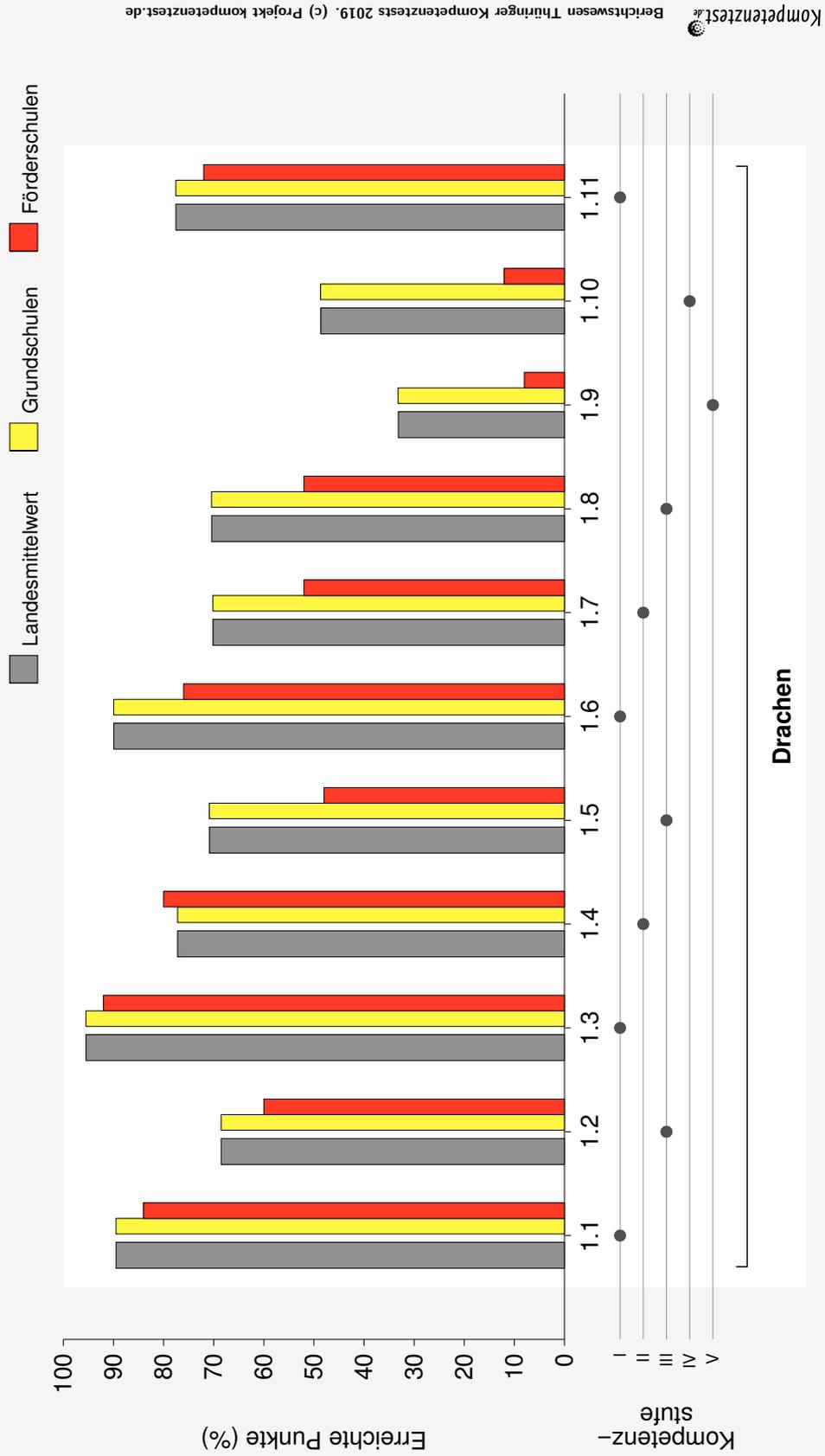
Test	zugeordnete Teilbereiche	Erläuterung
DK3L	Kompetenzstufe	<p>I Explizit angegebene Einzelinformationen identifizieren</p> <p>II Benachbarte Informationen miteinander verknüpfen</p> <p>III "Verstreute" Informationen miteinander verknüpfen und den Text ansatzweise als Ganzen erfassen</p> <p>IV Für die Herstellung von Kohärenz auf der Ebene des Textes wesentliche Aspekte erfassen</p> <p>V Auf zentrale Aspekte des Textes bezogene Aussagen selbstständig begründen</p>
DK3H	Kompetenzstufe	<p>I Kompetenzen, die auf dieser Stufe gezeigt werden, entsprechen nicht den Erwartungen der KMK-Bildungsstandards und liegen auch noch unter dem, was als Mindeststandard definiert werden kann.</p> <p>II Schüler auf dieser Stufe erfüllen noch nicht die in den Bildungsstandards der KMK für die vierte Jahrgangsstufe formulierten Erwartungen. Die Kompetenzstufe II kann jedoch als Erreichung des Mindeststandards für den Kompetenzbereich Sprache und Sprachgebrauch untersucht interpretiert werden.</p> <p>III Schülerinnen und Schüler, die diese Kompetenzstufe oder eine höhere erreicht haben, erfüllen die in den Bildungsstandards beschriebenen Erwartungen, d. h. sie erreichen den von der KMK festgelegten Regelstandard.</p> <p>IV Mit der Kompetenzstufe IV ist ein Leistungsniveau erreicht, das über den Erwartungen der Bildungsstandards liegt. Schülerinnen und Schüler, die sich auf dieser Stufe befinden, erreichen den Regelstandard plus.</p> <p>V Mit der Stufe V ist ein Leistungsbereich definiert, der nur bei sehr günstigen individuellen Voraussetzungen und optimalen schulischen und außerschulischen Lernangeboten erreicht werden kann. Die Kompetenzen auf dieser Stufe können als Erreichen des Optimalstandards interpretiert werden.</p>
MK3	Kompetenzstufe	<p>I Technische Grundlagen</p> <p>II Einfache Anwendungen von Grundlagenwissen</p> <p>III Erkennen und Nutzen von Zusammenhängen</p> <p>IV Sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang</p> <p>V Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien</p>
DK6	Anforderungsbereich	<p>I Informationen ermitteln / wiedergeben</p> <p>II Textbezogen interpretieren / Zusammenhänge herstellen</p> <p>III Reflektieren und Bewerten</p>
MK6	Anforderungsbereich	<p>I Reproduzieren</p> <p>II Zusammenhänge herstellen</p> <p>III Verallgemeinern und Reflektieren</p>
DK8	Anforderungsbereich	<p>I Informationen aus dem Stimulus wiedergeben</p> <p>II Verknüpfen von Informationen</p> <p>III Reflexion und Beurteilung von Informationen aus dem Stimulus</p>
MK8	Anforderungsbereich	<p>I Reproduzieren</p> <p>II Zusammenhänge herstellen</p> <p>III Verallgemeinern und Reflektieren</p>

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK3L (1/2)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
 Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

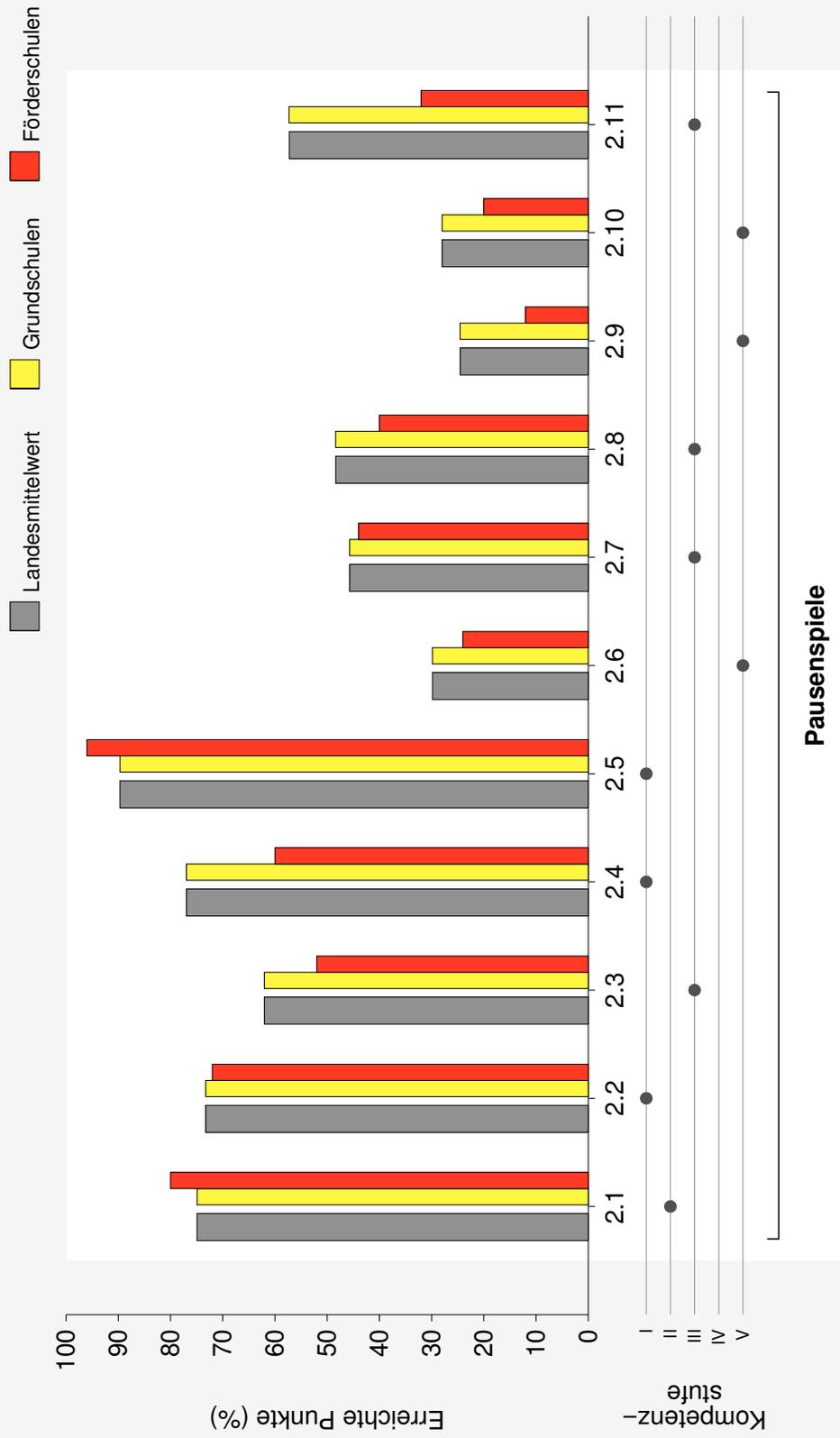
Abbildung 1.21: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest Lesen, Klassenstufe 3 – Teil 1)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK3L (2/2)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
 Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.22: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest Lesen, Klassenstufe 3 – Teil 2)

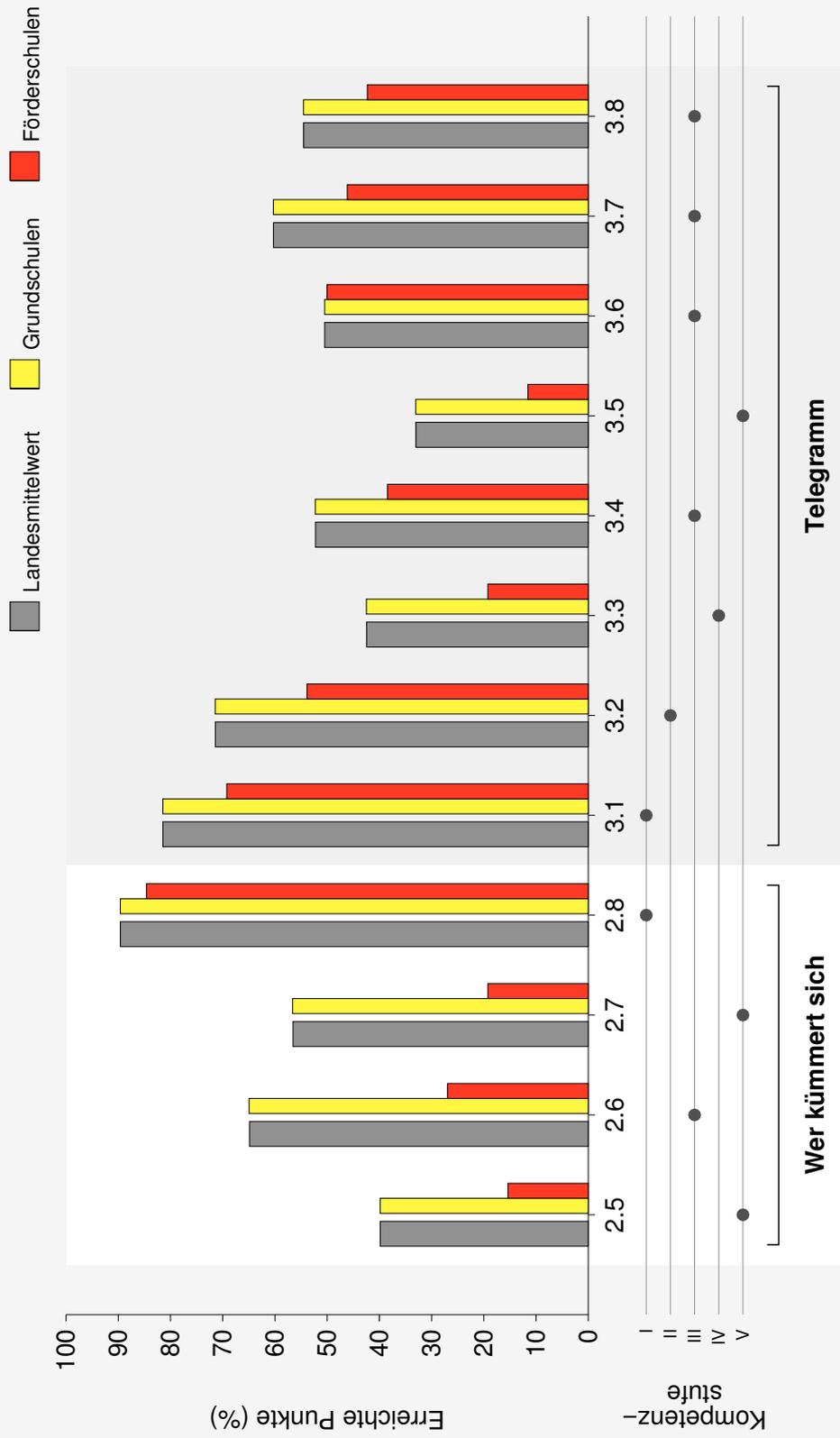


Abbildung 1.23: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest Sprachgebrauch, Klassenstufe 3 – Teil 1)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse DK3H (2/2)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
 Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.24: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest Sprachgebrauch, Klassenstufe 3 – Teil 2)

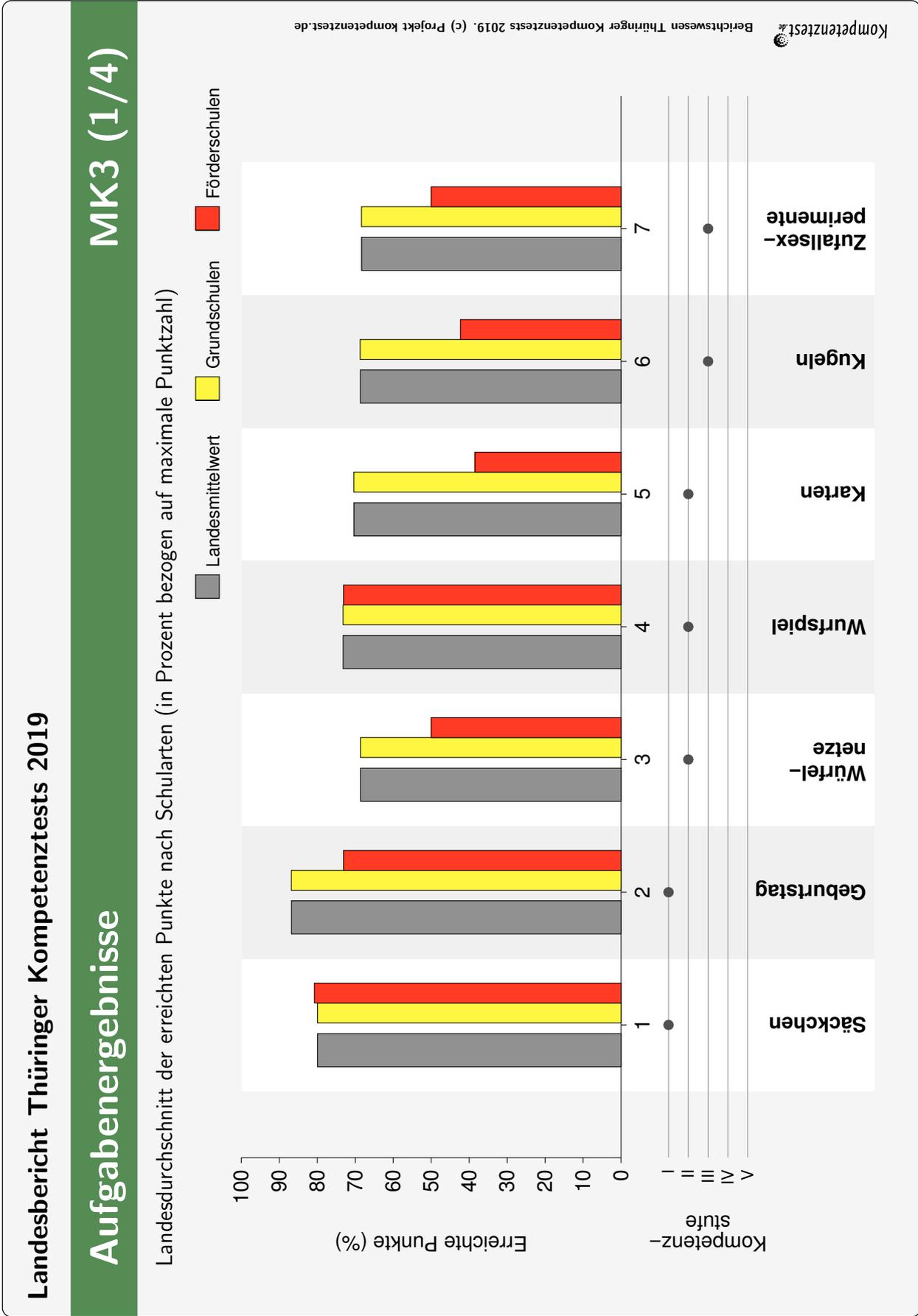


Abbildung 1.25: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Klassenstufe 3 – Teil 1)

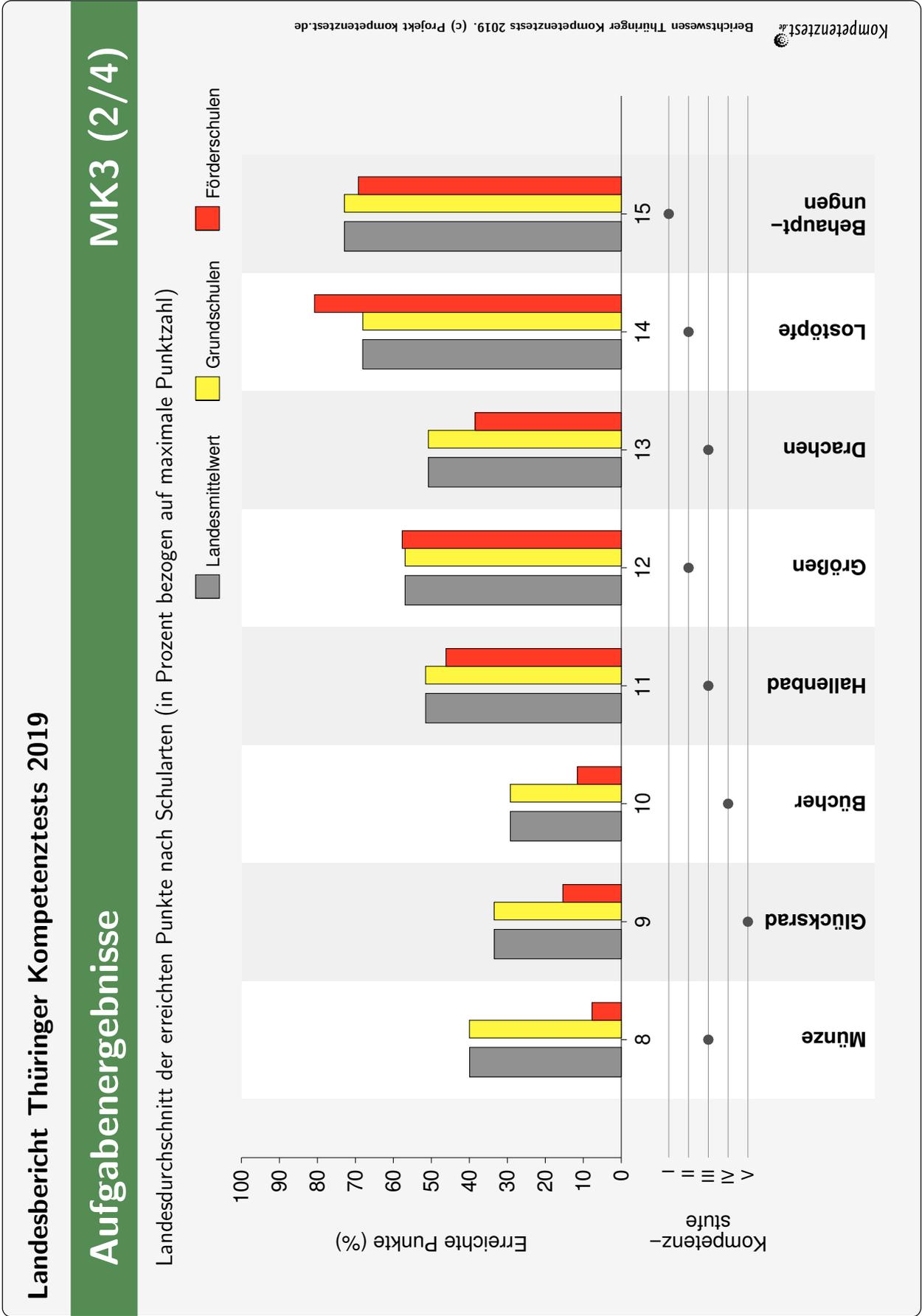


Abbildung 1.26: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Klassenstufe 3 – Teil 2)

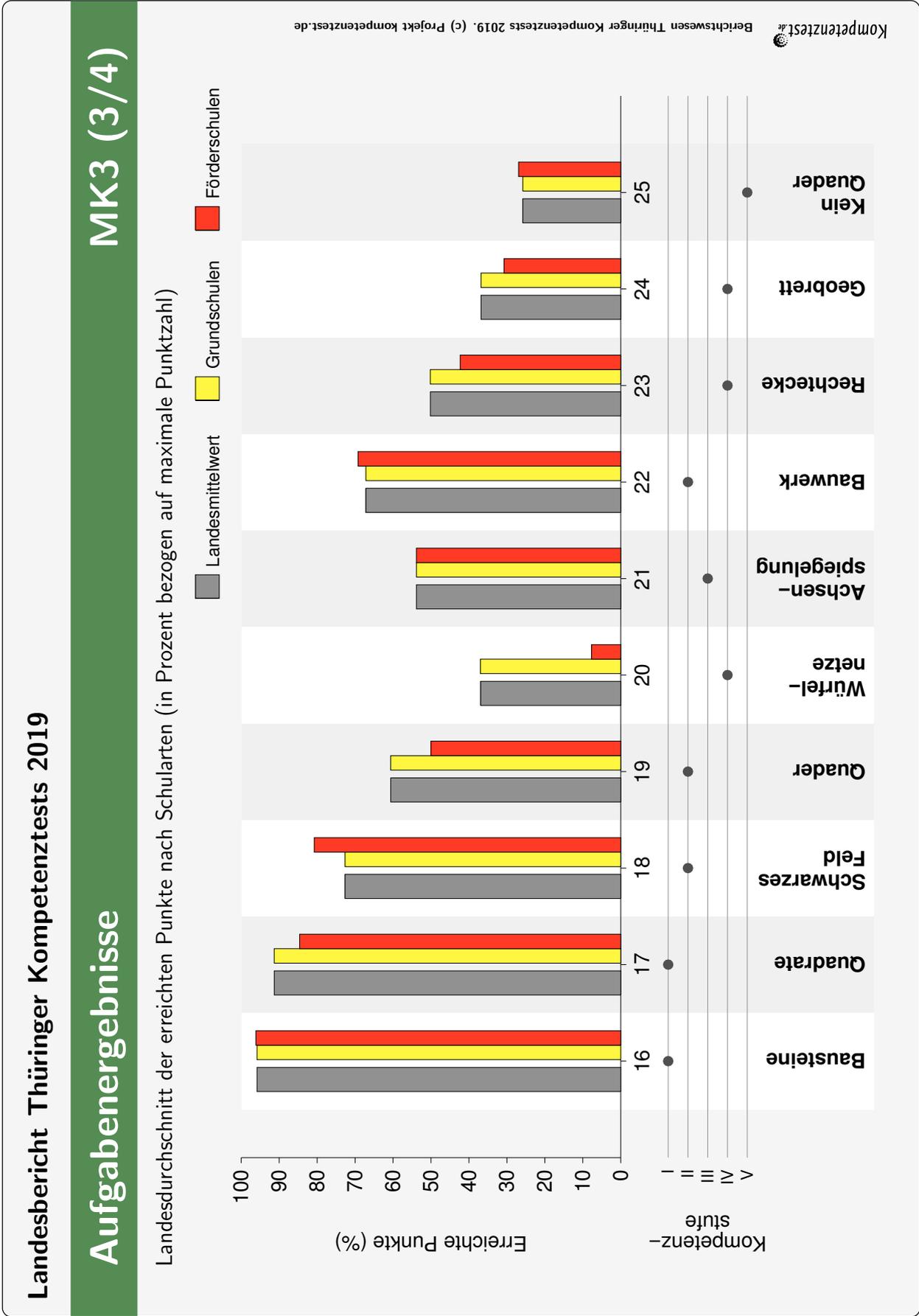


Abbildung 1.27: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Klassenstufe 3 – Teil 3)

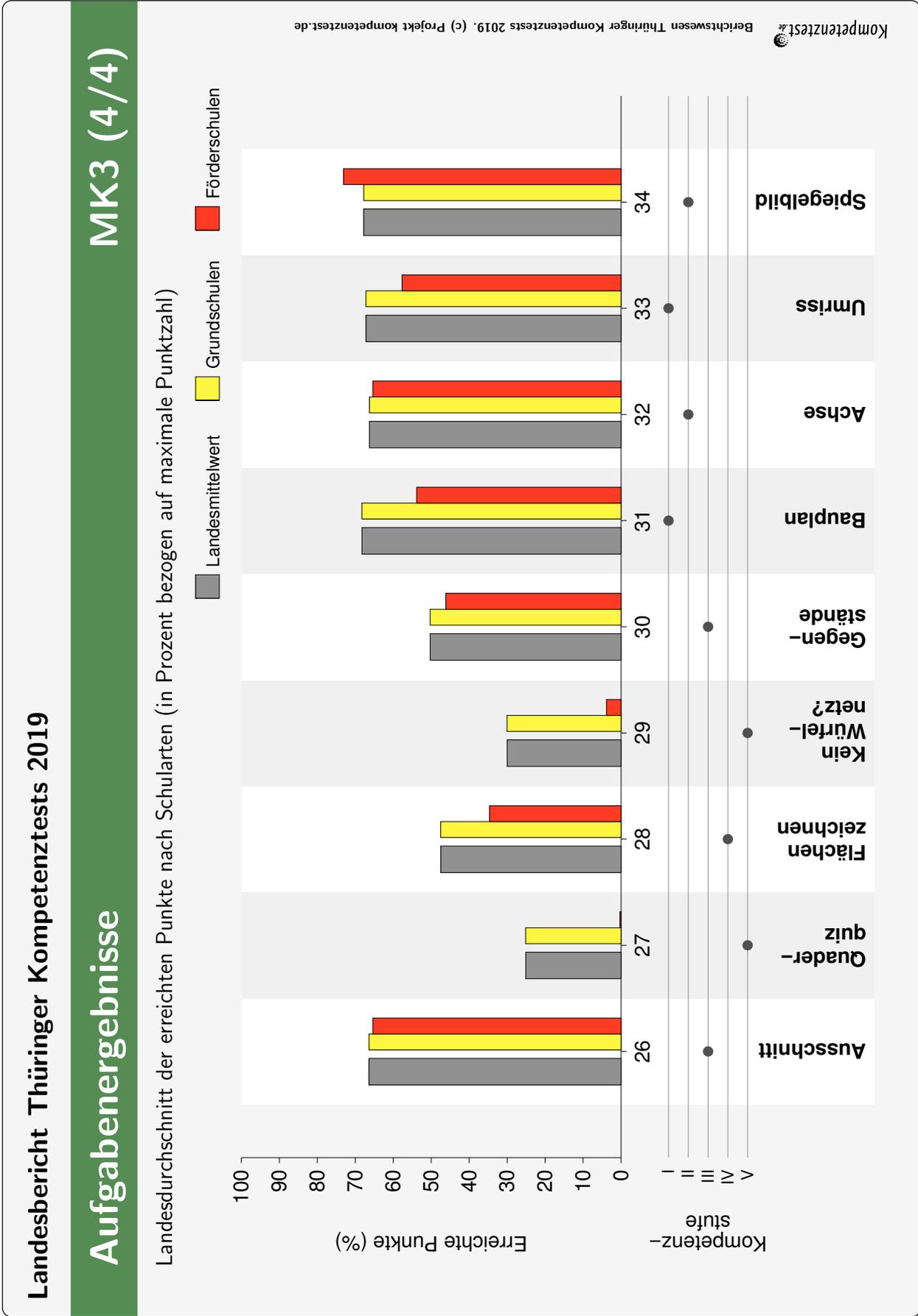


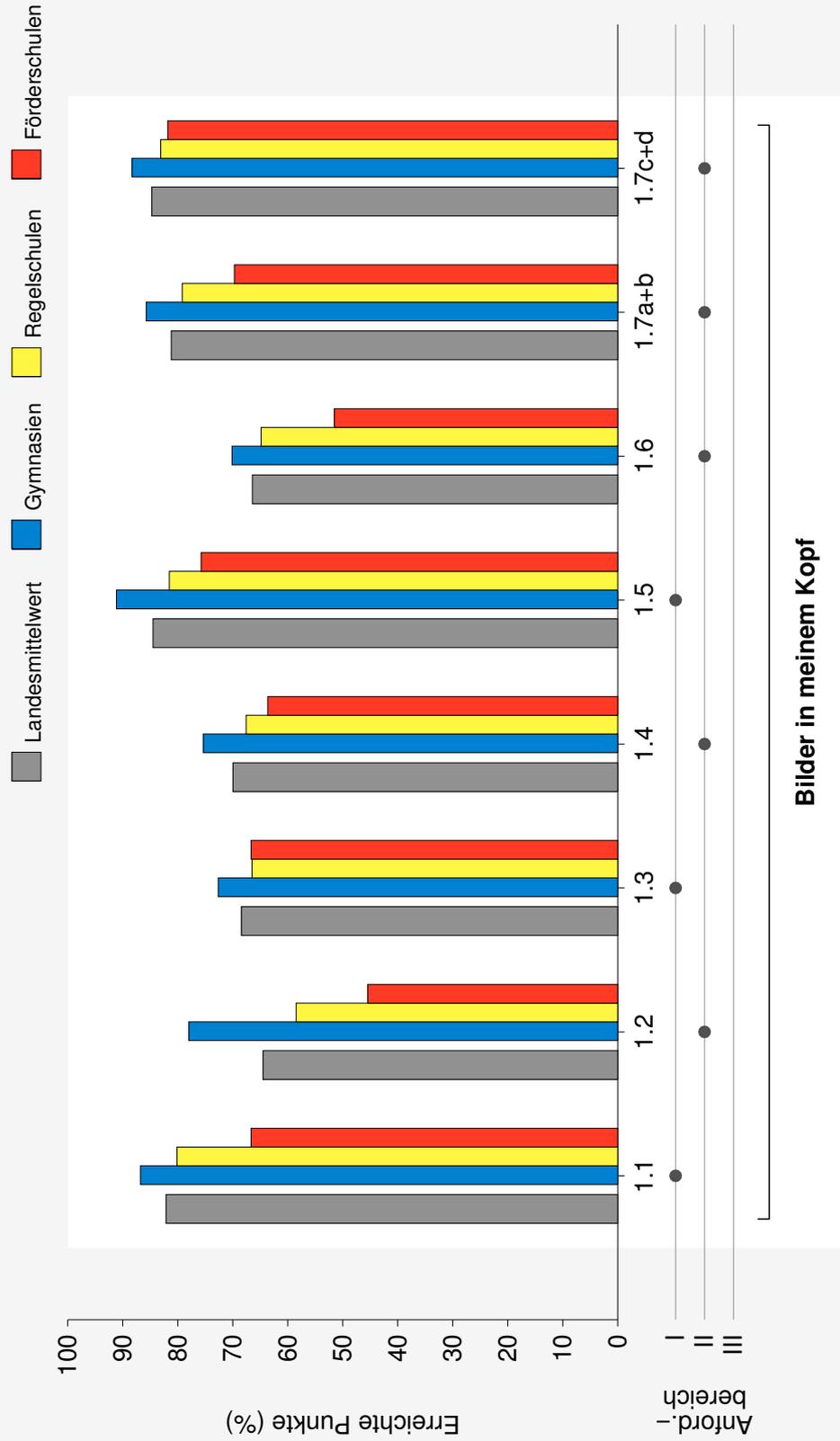
Abbildung 1.28: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Klassenstufe 3 – Teil 4)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK6 (1/6)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

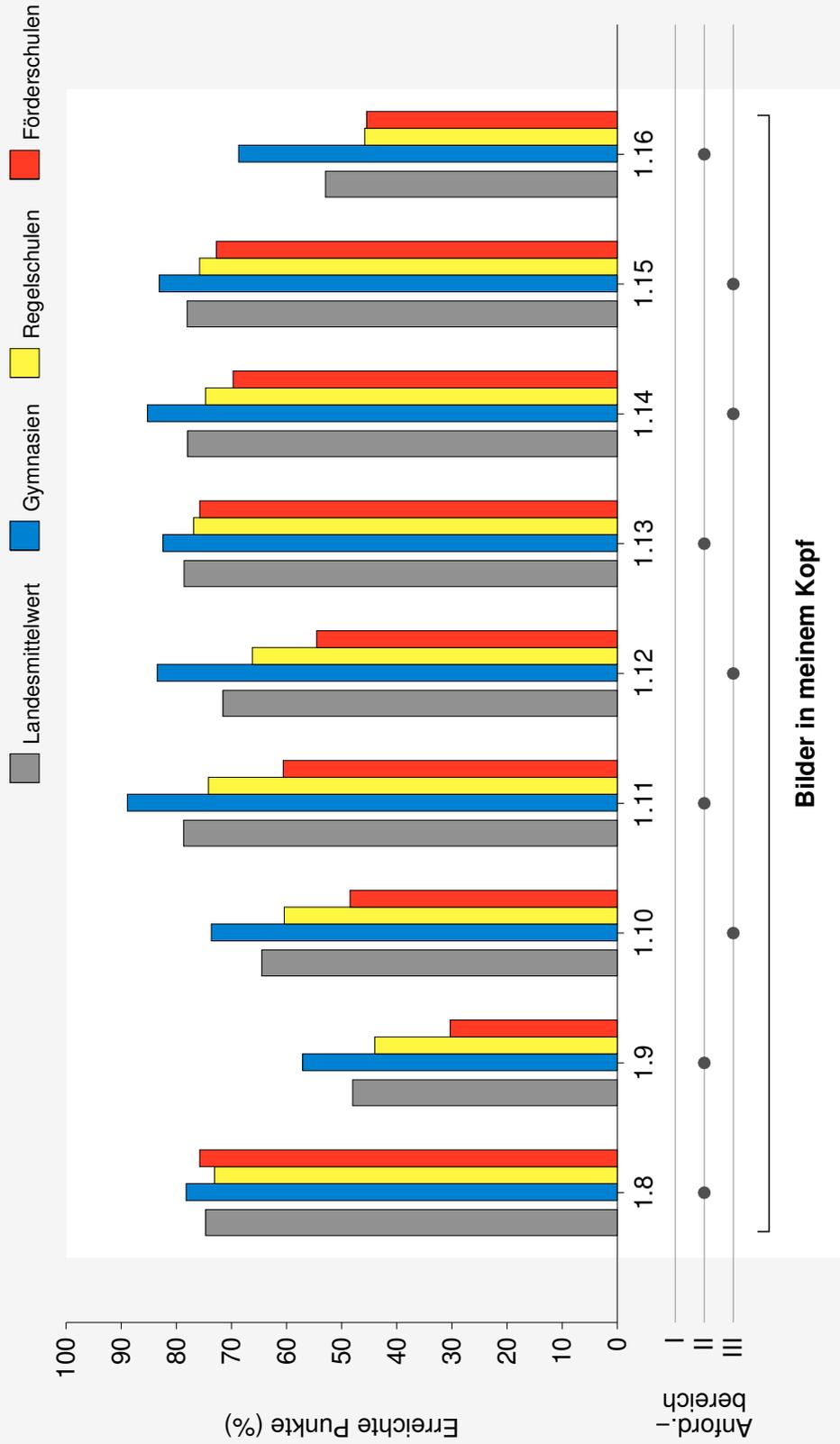
Abbildung 1.29: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 6 – Teil 1)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK6 (2/6)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.30: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 6 – Teil 2)

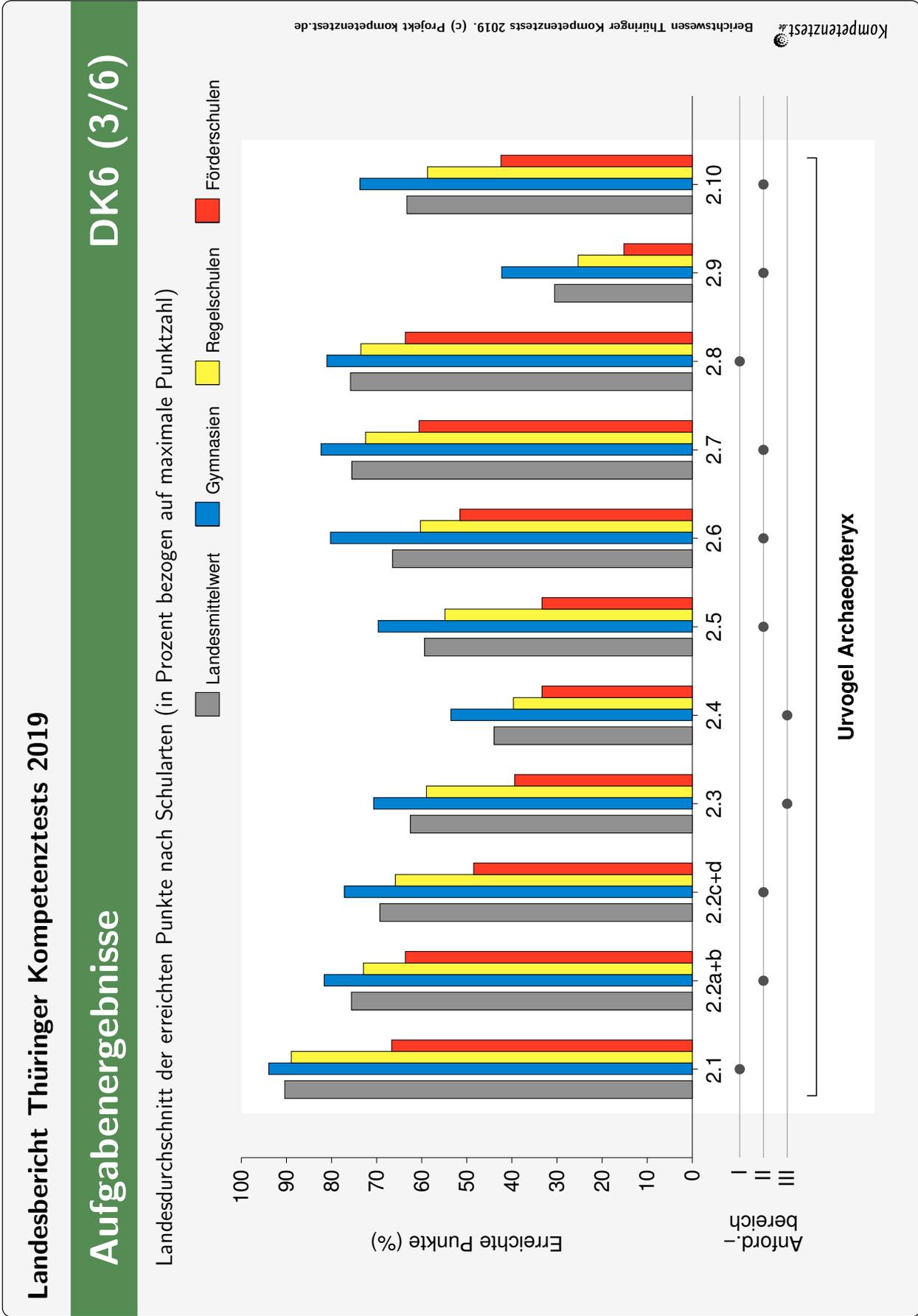


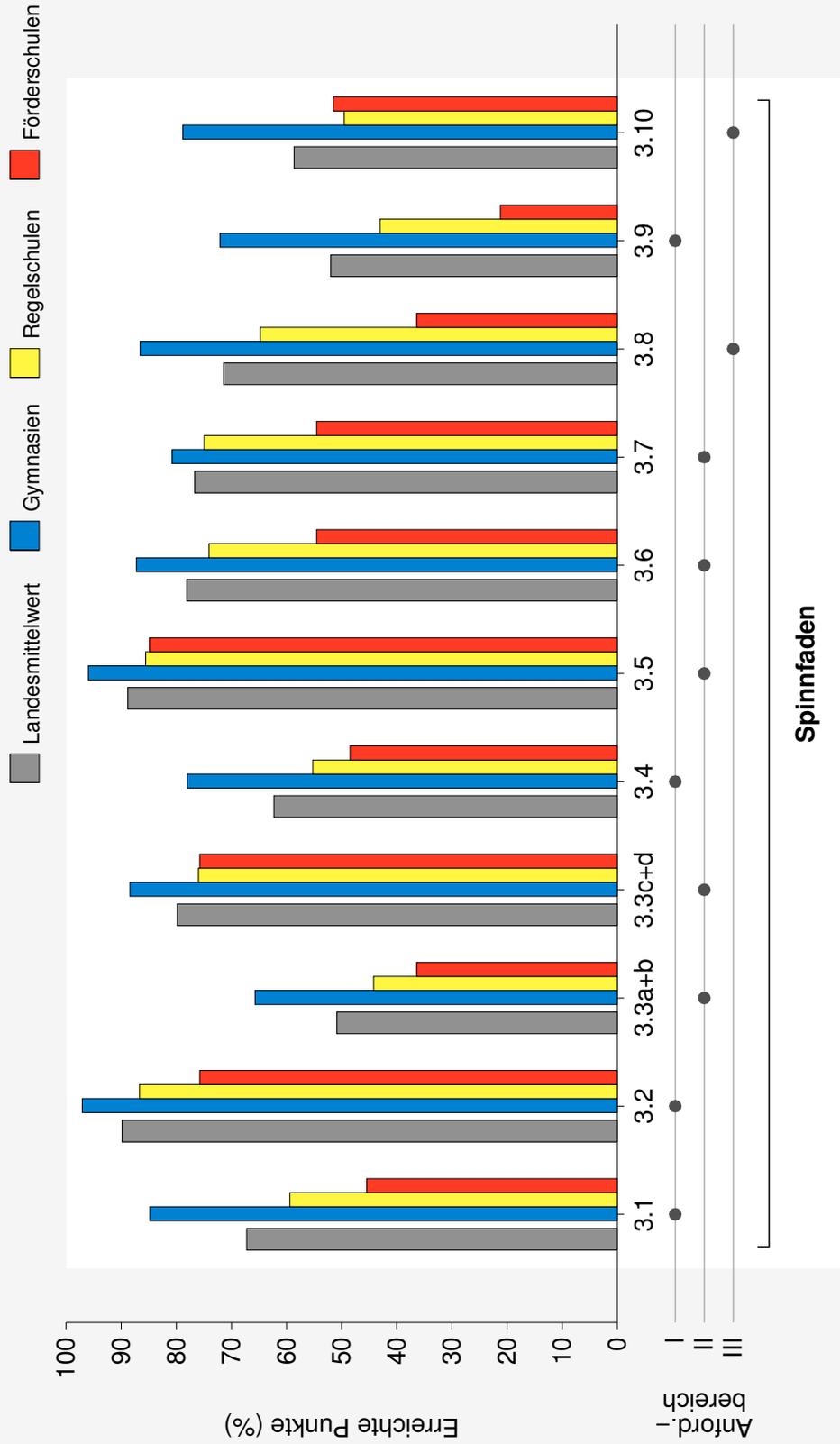
Abbildung 1.31: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 6 – Teil 3)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK6 (4/6)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.32: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 6 – Teil 4)

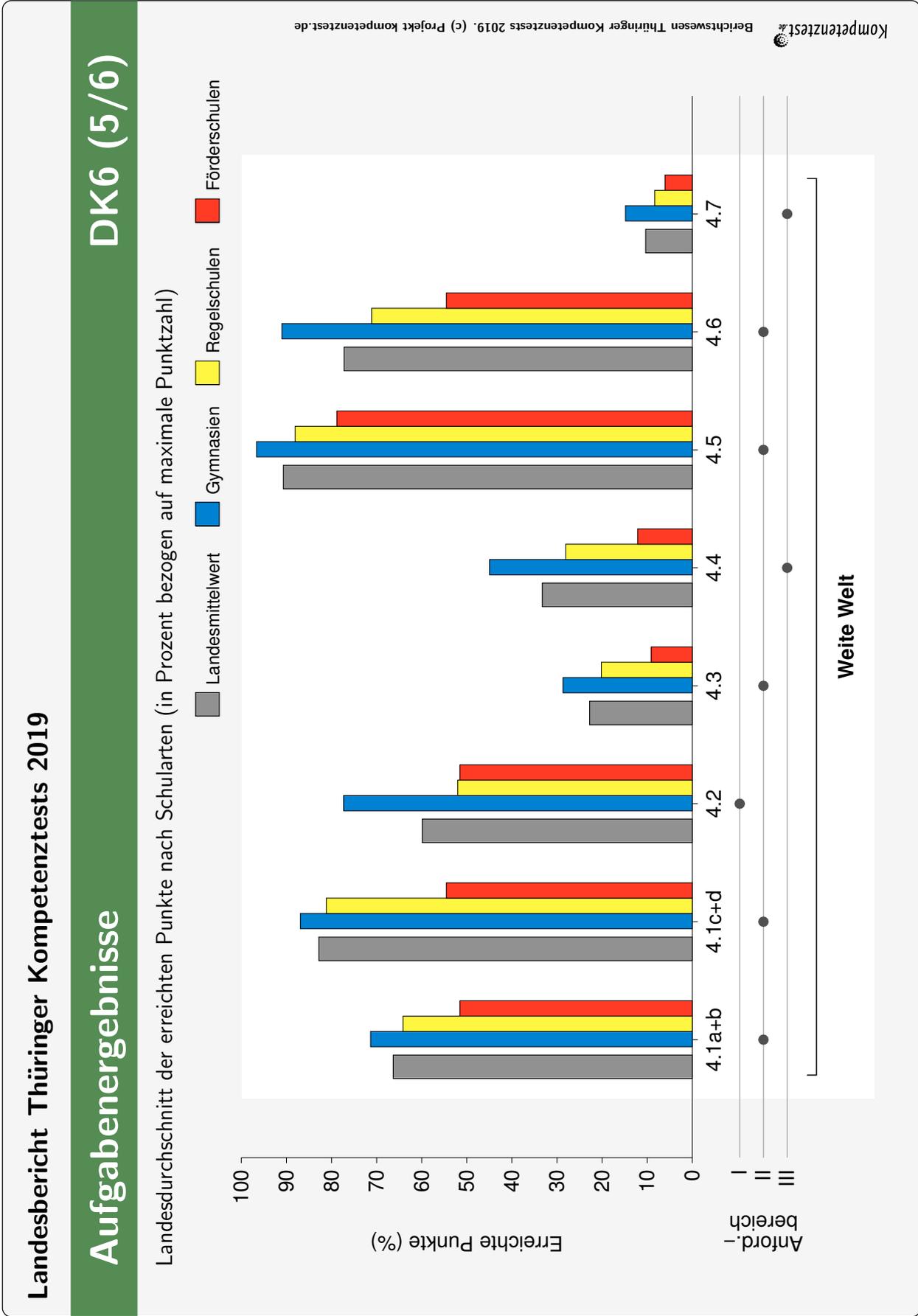


Abbildung 1.33: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 6 – Teil 5)

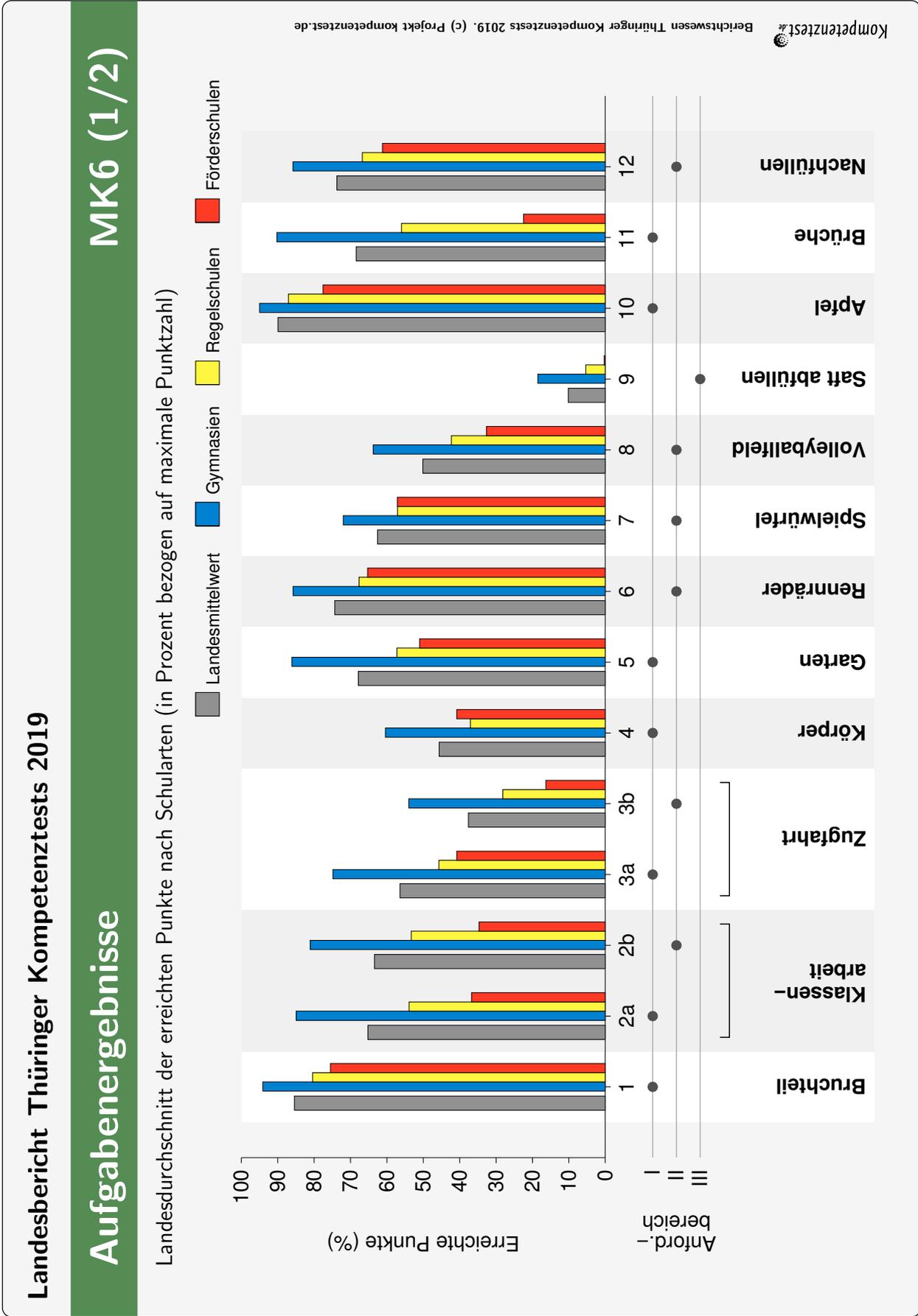


Abbildung 1.35: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Klassenstufe 6 – Teil 1)

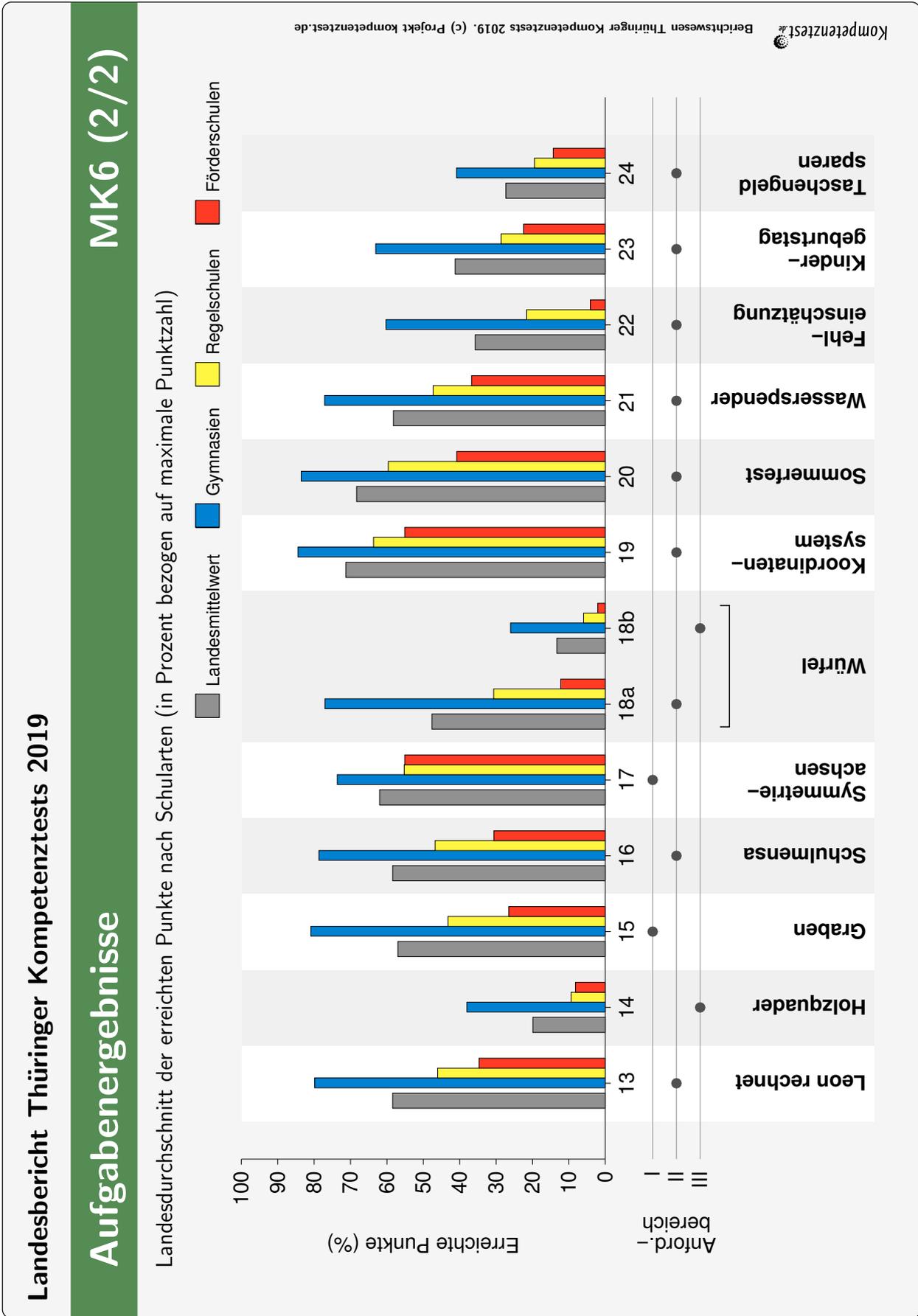


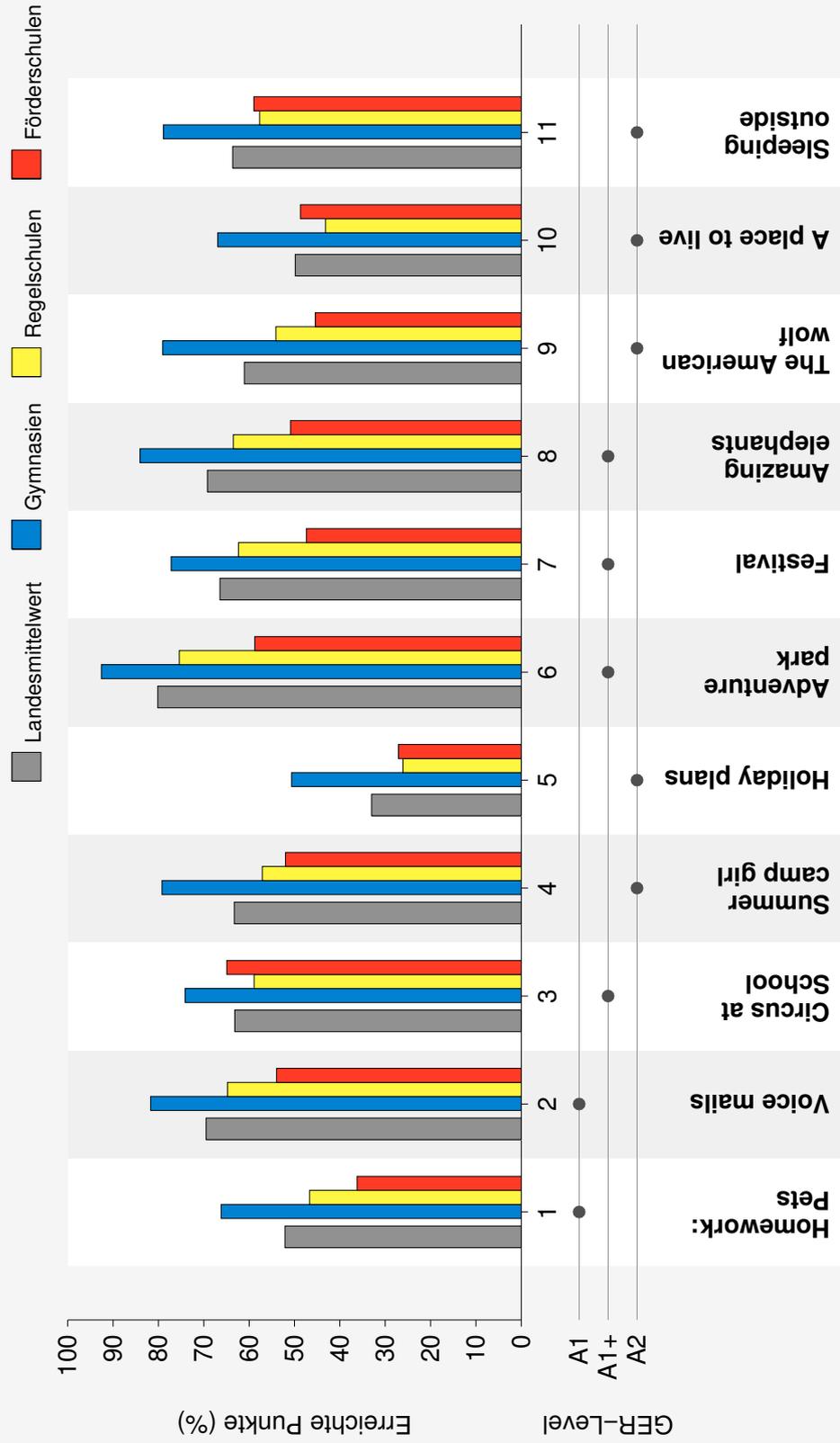
Abbildung 1.36: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Klassenstufe 6 – Teil 2)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

EK6

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.37: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englischtest Teilbereiche Hör- und Leseverstehen, Klassenstufe 6)

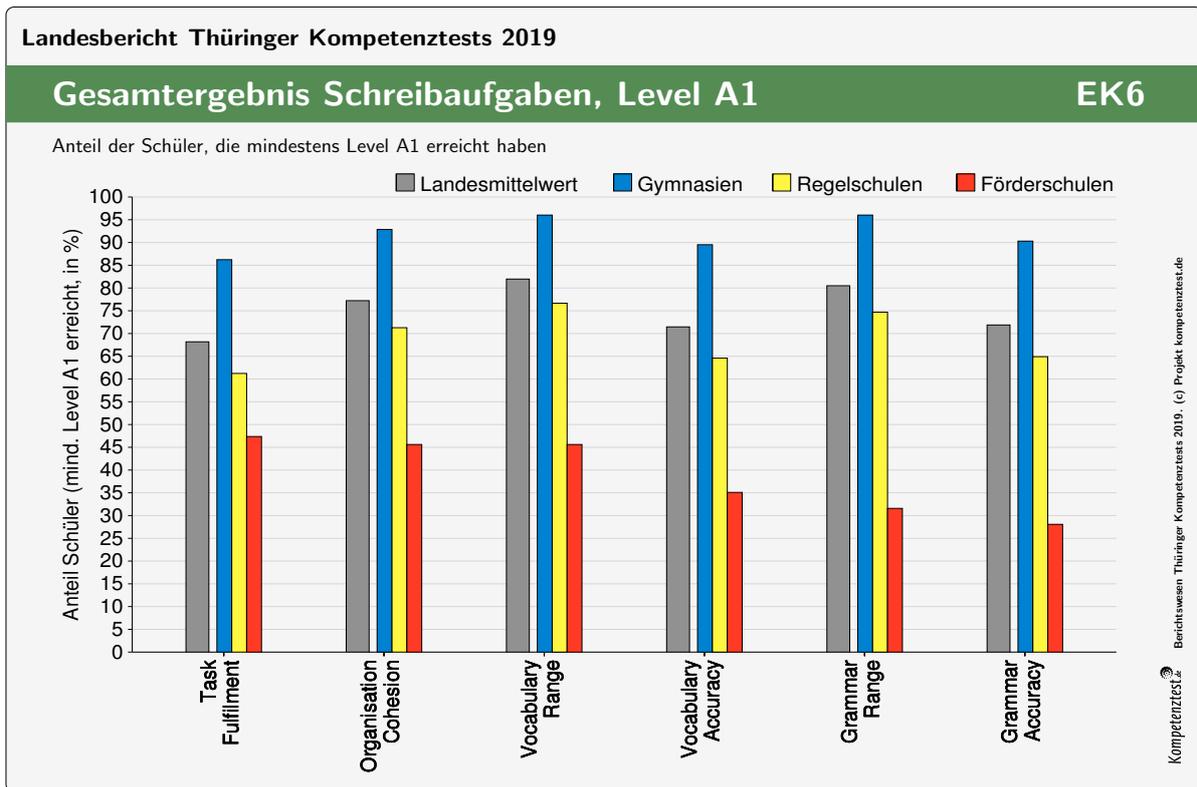


Abbildung 1.38: Ergebnisse der Schreibkriterien der Aufgaben 'Favourite food', 'My favourite season' und 'A nice dream' (Level A1)

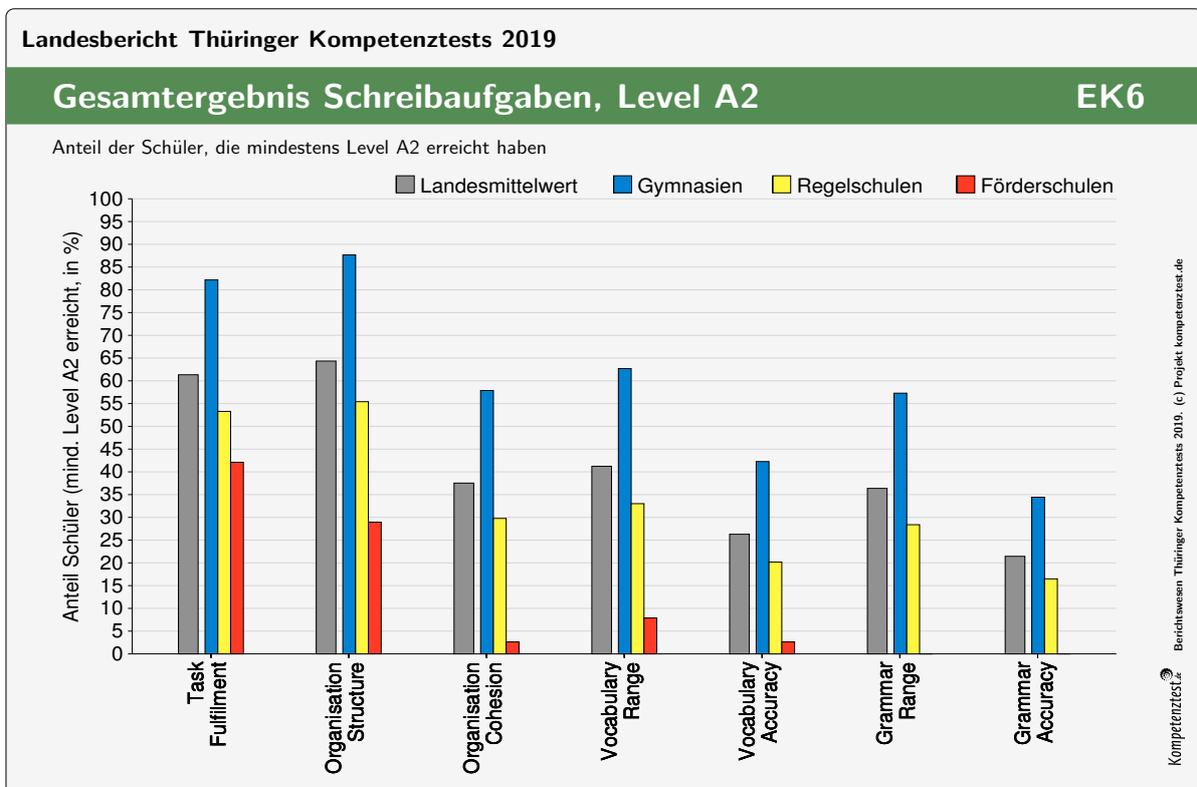
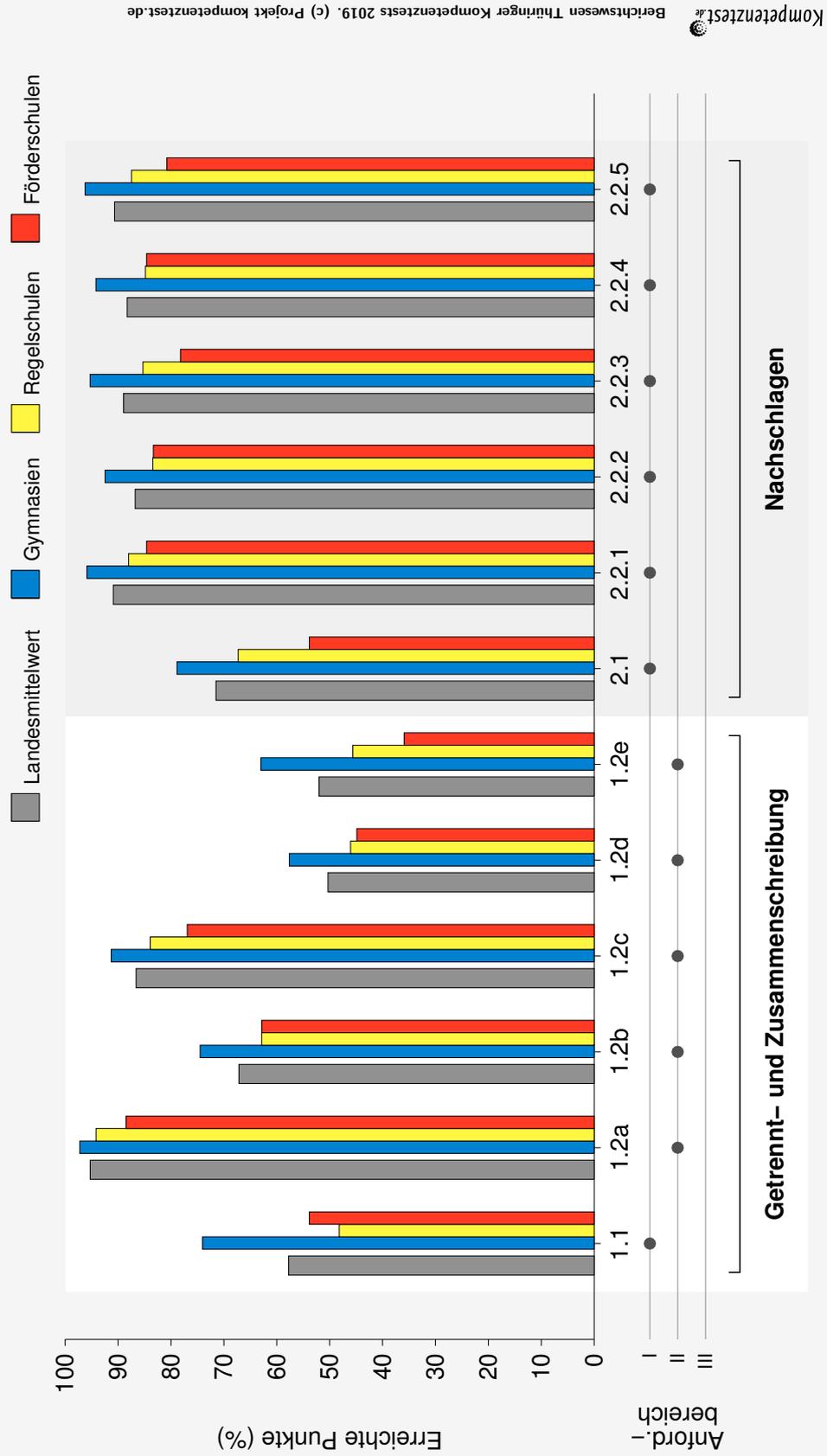


Abbildung 1.39: Ergebnisse der Schreibkriterien der Aufgaben 'My favourite season' und 'A nice dream' (Level A2)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse **DK8 (1/9)**

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
 Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

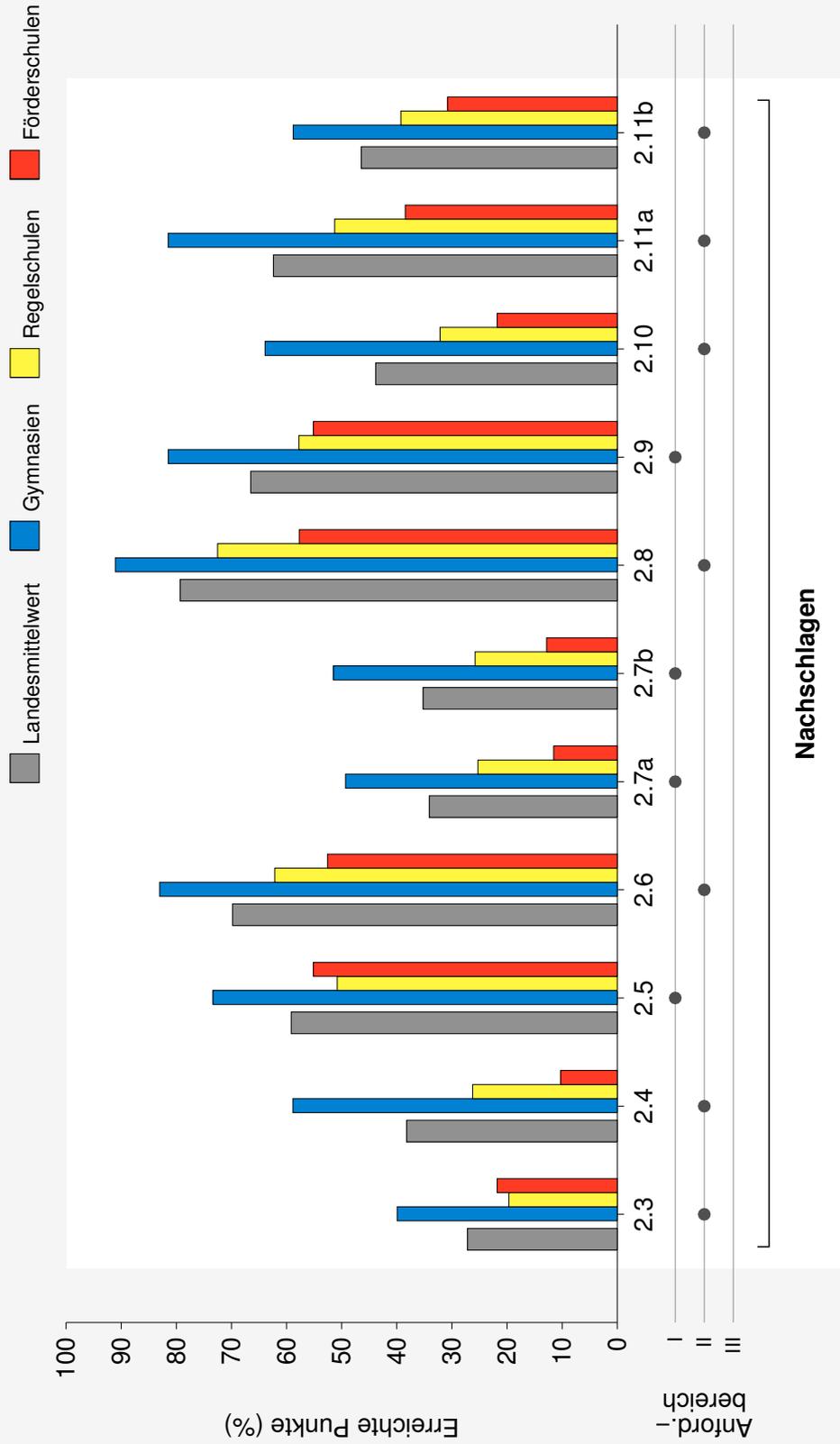
Abbildung 1.40: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 1)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK8 (2/9)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

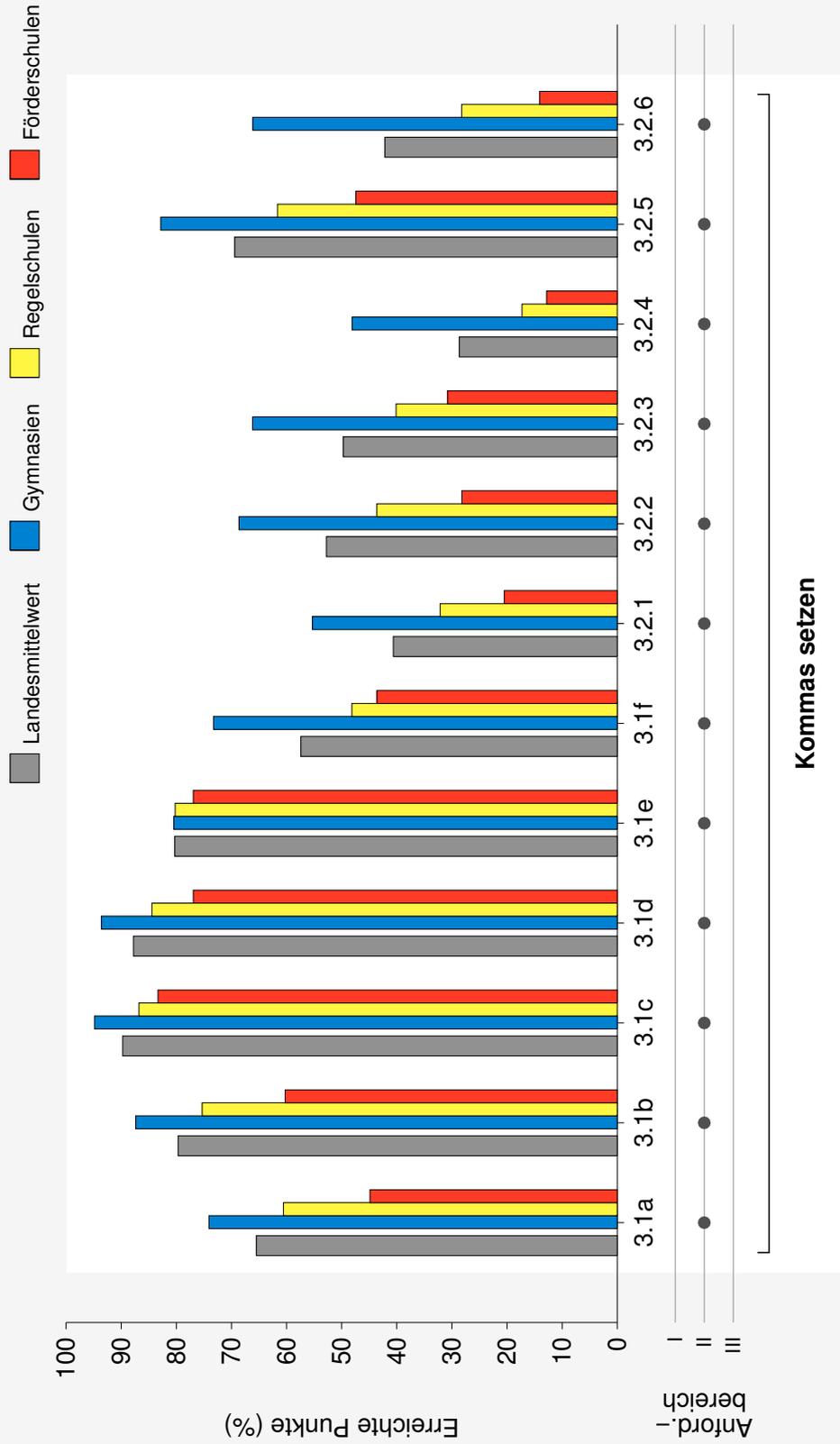
Abbildung 1.41: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 2)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK8 (3/9)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.42: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 3)

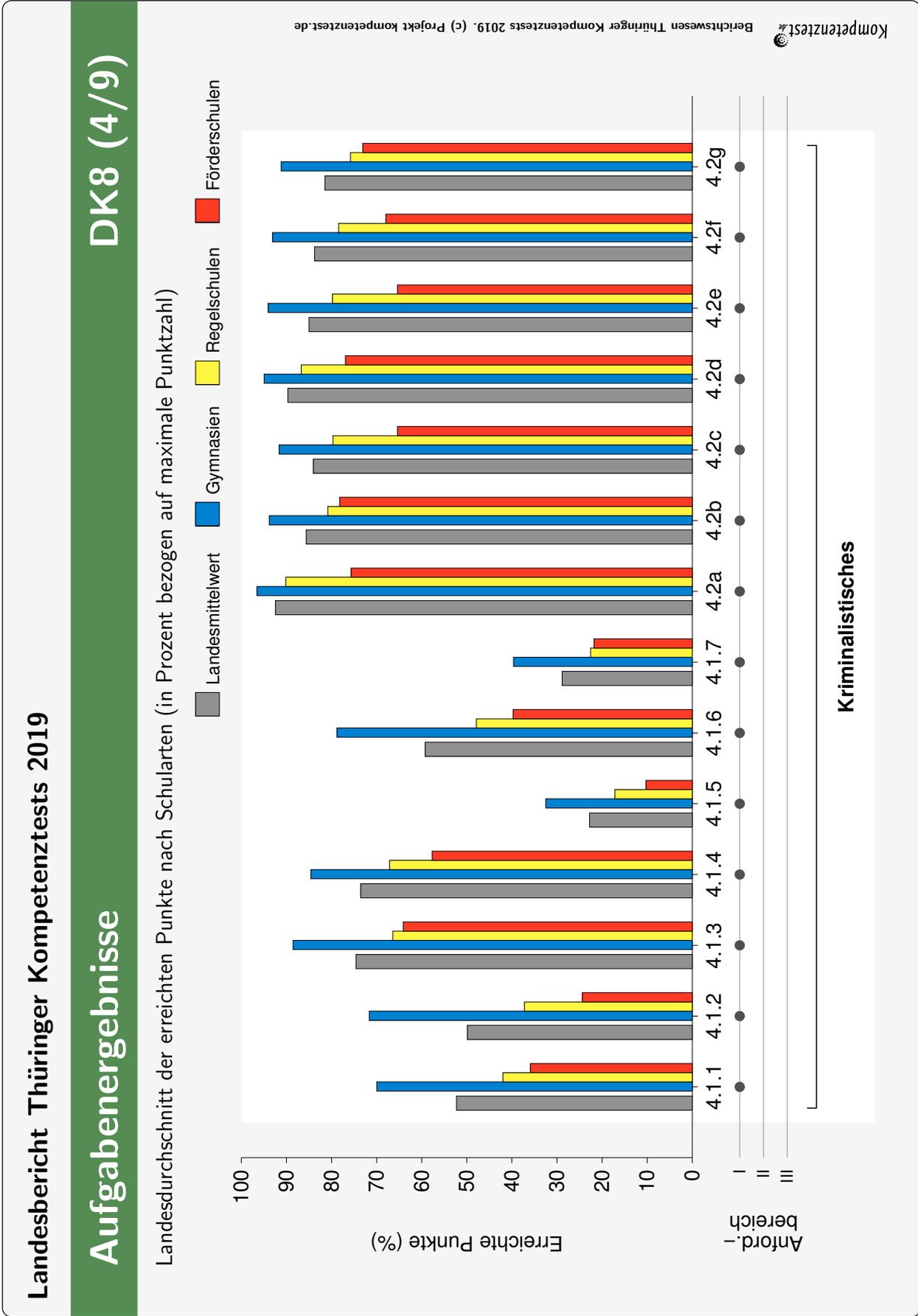


Abbildung 1.43: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 4)

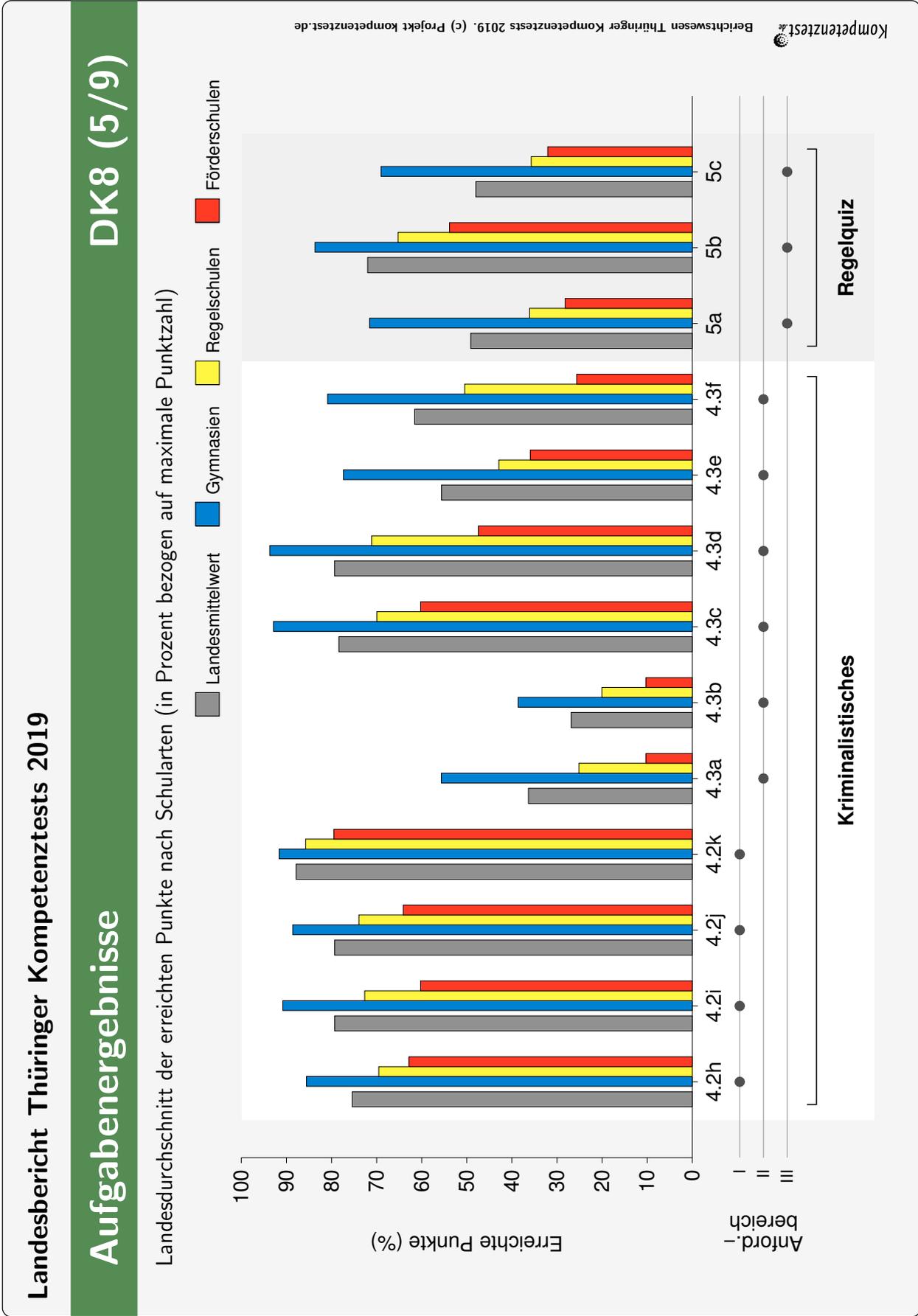
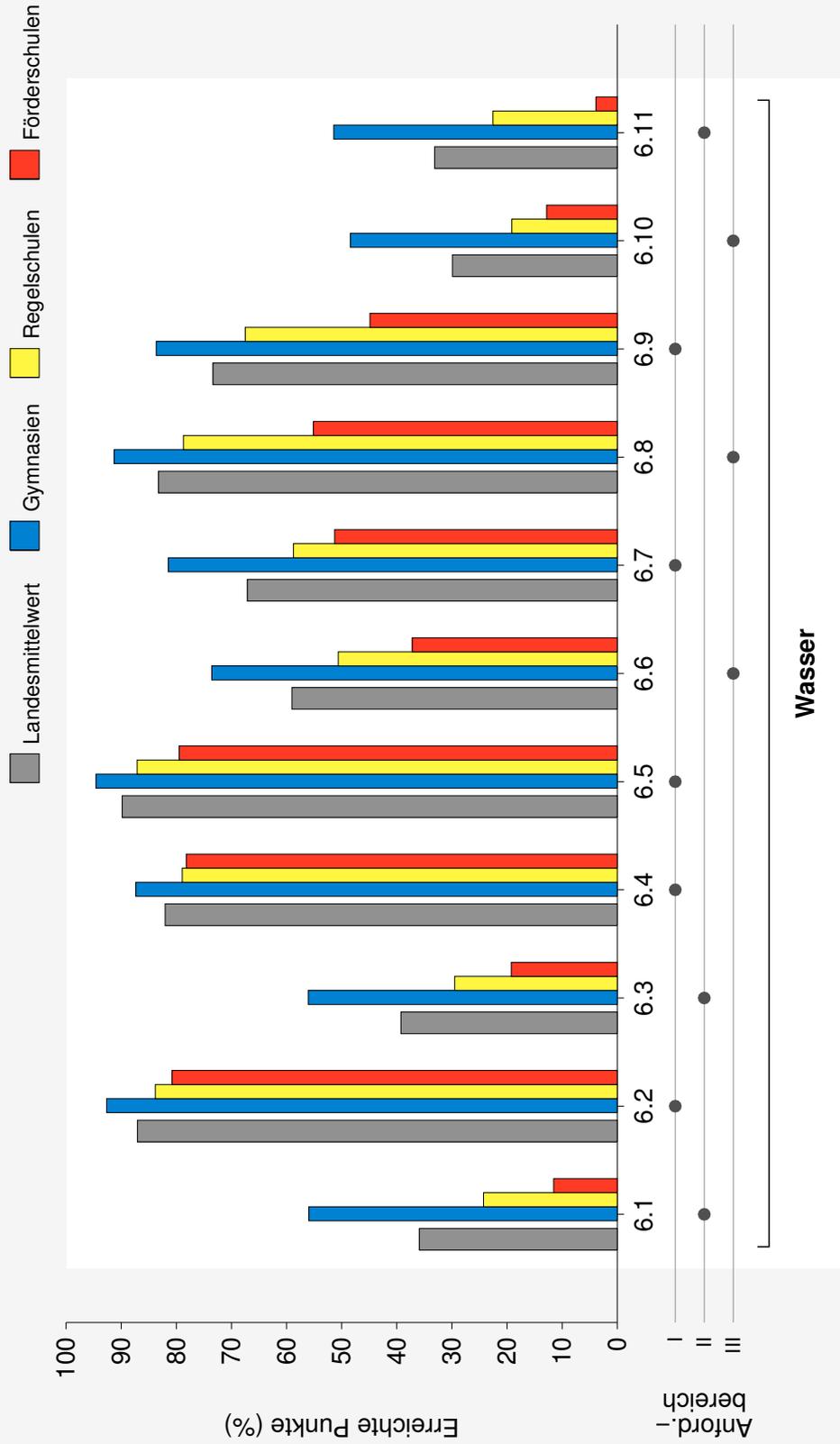


Abbildung 1.44: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 5)

Aufgabenergebnisse **DK8 (6/9)**

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
 Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

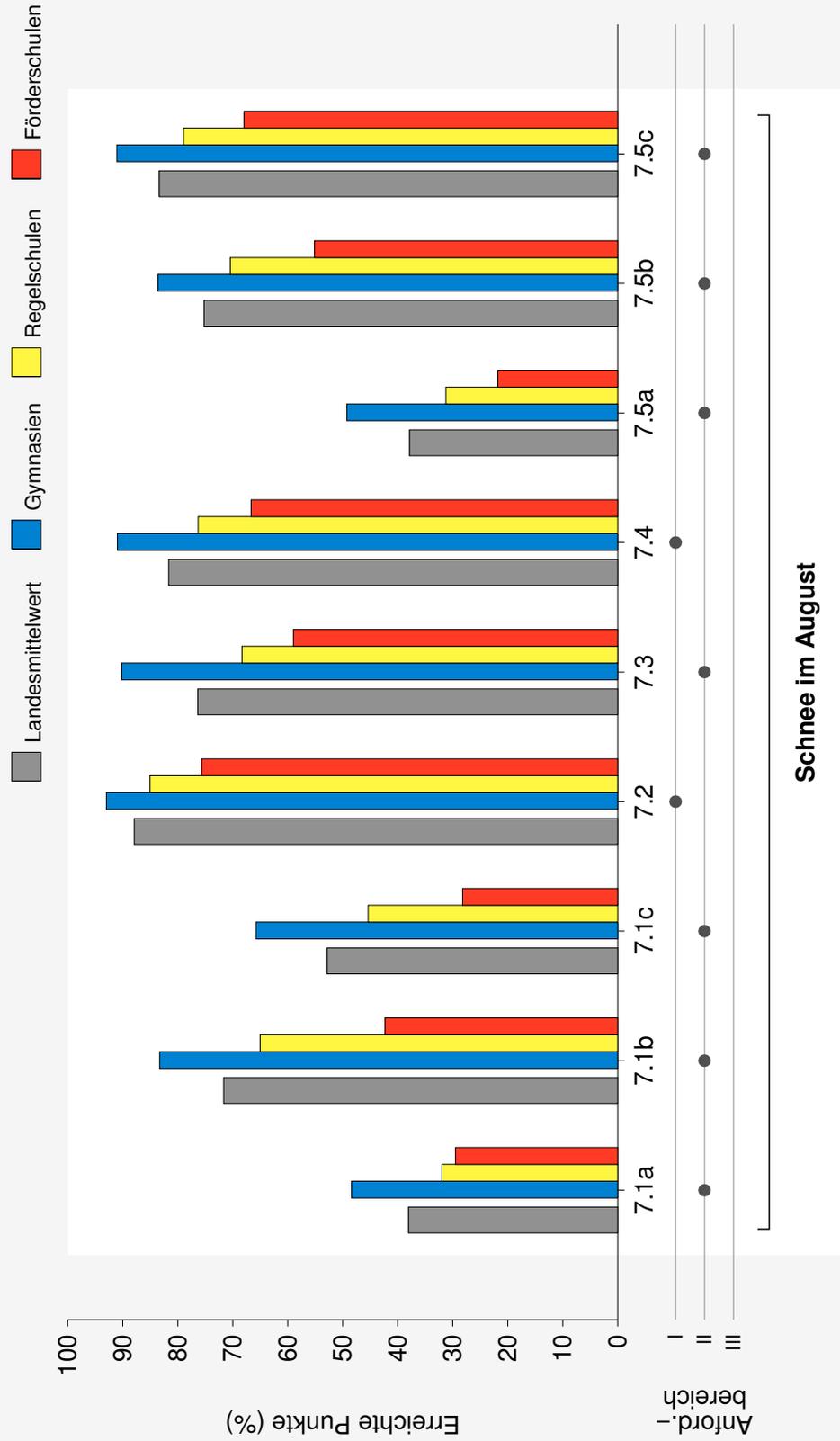
Abbildung 1.45: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 6)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

DK8 (7/9)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.46: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 7)

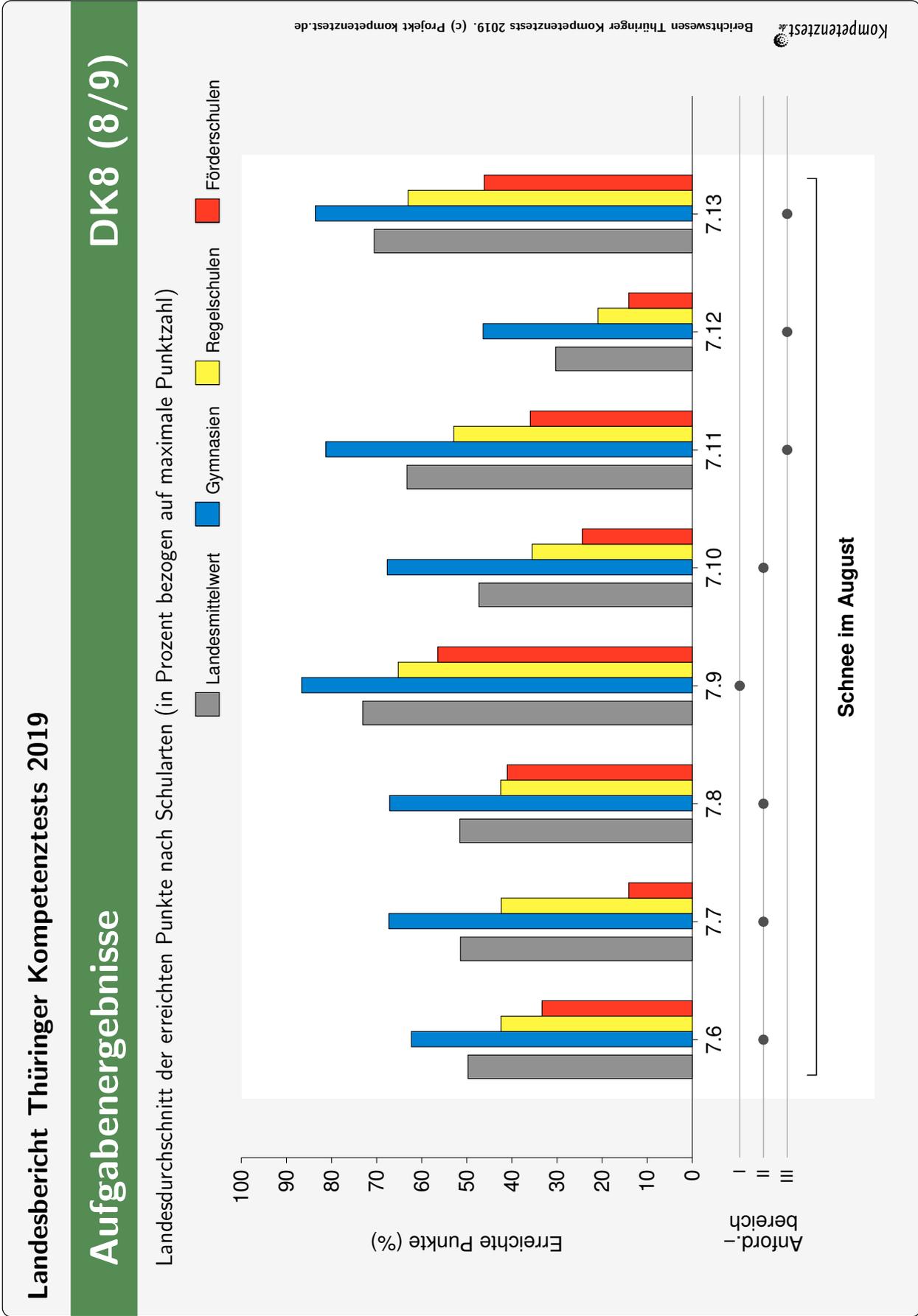


Abbildung 1.47: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 8)

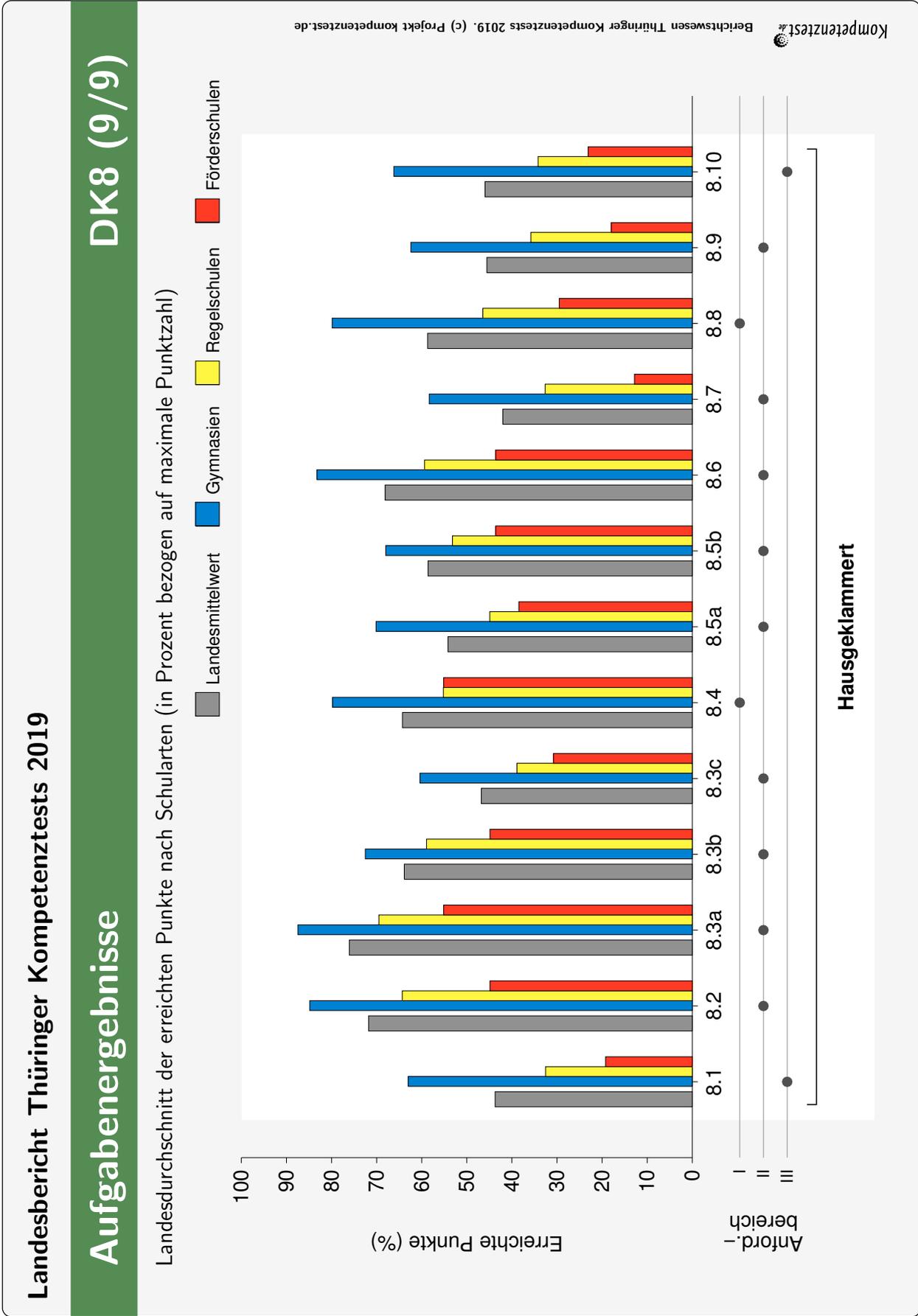


Abbildung 1.48: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Deutschtest, Klassenstufe 8 – Teil 9)

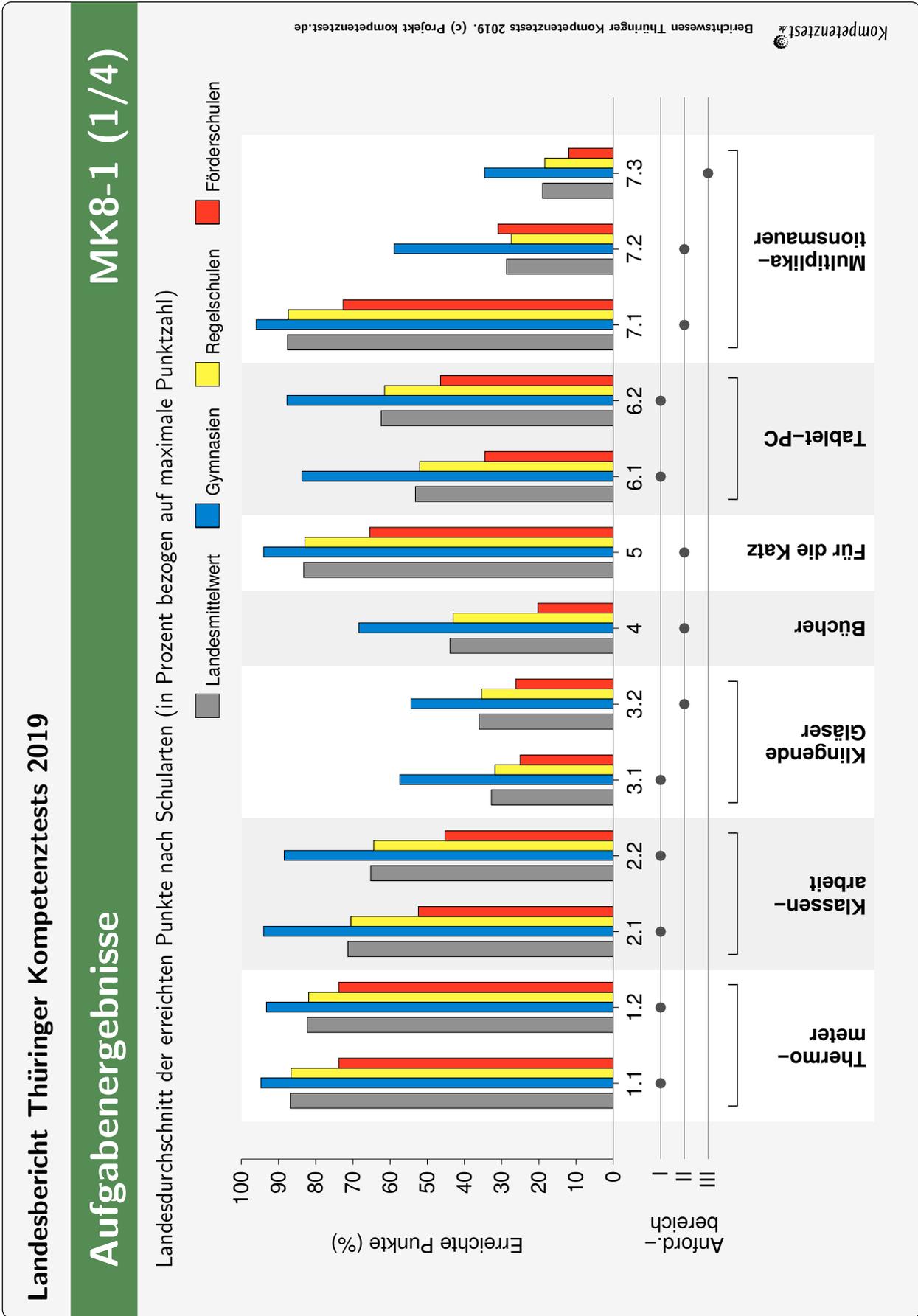


Abbildung 1.49: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 1, Klassenstufe 8 – Teil 1)

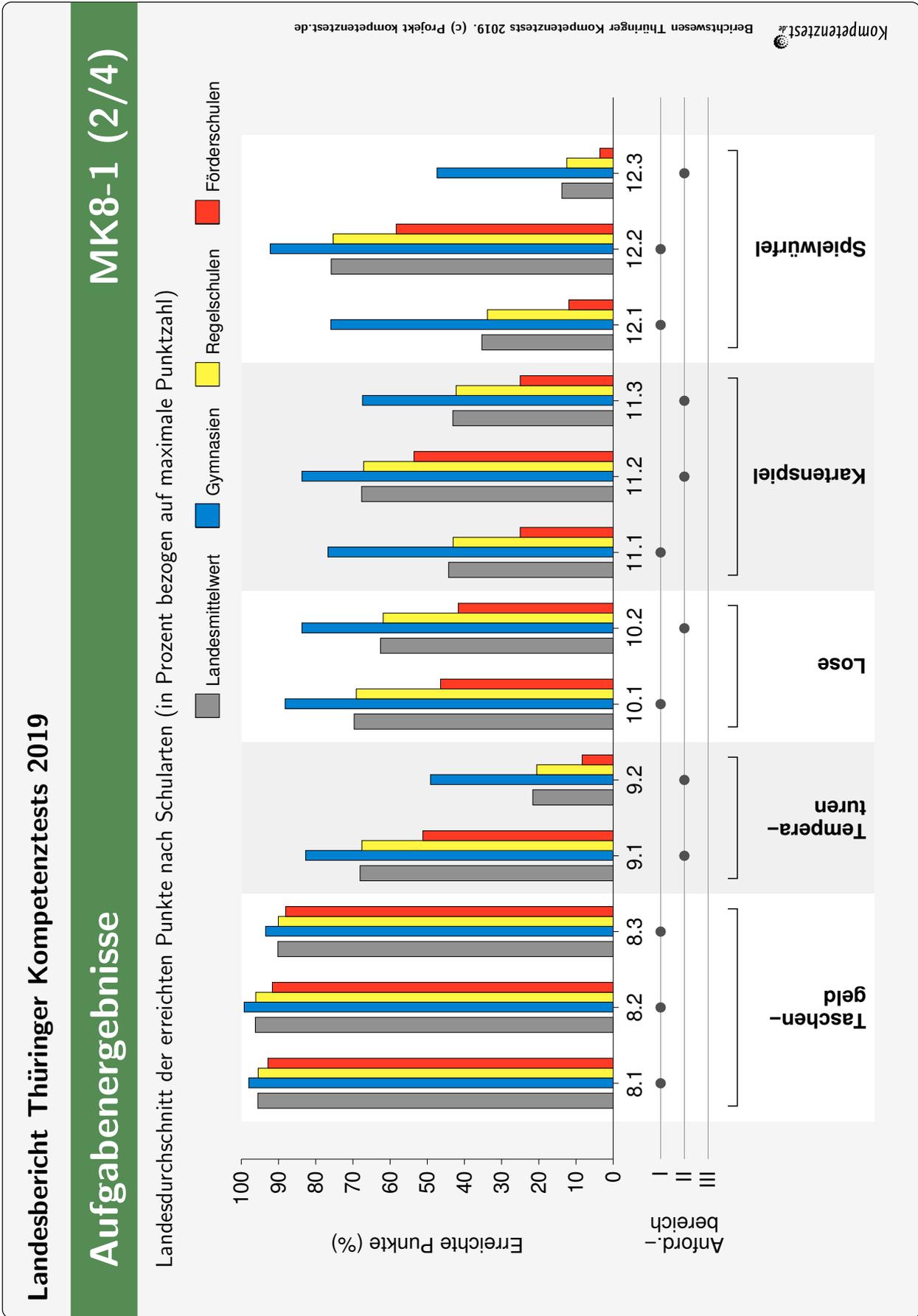


Abbildung 1.50: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 1, Klassenstufe 8 – Teil 2)

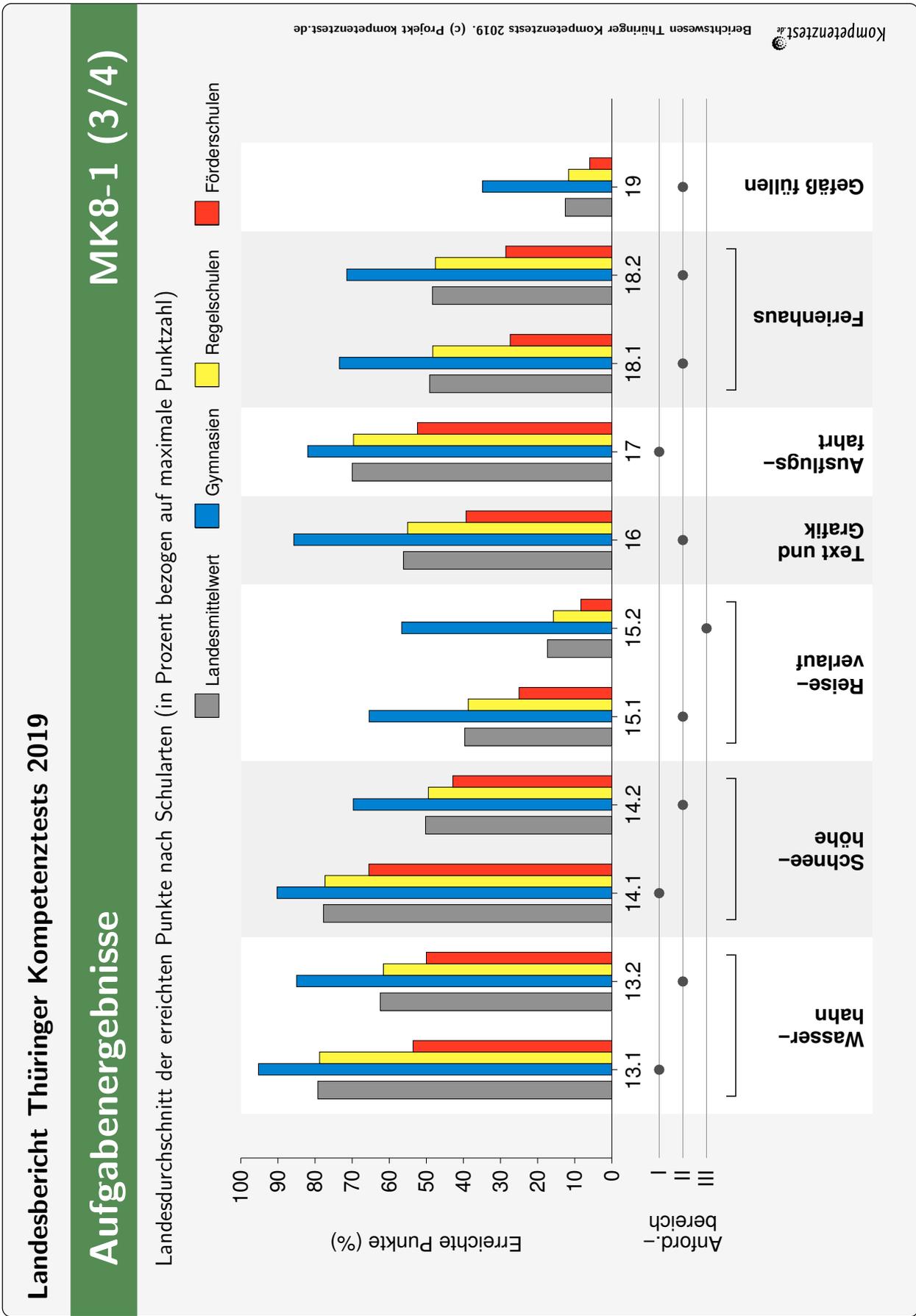


Abbildung 1.51: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 1, Klassenstufe 8 – Teil 3)

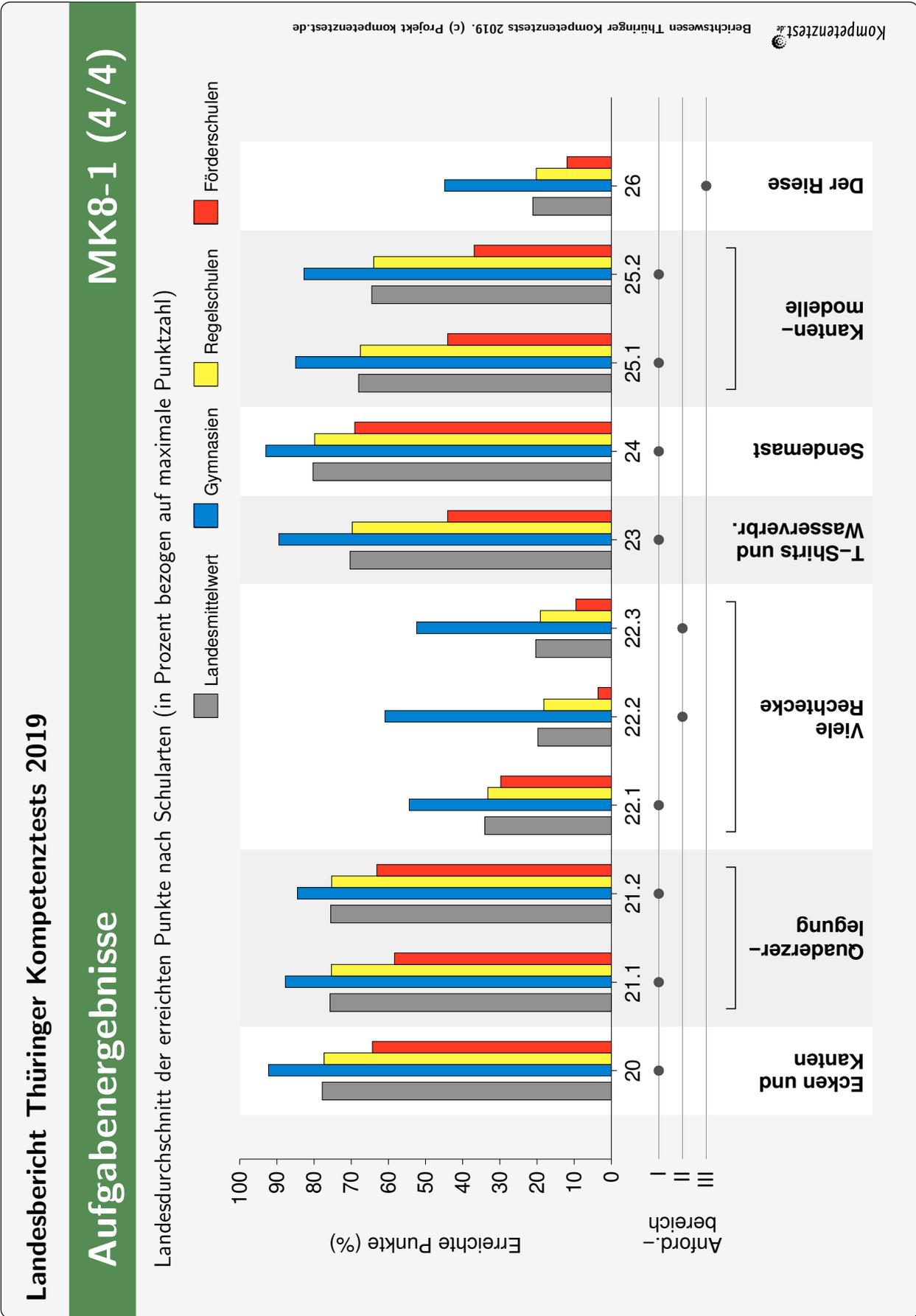


Abbildung 1.52: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 1, Klassenstufe 8 – Teil 4)

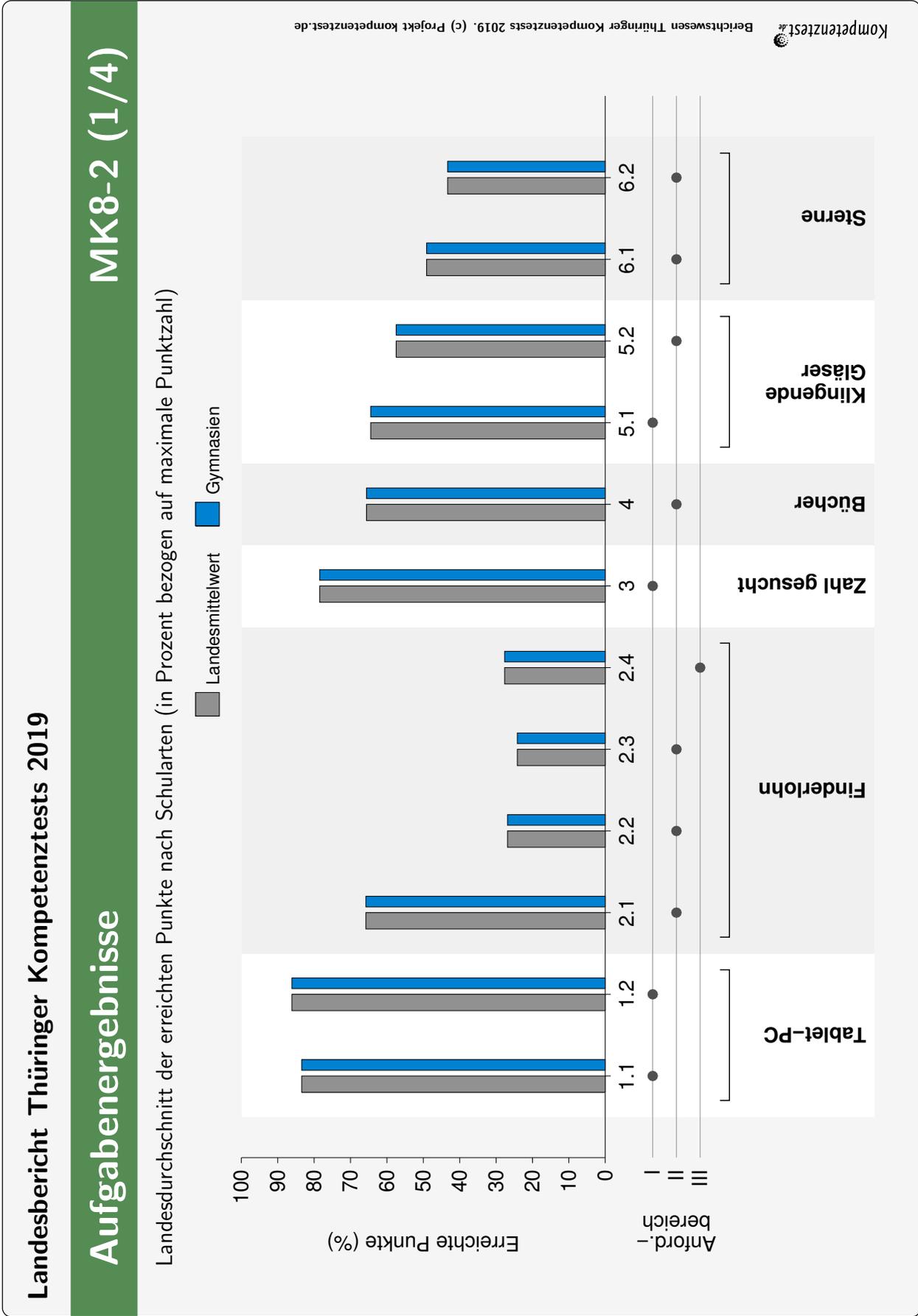


Abbildung 1.53: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 2, Klassenstufe 8 – Teil 1)

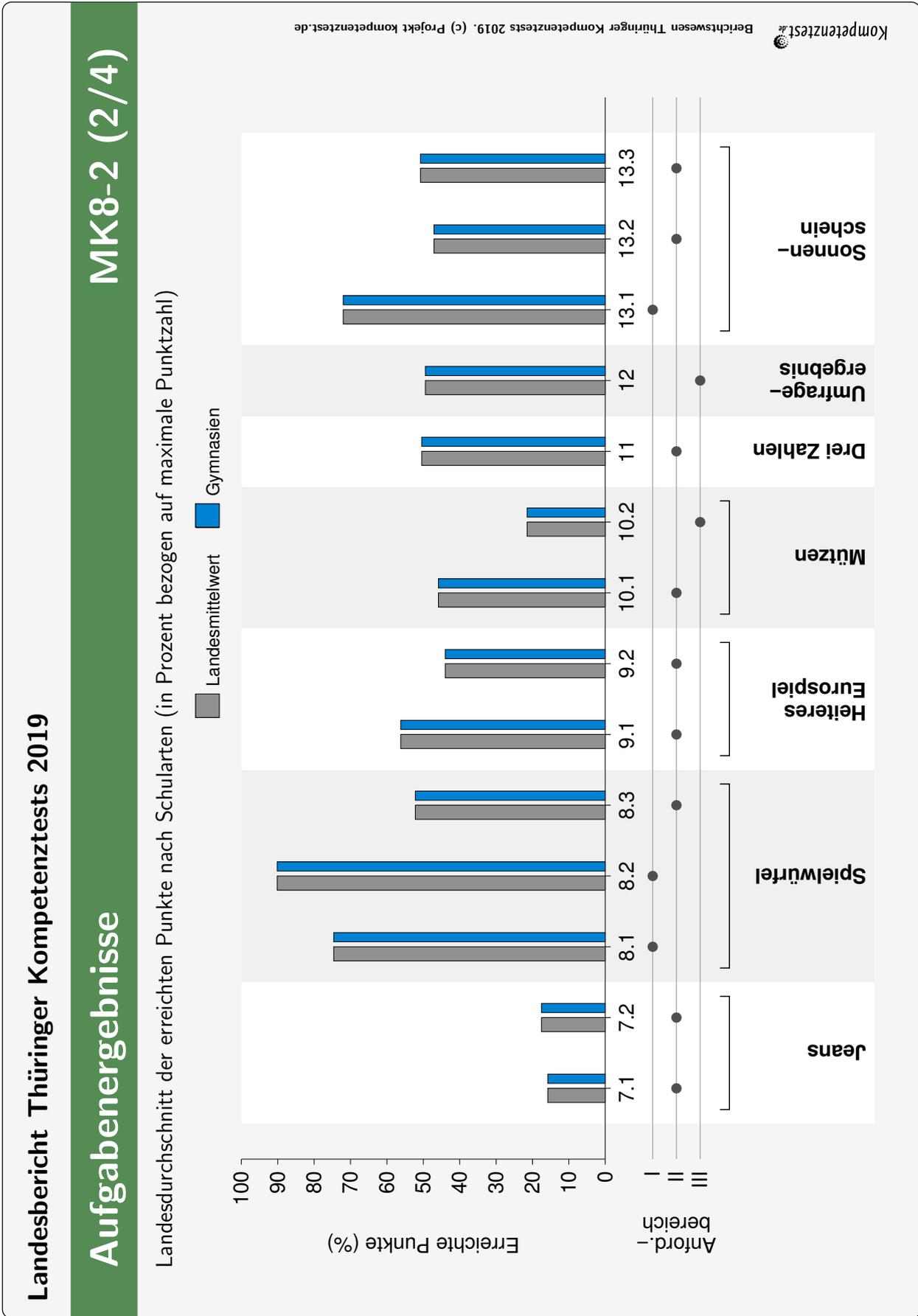


Abbildung 1.54: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 2, Klassenstufe 8 – Teil 2)

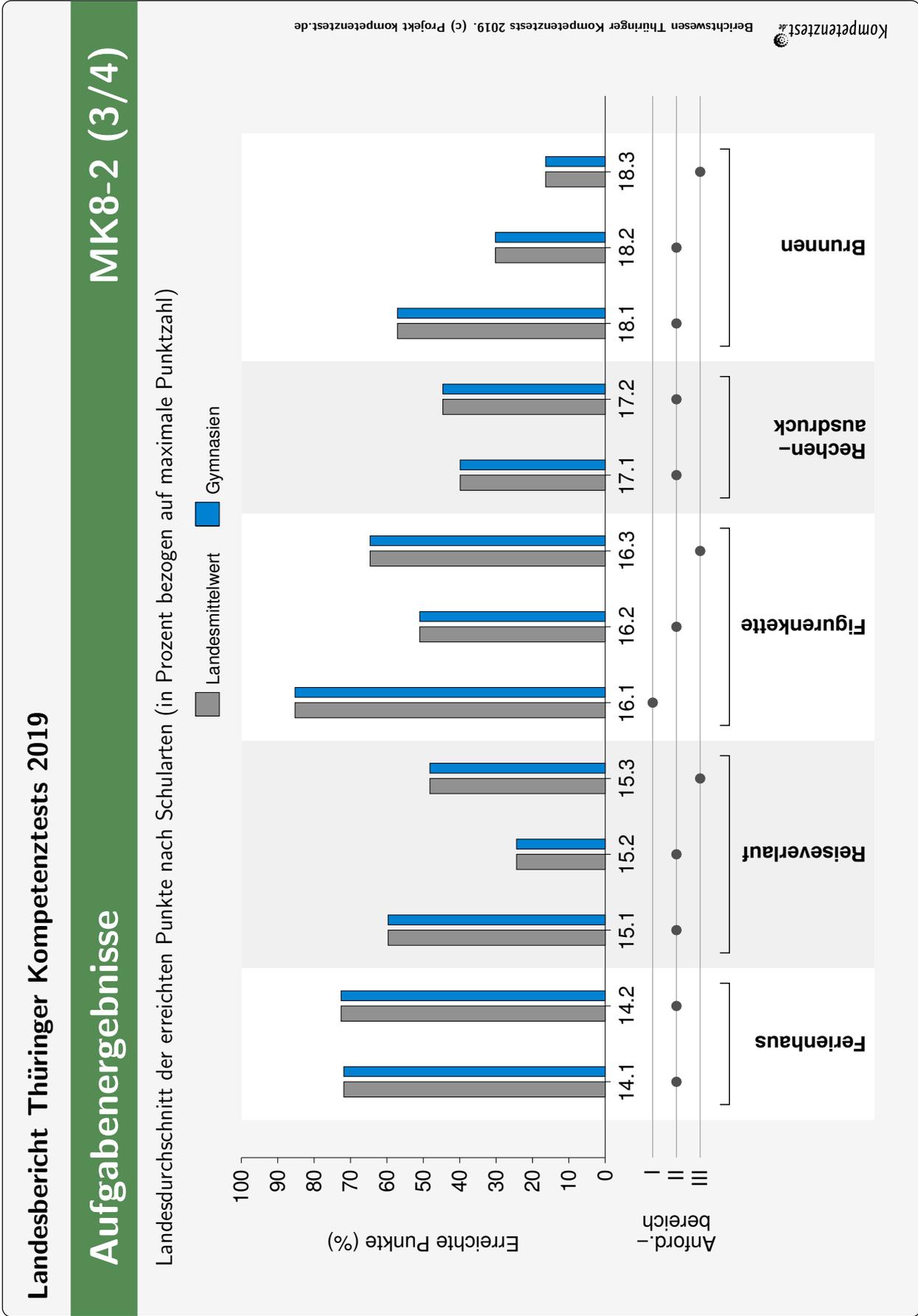


Abbildung 1.55: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 2, Klassenstufe 8 – Teil 3)

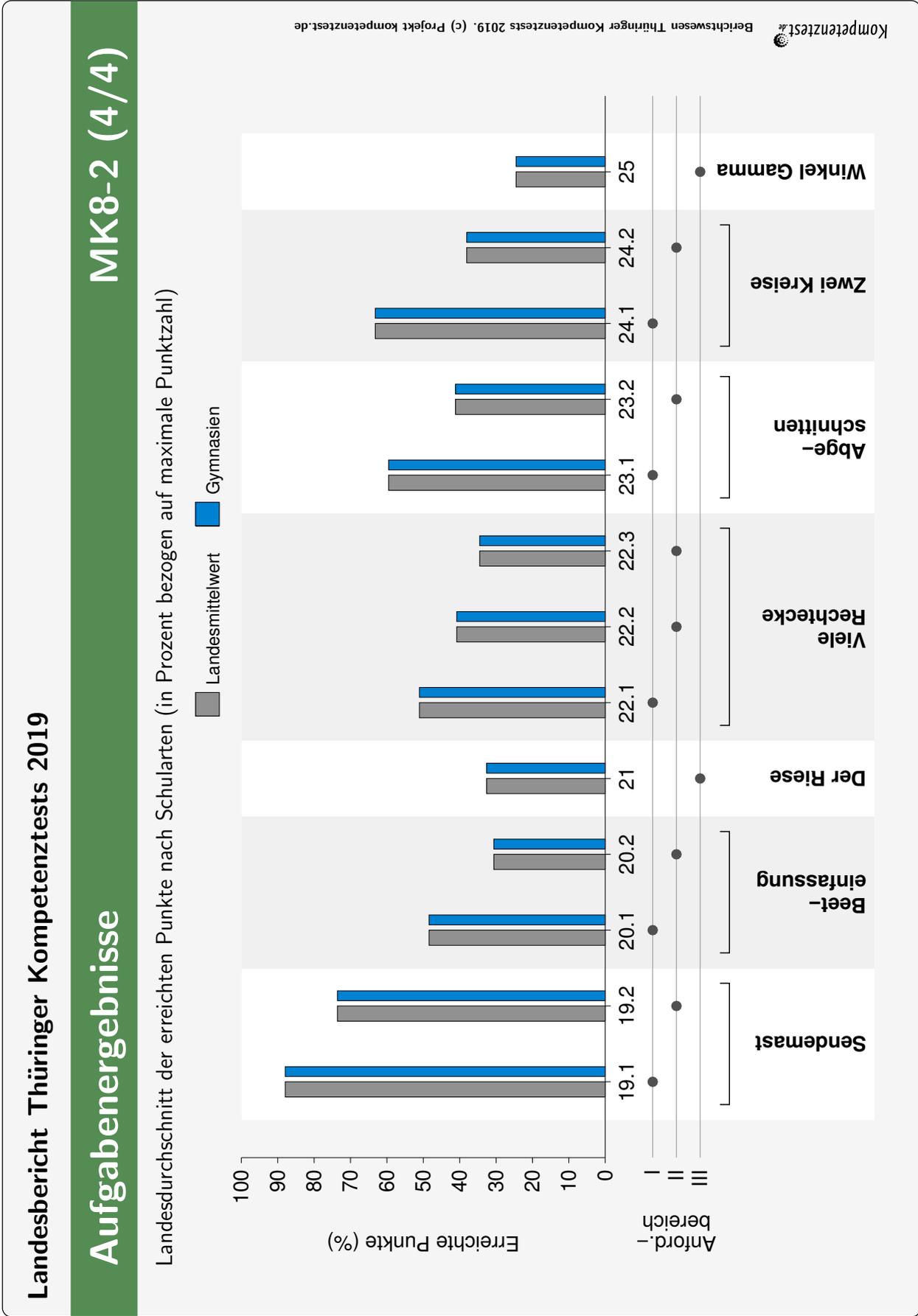


Abbildung 1.56: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Mathematiktest, Heft 2, Klassenstufe 8 – Teil 4)

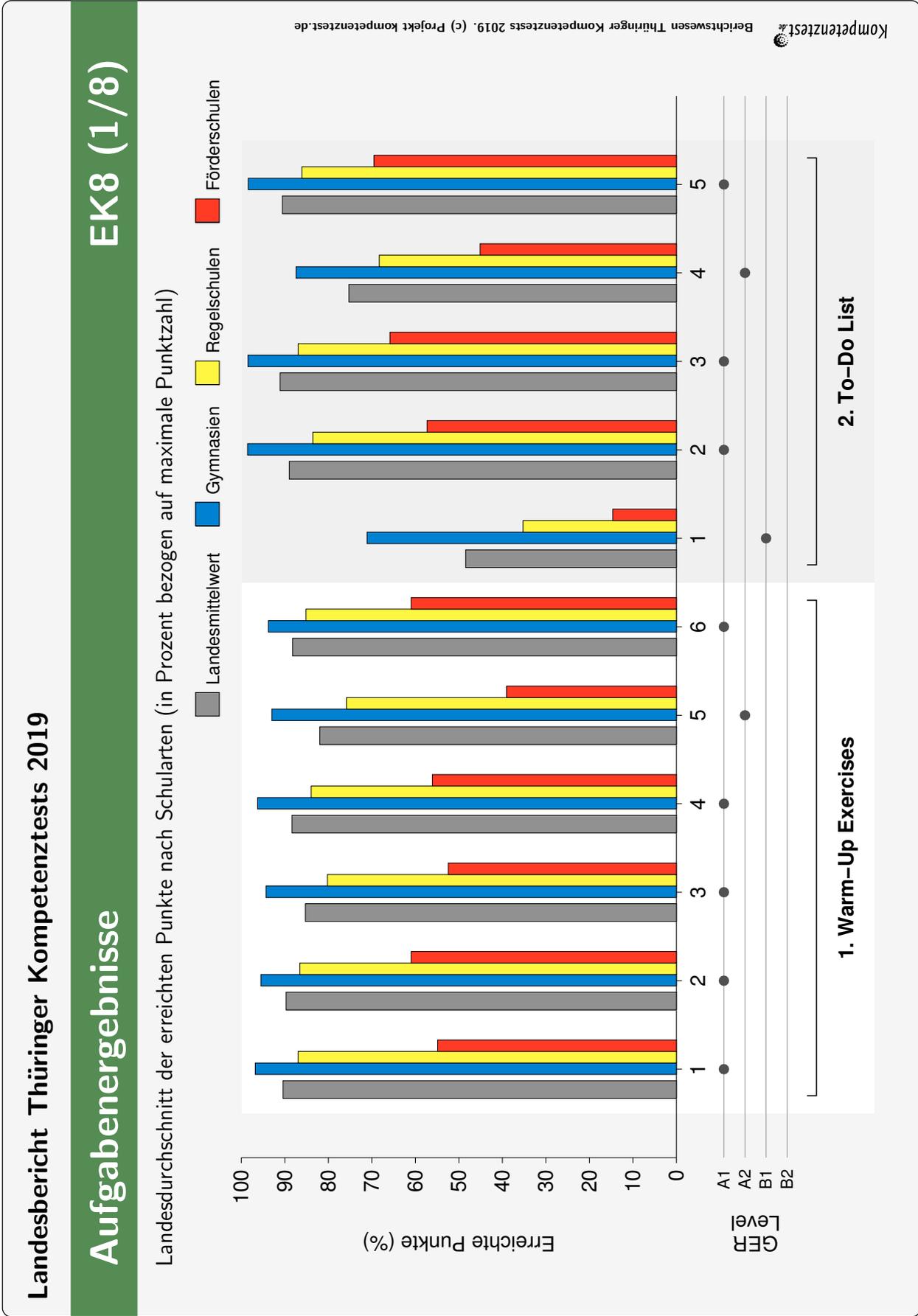


Abbildung 1.57: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englishtest, Klassenstufe 8 – Teil 1)

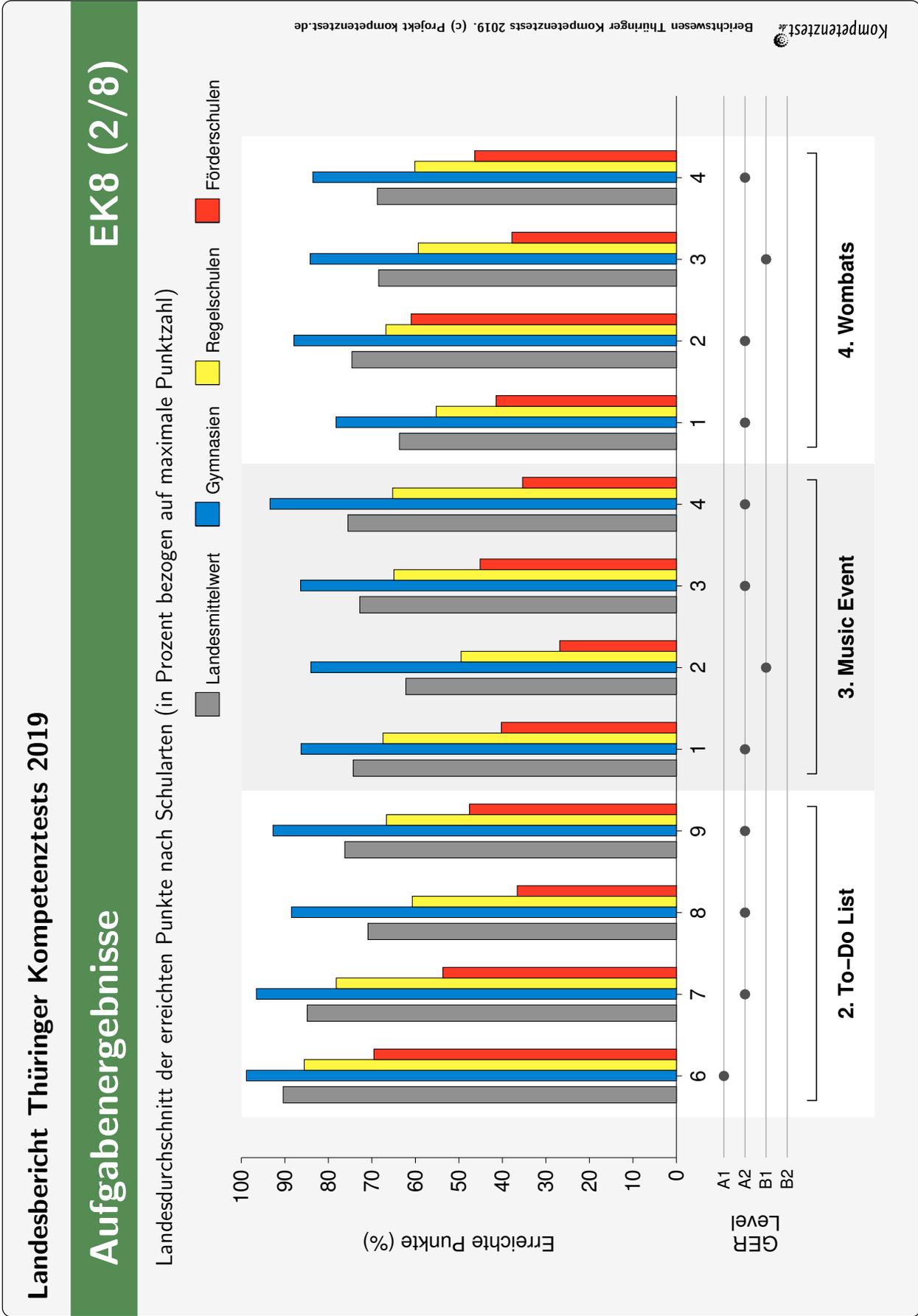


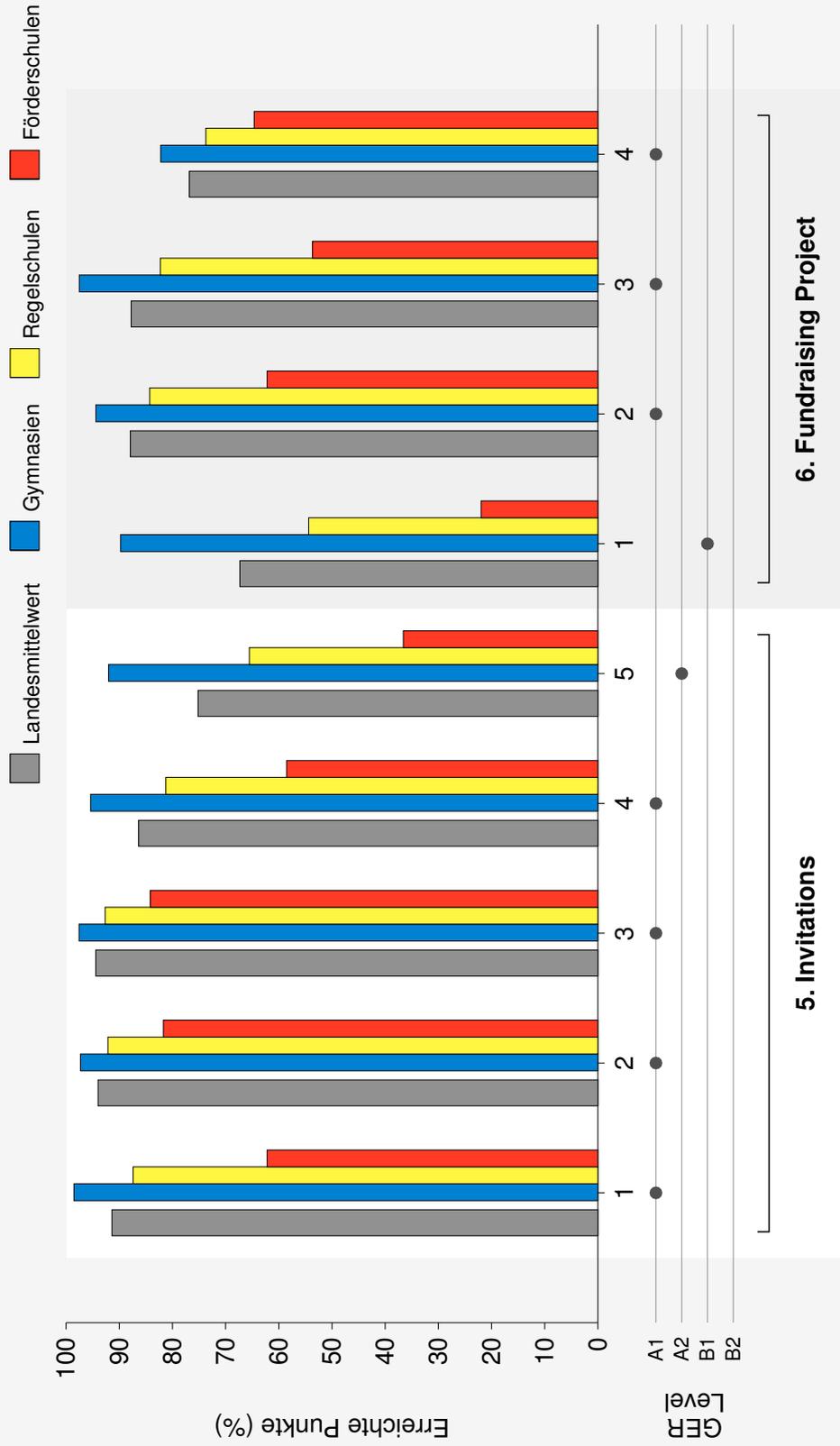
Abbildung 1.58: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englishtest, Klassenstufe 8 – Teil 2)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

EK8 (3/8)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

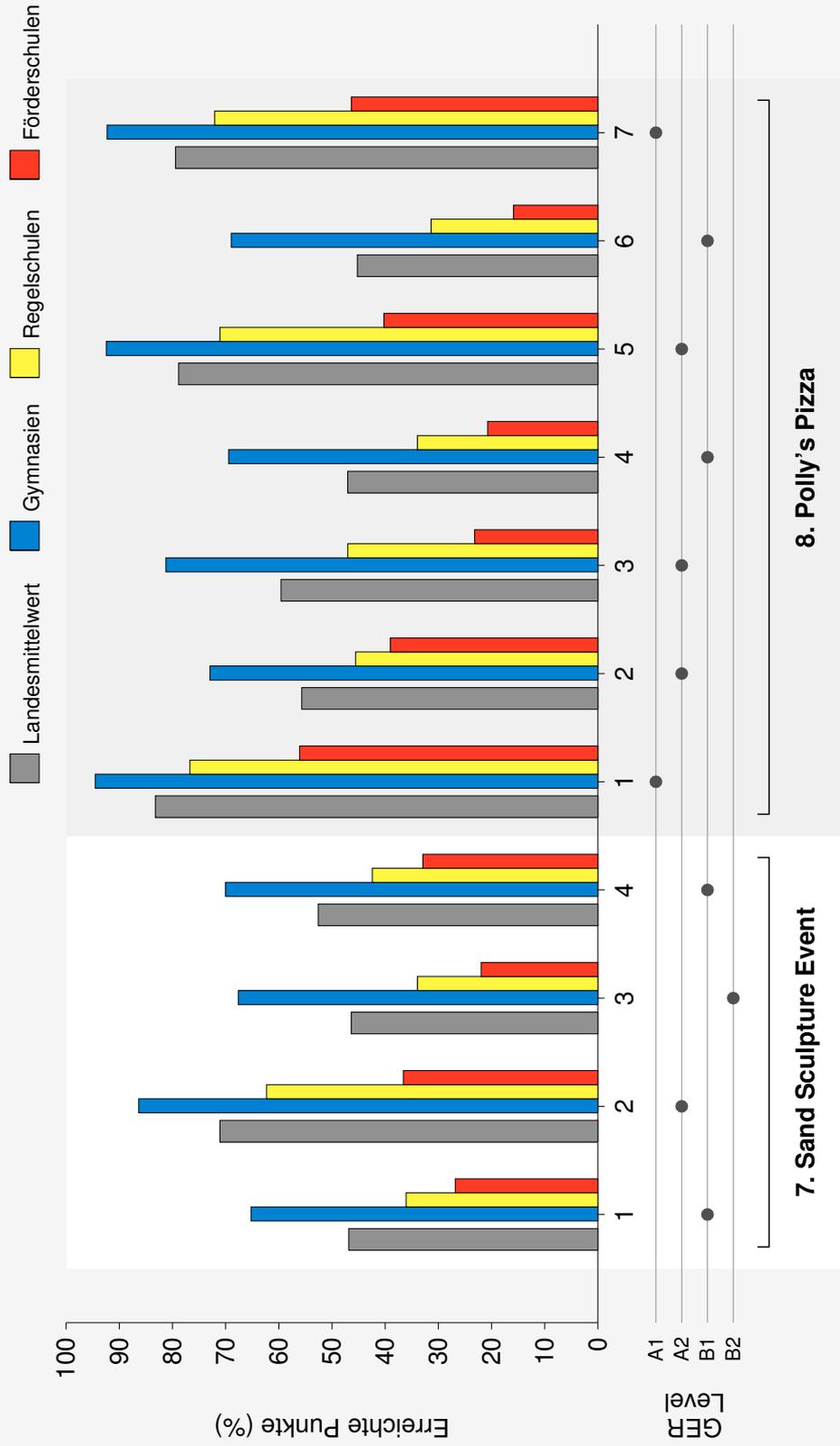
Abbildung 1.59: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englischtest, Klassenstufe 8 – Teil 3)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

EK8 (4/8)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

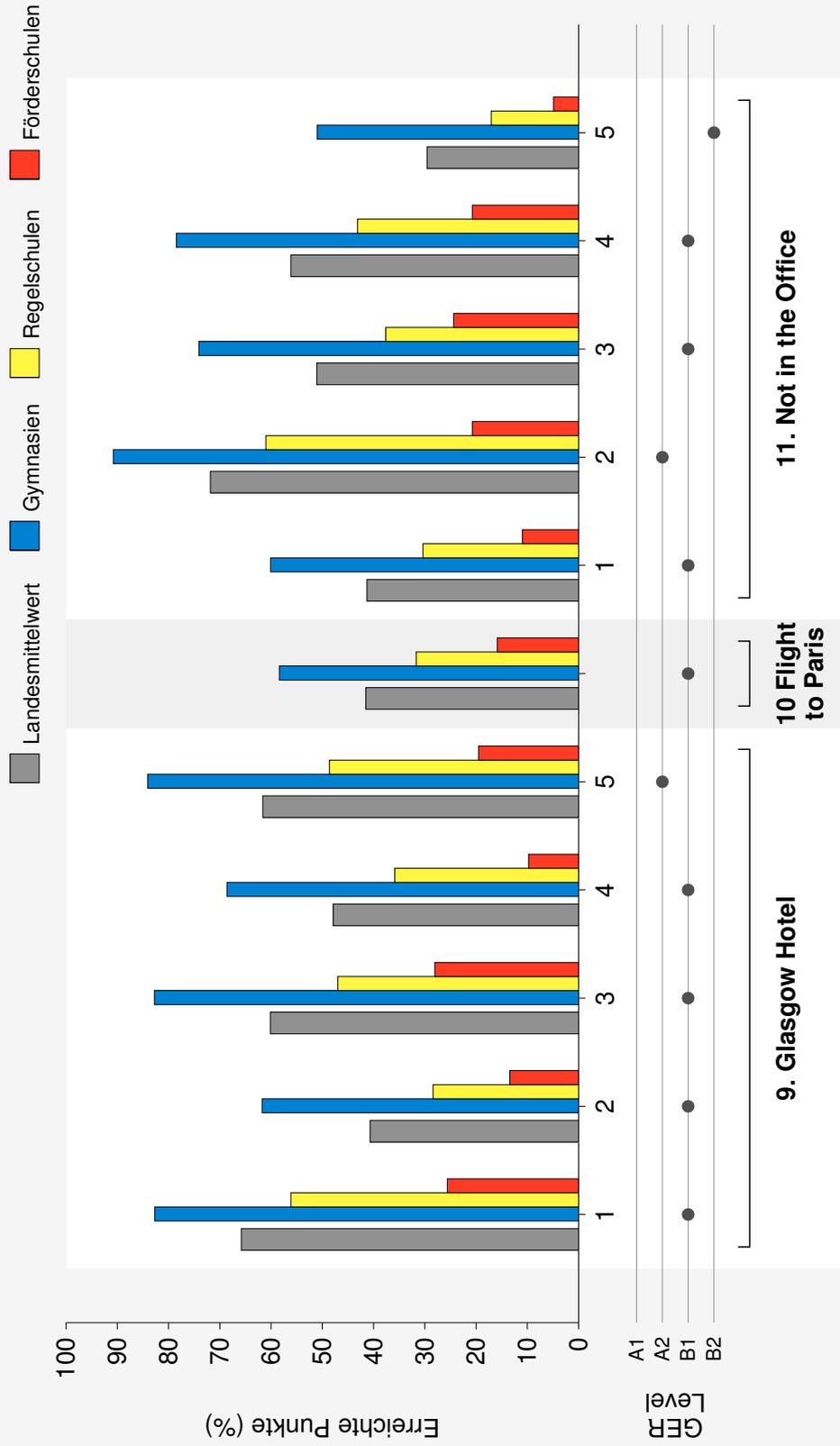
Abbildung 1.60: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englishtest, Klassenstufe 8 – Teil 4)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

EK8 (5/8)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.61: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englishtest, Klassenstufe 8 – Teil 5)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

EK8 (6/8)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)

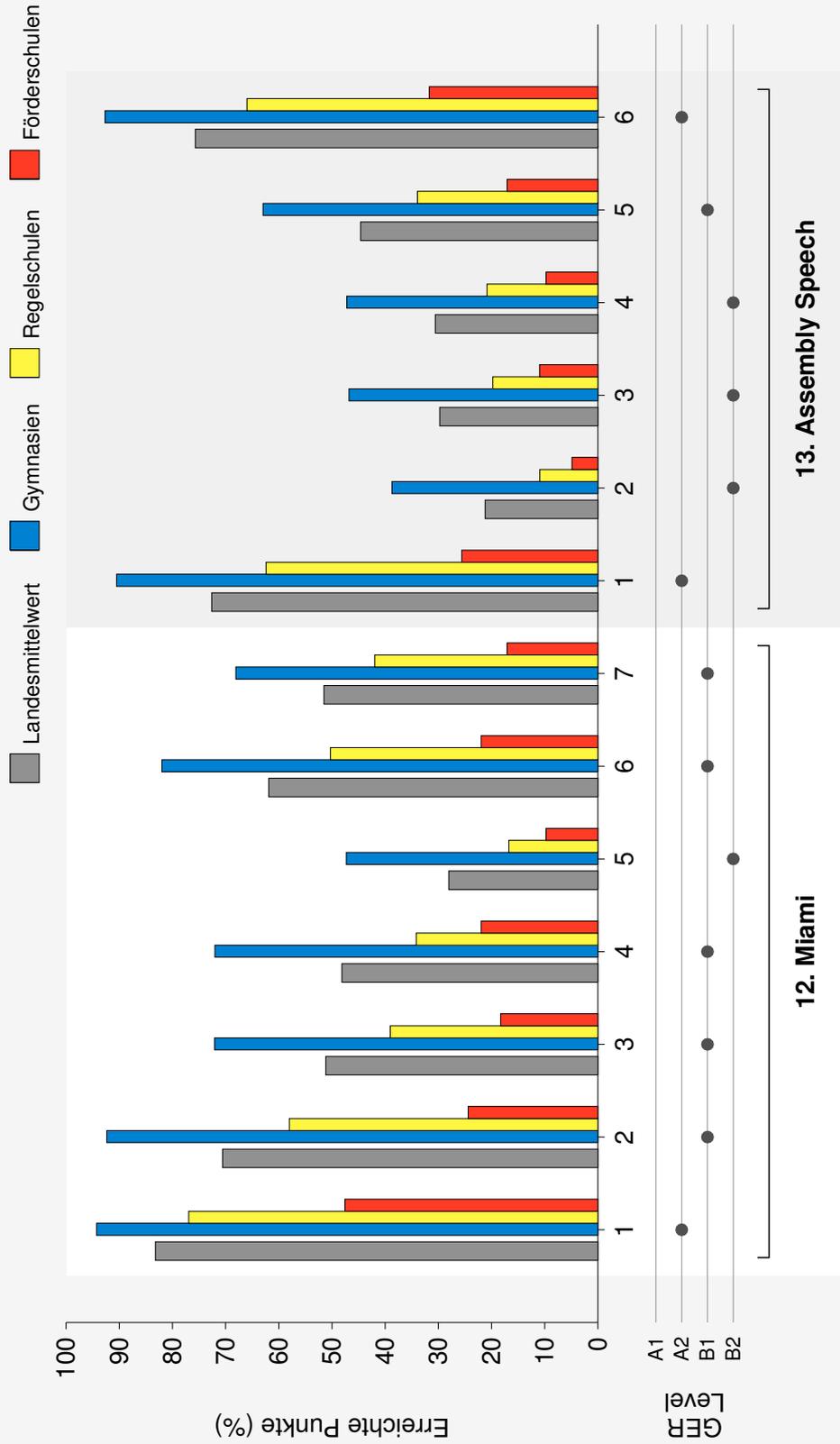


Abbildung 1.62: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englischtest, Klassenstufe 8 – Teil 6)

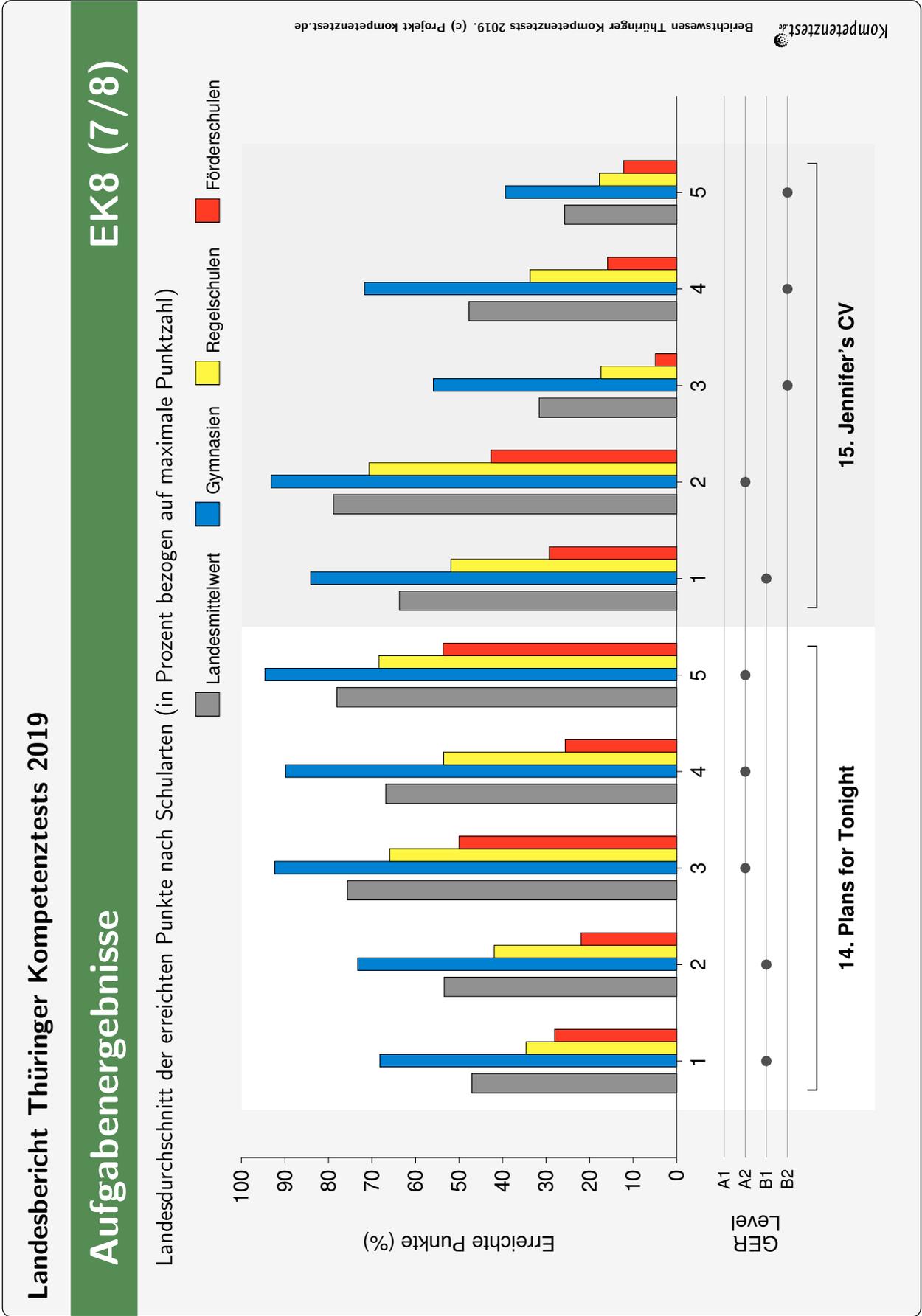


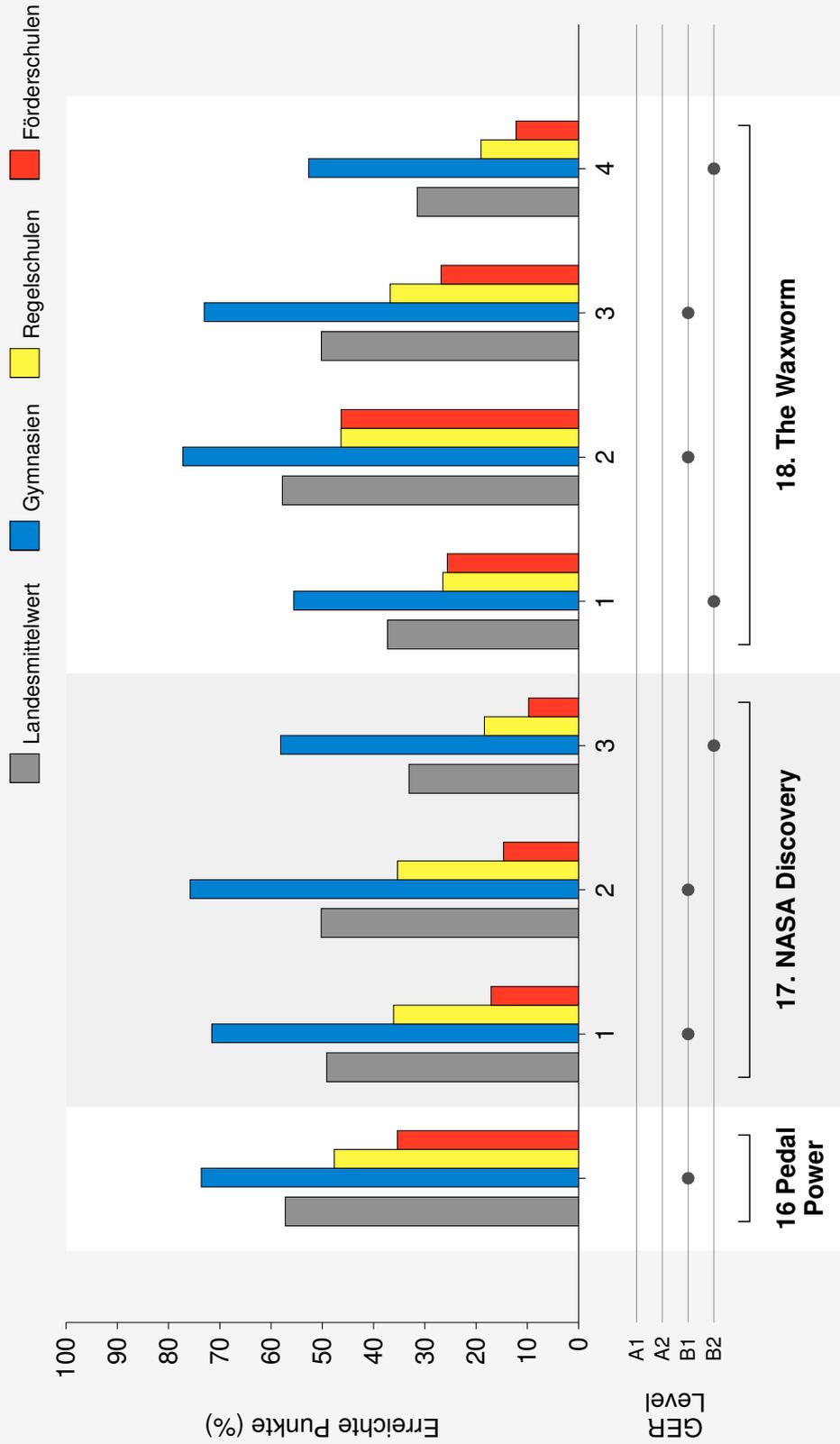
Abbildung 1.63: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englischtest, Klassenstufe 8 – Teil 7)

Landesbericht Thüringer Kompetenztests 2019

Aufgabenergebnisse

EK8 (8/8)

Landesdurchschnitt der erreichten Punkte nach Schularten (in Prozent bezogen auf maximale Punktzahl)



Kompetenztest.de
Berichtswesen Thüringer Kompetenztests 2019. (c) Projekt kompetenztest.de

Abbildung 1.64: Aufgabenschwierigkeiten im Landesdurchschnitt sowie getrennt nach Schulart (Englischtest, Klassenstufe 8 – Teil 8)

2 Schwerpunktthema 2019: Wiederholen von Klassenstufen

von Carolin Hellrung

Im Rahmen der Kompetenztests werden in jedem Jahr auch Einflussfaktoren auf die Testergebnisse untersucht. In diesem Jahr geht es dabei um die Auswirkung der Wiederholung einer Klassenstufe. Um deren Wirkung zu beurteilen, reicht es nicht aus, die Testleistungen von Schülern, die eine Klasse wiederholen, mit ihren Mitschülern zu vergleichen. Hier sind schwächere Testleistungen ohnehin zu erwarten, da schwächere Leistungen in der Regel der Grund für die Wiederholung einer Klassenstufe sind. Aufgrund des Thüringer Schülerlängsschnitts ist es jedoch möglich, frühere Testleistungen und die Lernentwicklungen zu berücksichtigen und sich so dem eigentlichen Effekt des "Sitzenbleibens" anzunähern.

2.1 Das Konstrukt "Sitzenbleiben"

Viele berühmte Personen haben es getan: Thomas Mann, Hermann Hesse, Winston Churchill, Harald Schmidt. Sie alle sind während ihrer Schulzeit „sitzengeblieben“. Seit der Einführung des Jahrgangsklassensystems im 18. Jahrhundert (vgl. Palowski, 2016) wiederholen Schüler ganze Klassenstufen. Das Konstrukt des Sitzenbleibens bzw. der Klassenwiederholung ist heutzutage fast jedem, der eine deutsche Schule besucht hat, bekannt. Nicht zwingend durch eigene Wiederholungserfahrungen, wohl aber durch jene von Klassenkameraden.

Wenn man Statistiken betrachtet, in denen die Zahlen der wiederholenden Schüler zusammengefasst werden, wird die Reichweite der Maßnahme „Nichtversetzung“ deutlich. Deutschlandweit wiederholten laut Statistischem Bundesamt (Destatis, 2018) im Schuljahr 2017/18 152.582 Schüler eine Jahrgangsstufe. Allein in Thüringen beliefen sich die Wiederholerzahlen auf 3.435 (Destatis, 2018). Somit befanden sich ca. 2 % aller Schüler an allgemeinbildenden Thüringer Schulen im Schuljahr 2017/18 in einem Wiederholungsjahr. Bei einer fiktiven Schülerzahl von 25 pro Klasse saß also durchschnittlich in jedem zweiten Klassenzimmer ein Sitzenbleiber.

Unter Sitzenbleiben wird das „nicht in die nächsthöhere Schulklasse versetzt werden“¹ verstanden. Die Entscheidung, einen Schüler nicht zu versetzen, wird aufgrund mangelnder Entwicklung bzw. Leistung des Schülers getroffen. Der Leistungsrückstand wird als zu groß eingeschätzt, als dass er im folgenden Schuljahr nebenbei aufgeholt werden könnte. Der Schüler wird zurückgestuft, um ihm damit die Möglichkeit zu geben, an seinen Schwächen in Ruhe zu arbeiten und die Anforderungen für die Versetzung in die nächsthöhere Klassenstufe zu erfüllen. Kurz gesagt: die Klassenstufenwiederholung soll zu einer Leistungsverbesserung des Schülers führen. Erfüllt die Maßnahme aber ihren Zweck? Welchen Einfluss hat das Wiederholen wirklich auf die Leistung des Schülers?

¹ https://www.duden.de/rechtschreibung/sitzen_bleiben

2.2 Forschungsstand über Folgen einer Klassenstufenwiederholung

Um die Auswirkungen der Wiederholung einer Klassenstufe kritisch diskutieren zu können, ist es interessant die bisherigen Erkenntnisse der empirischen Forschung zu diesem Thema zu betrachten.

Viele Studien zeigen auf, dass die Klassenstufenwiederholung für den Schüler eine große emotionale Belastung darstellt. Yamamoto (1980) fand beispielsweise, dass eine Klasse wiederholen stressvoller eingestuft wurde als „in der Klasse einnässen“ oder beim Stehlen erwischt zu werden. Eine Befragung von Kemmler (1976) ergab außerdem, dass sich Kinder mit Wiederholungserfahrung selbst als weniger konzentriert, angepasst und kontaktfreudig wahrnehmen. Weiterhin betont Bärsch (1987), dass die Wiederholung einer Klassenstufe „die größte Enttäuschung [ist], die ein Kind und auch seine Eltern in der Schule erleben“ (Bärsch, 1987, S. 27).

Am interessantesten sind jedoch die Erkenntnisse zu den Auswirkungen der Klassenstufenwiederholung auf die Leistung der betroffenen Schüler. Es stellt sich heraus, dass Wiederholer im Mittel schlechtere Leistungen zeigen, als immer versetzte Schüler (vgl. Nachtigall, 2018). Hattie (2015) sowie Holmes (1989) ermitteln einen negativen Effekt ($d = -.15$)² der Wiederholung sowohl auf die schulische Leistung als auch das Selbstkonzept des betroffenen Schülers. Auch wird ermittelt, dass Schüler, die zwei oder mehrere Male repetieren, schlechter seien als einmalige Wiederholer (vgl. Ingenkamp, 1976). Bei den erwähnten Studien bleibt fraglich, ob die gefundenen Werte so kausal als Effekt der Nichtversetzung interpretiert werden dürfen. Rein von der Definition her zeigen nicht versetzte Schüler sehr viel schwächere Leistungen als ihre Klassenkameraden. Deshalb ist der Zusammenhang mit den zuvor bereits sehr schwachen Leistungen zu beachten. Außerdem ist nicht klar, ob die schlechten Leistungen nach der Nichtversetzung zwingend durch diese bedingt werden oder die Leistungen durch andere Faktoren bedingt sind, welche auch zuvor eingewirkt haben. Mögliche andere Faktoren sind z. B. mangelhafte Leistungsmotivation oder eine schlechte Lehrer-Schüler-Beziehung.

Auch wenn die meisten Studien negative Auswirkungen der Wiederholung einer Klassenstufe ermitteln, gibt es Untersuchungen, welche unter bestimmten Bedingungen positive Leistungsänderungen aufzeigen. Über einen kurzen Zeitraum hinweg soll sich eine Klassenstufenwiederholung positiv auf die schulischen Leistungen auswirken (vgl. Jeske, 1981; Rust & Wallace, 1993). Ebenfalls soll die Schulform einen Moderator der Auswirkungen einer Wiederholung darstellen. Laut Kernade (1989) sei die Wiederholung einer Klassenstufe für Gymnasiasten förderlicher als für Haupt- bzw. Regelschüler. Offen bleibt bei diesen Untersuchungen, ob Selektionseffekte eine Rolle spielen. Eine mögliche Erklärung für die positiveren Auswirkungen der Klassenstufenwiederholung auf die Leistung von Gymnasialschülern ist, dass die Schüler, die sich für die Schulform des Gymnasiums entscheiden, auch eine höhere Leistungsmotivation haben. Dadurch kann die Chance der Wiederholung eher zur Verbesserung der eigenen Leistungen genutzt werden. Dem kommt hinzu, dass nicht versetzten Schülern mit schlechteren Prognosen eher zu einem Schulwechsel geraten wird. Auch dadurch haben Gymnasialklassen weniger Wiederholer und seltener wiederholende Schüler mit schlechter Prognose.

² Dies ist ein statistischer Kennwert, der auf Grundlage der Unterschiede zwischen Mittelwerten errechnet wird. Er wird im Anhang zusammen mit dem Begriff "Effekt" genauer erläutert.

2.3 Neue Untersuchungen

Der Datensatz des Thüringer Schülerlängsschnitts enthält u. a. Ergebnisse der Thüringer Kompetenztests der vergangenen Jahre und bietet die Möglichkeit Leistungsdaten von Schülern über die Zeit verknüpft zu betrachten. Damit eröffnet der Datensatz die Chance, den kausalen Bezug zwischen der Klassenstufenwiederholung und den Leistungen besser herauszustellen. Neben den Ergebnissen der Kompetenztests beinhaltet der Thüringer Schülerlängsschnitt auch weitere Information z. B. zu Schulform und Halbjahresnote. Vor dem Hintergrund dieses Datensatzes und mit Blick auf die vergangenen Studienergebnisse zum Thema der Klassenstufenwiederholung ergeben sich einige interessante Hypothesen, die im Folgenden kritisch untersucht werden:

1. Die Testleistungen der Wiederholer sind in den einzelnen getesteten Fächern signifikant schlechter als die Leistungen der Nichtwiederholer.
2. Die Leistungen eines Schülers, der in zwei aufeinanderfolgenden Schuljahren an einem Test derselben Klassenstufe teilnimmt, unterscheiden sich. Der wiederholende Schüler erzielt im zweiten Jahr bessere Ergebnisse.
3. Wiederholer erreichen im Test der Klassenstufe 8 ein schlechteres Ergebnis als Nichtwiederholer, selbst wenn Sie in den Kompetenztests der Klassenstufe 3 ähnliche Ergebnisse erzielten.
4. Gymnasiasten profitieren in ihrem Leistungszuwachs stärker von einem wiederholten Schuljahr als Real- bzw. Hauptschüler.
5. Wiederholer verbessern sich vor allem in den Fächern, in welchen vorher besonders schwache Noten vorlagen.

2.4 Methoden und Ergebnisse

Für die durchgeführten Untersuchungen wurden die Ergebnisse der Thüringer Kompetenztests aus den Jahren 2013, 2014 sowie 2018 und 2019 analysiert. Mit Hilfe eines statistischen Verfahrens (z-Transformation) wurden die Daten der verschiedenen Tests vergleichbar gemacht. Anschließend wurden verschiedene statistische Tests durchgeführt.

1. **Die Testleistungen der Wiederholer sind in den einzelnen getesteten Fächern signifikant schlechter als die Leistungen der Nichtwiederholer.**

Für die Überprüfung dieser Hypothese wurden die Testergebnisse der Kompetenztests 2019 von Schülern mit Wiederholungserfahrung und solchen ohne vergangener Klassenstufenwiederholung verglichen. Dabei zeigten sich im Schnitt Effektstärken mittlerer Größe ($d = -.70$)³. Schüler mit Wiederholungserfahrung erreichen demnach im Durchschnitt schlechtere Ergebnisse im selben Test als Schüler, die immer versetzt wurden. Hierbei ist jedoch zu beachten, dass die Wiederholer schließlich wegen nicht ausreichenden Leistungen eine Jahrgangsstufe wiederholen mussten. Deshalb ist dieses Ergebnis nicht überraschend. Es zeigt, die vorher vorhandenen Leistungsunterschiede nicht verschwinden. Für die Ermittlung des eigentlichen Effektes der Wiederholung werden zusätzliche Untersuchungen benötigt.

³ Erläuterung der Effektstärken im Anhang.

- 2. Die Leistungen eines Schülers, der in zwei aufeinanderfolgenden Schuljahren an einem Test derselben Klassenstufe teilnimmt, unterscheiden sich. Der wiederholende Schüler erzielt im zweiten Jahr bessere Ergebnisse.**

Für die Untersuchung dieser Aussage wurden die Testdaten der Kompetenztests 2018 bzw. 2019 in der Klassenstufe 8 verwendet. Dabei wurde geschaut, ob Schüler, die in beiden Jahren am Test teilgenommen haben, eine Veränderung in der Leistung zeigen und wenn ja welche. Es zeigt sich, dass Wiederholer beim zweiten Mal im Mittel ein signifikant besseres Ergebnis in den Tests erzielen, als beim ersten Durchgang. Die dabei ermittelten Effektstärken liegen im mittleren Bereich ($d = .49$).

Auch bei diesem Ergebnis gibt es einige Aspekte, denen Beachtung geschenkt werden sollte. Dieser Lernzuwachs muss keineswegs die Wirkung der Klassenstufenwiederholung darstellen. In jedem Jahr finden Reifungsprozesse statt, welche den Zuwachs erklären könnten. Für die Ermittlung des kausalen Effekts einer Klassenstufenwiederholung wäre zu prüfen, inwiefern sich das Wiederholen bzw. Nichtwiederholen einer Klassenstufe, unter sonst gleichen Bedingungen, auf die Leistung eines Schülers auswirken würde. Eine Annäherung an eine solche Analyse stellt die Prüfung der nächsten Hypothese dar.

- 3. Wiederholer erreichen im Test der Klassenstufe 8 ein schlechteres Ergebnis als Nichtwiederholer, selbst wenn Sie in den Kompetenztests der Klassenstufe 3 ähnliche Ergebnisse erzielten.**

Ausgangspunkt für die Untersuchung dieser Hypothese sind die Ergebnisse der Kompetenztests der Klassenstufe 8 im Jahr 2019. Von dort aus wurden rückblickend verknüpfte Daten der anderen Jahrgänge verwendet. Die Testergebnisse aus den Jahren 2013 und 2014 stellen die sogenannten Vortestwerte dar. Dies sind die Ergebnisse der Kompetenztests der gleichen Schüler in der Klassenstufe 3. Für Schüler, die immer versetzt worden sind, ist das Ergebnis der 3. Klassenstufe im Datensatz der Thüringer Kompetenztests 2014 enthalten. Für Schüler mit Wiederholungserfahrung (ein Jahr) ist das Kompetenztestergebnis der 3. Klassenstufe im Jahr 2013 zu finden.

Bei gleichem Vortestwert zeigt sich im Mittel für alle Tests (Deutsch, Mathematik) ein signifikanter negativer Effekt des Wiederholerstatus auf die Leistung der Schüler. Schüler mit Wiederholerstatus schneiden in Klassenstufe 8 systematisch schwächer ab als ihre Mitschüler, selbst wenn sie in Klassenstufe 3 vergleichbare Leistungen zeigten. Bei der Interpretation dieses Ergebnisses ist jedoch zu beachten, dass es sich um eine sehr große Zeitspanne handelt. In den 5 bzw. 6 Jahren zwischen den einbezogenen Tests gibt es viele Möglichkeiten für Einflüsse, die in diesem Format nicht erfasst und kontrolliert werden konnten. Relevante Einflüsse sind z. B. Schulwechsel, Unterrichtsform (z. B. Frontalunterricht) oder auch das Geschlechterverhältnis im Klassenverband.

Bei den nächsten beiden Hypothesen wird zusätzlich nach weiteren Einflussfaktoren wie der Schulform oder dem Fach differenziert.

- 4. Gymnasiasten profitieren in ihrem Leistungszuwachs stärker von einem wiederholten Schuljahr als Real- bzw. Hauptschüler.**

Für die vierte Hypothese wurde geprüft, ob sich die Differenzen der Testleistungen von Wiederholern zwischen den zwei aufeinanderfolgenden Schuljahren der 8. Klassenstufe (2018, 2019) nach Beschulungsform unterscheiden. Da bisherige Studienergebnisse dar-

auf hinweisen, dass Gymnasialschüler positive Auswirkungen der Wiederholung auf die schulische Leistung zeigen, wurden die beiden Gruppen "Gymnasium" und "nicht Gymnasium" im Zuge dieser Untersuchung verglichen. In den Testfächern Deutsch und Mathematik zeigt sich kein signifikanter Unterschied zwischen den Gruppen. Für das Fach Englisch stellt sich jedoch ein bedeutsamer Unterschied (mit mittlerer Effektstärke) heraus. Demzufolge haben Gymnasiasten im Fach Englisch im Mittel einen signifikant größeren Leistungszuwachs in einem wiederholten Schuljahr als Schüler der anderen Schulformen.

Bei der Betrachtung von Unterschieden zwischen Beschulungsformen sind jedoch mögliche Selektionseffekte zu beachten. Schülern mit schlechterer Prognose wird eher zu einem Schulwechsel geraten, wodurch Gymnasialklassen weniger Wiederholer und seltener wiederholende Schüler mit schlechter Prognose haben.

5. Wiederholer verbessern sich vor allem in den Fächern, in welchen vorher besonders schwache Noten vorlagen.

Für die Untersuchung der Hypothese Nummer fünf wurde überprüft, ob sich die zuvor schwächste Note von wiederholenden Schülern im Vergleich zum Durchschnitt der übrigen Schulnoten dieser Schüler im Folgejahr am meisten verbessert hat. Genauer gesagt wurden die Differenzwerte der Zeugnisnoten von wiederholenden Schülern in den beiden aufeinanderfolgenden Schuljahren verglichen und untersucht, ob die zuvor schwächste Note auch die größte positive Veränderung aufweist. Dabei wurden die Daten aus den Jahren 2018 und 2019 verwendet. Fälle, bei denen mehrere Noten gleichzeitig die "schwächste" darstellten, wurden ausgelassen.

Es stellte sich ein signifikanter Effekt der schlechten Note heraus. Dieser ist jedoch von kleiner Größe ($d = .36$). Bei diesem Ergebnis ist zu beachten, dass bei einer schlechteren Note eher die Möglichkeit zu einer Verbesserung besteht bzw. eine größere Verbesserung möglich ist als bei guten Noten.

2.5 Fazit

Die durchgeführten Analysen bestätigen weitestgehend den bisherigen Forschungsstand zum Thema Klassenstufenwiederholung. Die erste Hypothese hinterfragt die Leistungsunterschiede zwischen Schülern mit Wiederholungserfahrung und solchen ohne vergangener Wiederholung. Es stellt sich heraus, dass Wiederholer wie erwartet schwächere Testleistungen zeigen als immer versetzte Mitschüler. Das ist jedoch vor dem Hintergrund der Definition von "Sitzenbleibern" nicht überraschend.

Die zweite Hypothese erfragt die Veränderung der Leistung im Zeitraum der Wiederholung. Die Ergebnisse lassen vermuten, dass das wiederholte Schuljahr eine Leistungssteigerung für den wiederholenden Schüler nach sich zieht. Dies ist jedoch kritisch zu betrachten, da innerhalb dieser Zeit auch eine alterstypische Weiterentwicklung stattfindet. Somit ist aus diesem Ergebnis nicht abzuleiten, dass die Leistungssteigerung auf die Wiederholung an sich zurückzuführen ist. Ermittelt wird nur, dass eine Veränderung stattgefunden hat. In den Hypothesen vier und fünf sind spezielle Aspekte auf Basis der Ergebnisse der zweiten Hypothese untersucht worden. Es zeigt sich, dass nur bedingt Unterschiede zwischen den Leistungsverbesserungen von Wiederholern der unterschiedlichen Beschulungsformen bestehen und, dass die schlechtesten Schulnoten aus dem Vorjahr im Schnitt die größte Verbesserung verzeichnen.

Mit der dritten Hypothese findet eine Annäherung an den eigentlichen Effekt der Wiederholung auf die schulische Leistungen statt. Die Ergebnisse deuten darauf hin, dass sich die Leistungen von Schülern je nach Wiederholerstatus systematisch unterscheiden, obwohl in der Vergangenheit ein ähnlicher Leistungsstand vorhanden war. Schüler mit Wiederholungserfahrung (ein Jahr zwischen der dritten und achten Klassenstufe) weisen im Schnitt ein schwächeres Testergebnis auf, als Mitschüler, die immer versetzt wurden. Dies deutet darauf hin, dass die Klassenstufenwiederholung einen negativen Einfluss, nicht nur auf das Befinden, sondern auch auf die schulische Leistung hat. Wenngleich eine abschließende Bewertung der Wirkung einer Wiederholung auch im Rahmen dieser Untersuchung nicht möglich ist, geben die vorliegenden Ergebnisse Anlass zum genauen und kritischen Hinterfragen der Klassenstufenwiederholung.

3 Rezeption und Evaluation der Tests

3.1 Befragung der Lehrkräfte nach der Dateneingabe

Lehrkräfte konnten auch dieses Jahr nach Abschluss der Dateneingabe die Kompetenztests bzgl. folgender Aspekte bewerten:

- Lehrplanadäquatheit der Aufgaben
- Nützlichkeit als Instrument zur Lernstandsdiagnostik
- Nützlichkeit für die eigene Unterrichtsentwicklung
- Nützlichkeit für die Schulentwicklung

Die vier genannten Aspekte wurden mit Hilfe einer 5-stufigen Skala erfasst, wobei 1 = volle Zustimmung und 5 = gar keine Zustimmung in Bezug auf die jeweilige Aussage bedeutet. Insgesamt wurden bei 5030 Dateneingaben Bewertungen vorgenommen.

In Tabelle 3.1 sind testspezifisch die Mittelwerte der Einschätzungen der Fachlehrkräfte zu den Kompetenztests wiedergegeben. Auf eine separate Darstellung der Ergebnisse nach Gymnasium und Regelschule wurde hier verzichtet, da sich keine bedeutsamen Unterschiede in den Einschätzungen zeigen.

Tabelle 3.1: Einschätzung der Kompetenztests im Hinblick auf Lehrplanadäquatheit sowie die Nützlichkeit für Lernstandsdiagnostik, Unterrichtsentwicklung und Schulentwicklung, Skala von 1 = volle Zustimmung bis 5 = gar keine Zustimmung.

Kompetenz-test	Lehrplan-adäquatheit	Lernstands-diagnostik	Unterrichts-entwicklung	Schulent-wicklung
DK3L	2,21	2,66	2,70	2,96
DK3H	2,32	2,59	2,71	2,91
MK3	2,40	2,71	2,76	2,99
DK6	2,28	2,72	2,92	3,12
EK6	2,44	2,88	3,04	3,22
MK6	2,21	2,68	2,86	3,06
DK8	2,22	2,85	3,07	3,32
EK8	2,49	2,94	3,20	3,42
MK8	2,31	2,71	2,95	3,16

Im Fach Deutsch werden die Kompetenztests durch die Lehrkräfte im Hinblick auf die Lernstandsdiagnostik in Klassenstufe 3 mit einem kleinen Effekt besser als in Klassenstufe 8 eingeschätzt. Der Bezug zur Unterrichtsentwicklung wird in Klassenstufe 3 ebenfalls mit einem kleinen Effekt besser bewertet als in den Klassenstufen 6 und 8. Im Hinblick auf die Schulentwicklung werden die Kompetenztests in Klassenstufe 3 (Lesen) und Klassenstufe 6 mit kleinem Effekt besser als in Klassenstufe 8 und zusätzlich in Klassenstufe 3 (Hören) mit kleinem Effekt besser als in Klassenstufe 6 und 8 eingeschätzt. Während sich im Fach Englisch keine Effekte zwischen den verschiedenen Klassenstufen zeigen, wird in Mathematik die Lehrplanadäquatheit in Klassenstufe 6 mit einem kleinen Effekt besser als in Klassenstufe 3 eingeschätzt. In Bezug auf die Unterrichtsentwicklung wird der Mathematik-Kompetenztest in Klassenstufe 3 mit kleinem Effekt besser eingeschätzt als in Klassenstufe 8. Die Kompetenztests Deutsch und Englisch konnten in Klassenstufe 8 sowohl als Online-Test als auch klassisch als Papierversion durchgeführt werden. Hinsichtlich der in der obigen Tabelle erhobenen Merkmale ergeben sich keine signifikanten Unterschiede in Bezug auf den Test-Modus.

Nützlichkeitsbewertungen im Vergleich zum Vorjahr

Vergleicht man die Mittelwerte aller diesjährigen Antworten der Lehrkräfte mit denen des letzten Jahres, so zeigen sich keine bedeutsamen Unterschiede.

Bei einer differenzierten, fach- und schulformspezifischen Betrachtung finden sich in der Regel minimale oder keine Effekte. Die größten Unterschiede werden im Folgenden dargestellt:

- DK6 an Gymnasien wird in Bezug auf Schulentwicklung mit kleinem Effekt besser als im Vorjahr bewertet.
- EK6 an Gymnasien wird im Hinblick auf die Lehrplanadäquatheit mit kleinem Effekt besser als im Vorjahr eingeschätzt.
- DK8 an Regelschulen und Gymnasien werden mit kleinem Effekt besser als im Vorjahr in Bezug auf Lehrplanadäquat, an Regelschulen zusätzlich in Bezug auf Lernstandsdiagnostik bewertet.
- MK8 Testheft 2 an Gymnasien wird bezogen auf die Lehrplanadäquatheit mit kleinem Effekt besser als im Vorjahr eingeschätzt.

3.2 Rezeption der Testergebnisse an den Schulen

3.2.1 Formen der Rückmeldungen

Im Jahr 2019 wurden für Thüringer Klassen und Schüler für die Kompetenztests insgesamt 17906 Berichte im PDF-Format geliefert. Die Ergebnismrückmeldungen erfolgen in zwei Wellen. Unmittelbar nach der Dateneingabe können die teilnehmenden Lehrkräfte die Ergebnisse der eigenen Klasse auf dem Schulportal abrufen:

- Die **tabellarische Auswertung** gibt einen Überblick über die eingegebenen Daten und enthält darüber hinaus Schülerergebnisse in wichtigen Teilbereichen. Sie kann als Excel-Datei heruntergeladen und für eigene Analysen verwendet werden.
- Der **Sofortbericht** zeigt die Ergebnisse der Lerngruppe in grafisch aufbereiteter Form. Er enthält neben den Lösungshäufigkeiten der einzelnen Aufgaben, Angaben zu wichtigen Teilbereichen sowie die Verteilung der Schülerleistung.

Etwa 6 Wochen nach Ende der Dateneingabe werden verschiedene Berichte mit Vergleichswerten veröffentlicht:

- Die **Ergebnisberichte** enthalten Vergleichswerte von Klassen mit ähnlicher Schülerschaft die zur Reflektion über die Stärken und Schwächen des eigenen Unterrichts genutzt werden sollen.
- In den **Schülerberichten** sind die Ergebnisse der einzelnen Schüler im Test insgesamt und in verschiedenen Teilbereichen im Vergleich zu den landesweiten Ergebnissen dargestellt.
- Die **Schulberichte** enthalten die Testergebnisse der gesamten Schule im Vergleich zu Schulen mit ähnlicher Schülerschaft über mehrere Jahre hinweg. Sie dienen vorrangig der Schulentwicklung.

3.2.2 Download-Quoten 2019

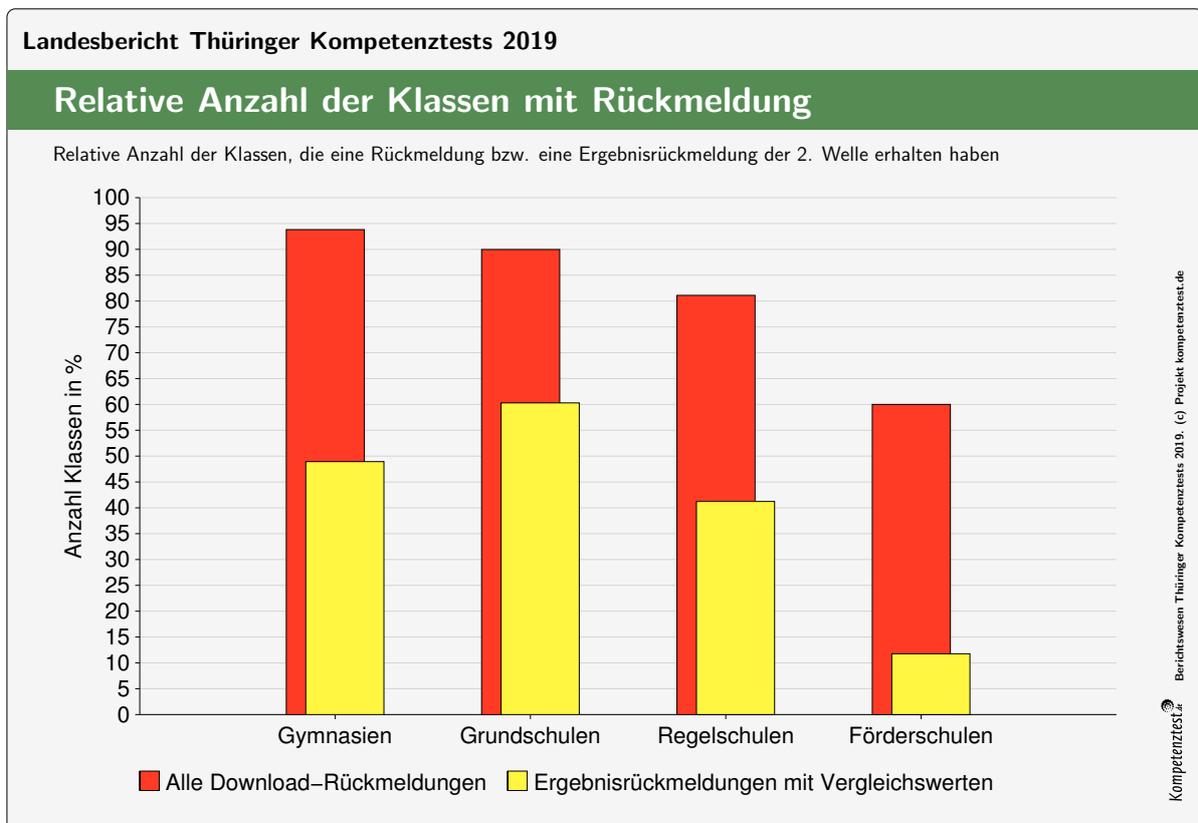


Abbildung 3.1: Relative Anzahl von Klassen mit Downloads mindestens einer Sofort-Rückmeldung vs. Ergebnissrückmeldung mit korrigierten Landesmittelwerten (Stand: 14. Oktober 2019).

In Abbildung 3.1 ist die relative Gesamtzahl der Klassen mit Downloads von Rückmeldungen, getrennt nach der Schulart, dargestellt. Die roten Balken zeigen den Anteil an Klassen, die mindestens eine der verschiedenen Rückmeldungen abgerufen haben. Die gelben Balken stellen den Anteil der Klassen dar, die in der zweiten Rückmeldewelle mindestens einen Bericht mit korrigierten Landesvergleichsdaten heruntergeladen haben.

Die Abrufquoten der Berichte sind vergleichbar mit dem Vorjahresniveau. Es ergeben sich einzelne Schwankungen im Bereich weniger Prozente. Z. B. ist die Abrufquote der Ergebnisse mit Vergleichswerten bei den Gymnasien ca. 5 % höher als im Vorjahr. Insgesamt fallen die Abrufquoten auch dieses Jahr an Förderschulen am geringsten aus, insbesondere bei den finalen

Ergebnisberichten mit Landesvergleichswerten. Da für Förderschulen die Vergleichbarkeit mit anderen Schulen aufgrund der Heterogenität der Schülerschaft oft nur eingeschränkt sinnvoll ist, ist dieser Befund auch zu erwarten.

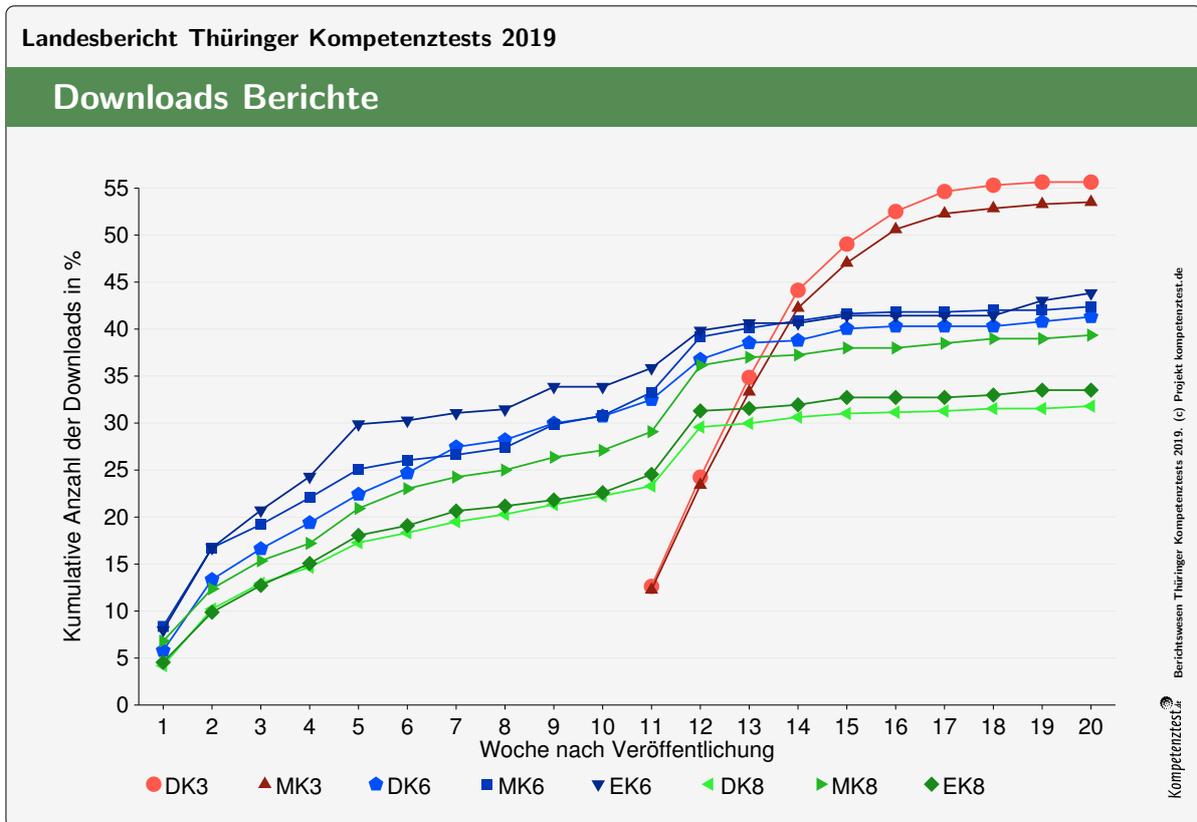


Abbildung 3.2: Downloads der Ergebnisberichte mit Vergleichsdaten der verschiedenen Tests im zeitlichen Verlauf (pro Woche nach Veröffentlichung K6/K8, Stand: 14. Oktober 2019).

Die Abrufquoten, speziell für die (Klassen-)Ergebnisberichte mit Vergleichswerten, getrennt nach Fach und Klassenstufe, zeigt Abbildung 3.2 im zeitlichen Verlauf. Die Rückmeldungen zu den Tests der Klassenstufen 6 und 8 wurden bereits am 27. Mai 2019 veröffentlicht, also zehn Wochen vor der Veröffentlichung der Berichte in der Klassenstufe 3.

Beim zeitlichen Verlauf der Downloads zeigt sich das gleiche Bild wie in den Vorjahren: Die Abrufquoten in der Klassenstufe 3 sind bereits nach wenigen Wochen am höchsten, d. h. in der Klassenstufe 3 werden prozentual mehr Berichte heruntergeladen und dies auch in einem kürzeren Zeitraum nach der Veröffentlichung. Die mit etwas Abstand geringsten Abrufquoten finden sich auch in diesem Jahr für die Berichte zu den Kompetenztests Deutsch und Englisch in der Klassenstufe 8. Im Vorjahresvergleich sind die Quoten vor der Veröffentlichung der Berichte in Klassenstufe 3 etwas heterogener als im Vorjahr.

Die Abrufzahlen der Schülerberichte liegen erneut unter denen der Klassenberichte. Wie in den letzten Jahren auch, wurde in nur ca. 50 % der Fälle auch ein Schülerbericht für das gleiche Fach heruntergeladen, wenn für eine Klasse ein Ergebnisbericht mit Vergleichswerten in diesem Fach abgerufen wurde.

Für 764 Schulen konnten in diesem Jahr Schulberichte erzeugt werden, die gemittelte Vergleiche der Klassenergebnisse mit den jeweils korrigierten Landeswerten über mehrere Jahre hinweg enthalten. Von diesen Schulen haben wie auch im letzten Jahr ca. 41 % die Berichte innerhalb von zehn Wochen nach dem letzten Veröffentlichungstermin abgerufen. Betrachtet

3.2 Rezeption der Testergebnisse an den Schulen

man die verschiedenen Schularten¹, so sind die Quoten mit 45% für Grundschulen zum Vorjahr vergleichbar, mit 43,5 % für Gymnasien moderat gestiegen und mit 32 % für Regelschulen etwas geringer als im Vorjahr.

¹ Für Förderschulen werden keine Schulberichte erzeugt.

3.3 Evaluationsbefragung zu den Kompetenztests

Jedes Jahr wird nach der Bereitstellung der Berichte eine Nutzerbefragung an 10 % zufällig ausgewählter Thüringer Schulen durchgeführt. In diesem Jahr nahmen 308 Lehrkräfte und 82 Schulleiter an der Evaluationsbefragung teil. Es werden Schulleiter und die am Kompetenztest beteiligten Lehrer um ihre Meinung zu den Kompetenztests und ihrem Umgang mit den Ergebnisberichten gebeten. Dabei werden drei zentrale Komponenten erfasst: Die Gesamtbewertung der Kompetenztests, die Auswirkungen auf die Unterrichtsentwicklung und die Auswirkungen auf die Schulentwicklung.² Jedes dieser Hauptthemen besteht aus mehreren Items, deren zusammengefasste Mittelwerte letztendlich den Skalenwert der Komponente ergeben. Diese den einzelnen Skalen zugrundeliegenden Items sind in den nachfolgenden Tabellen 3.2 bis 3.4 mit ihrem Wertebereich aufgeführt.

Tabelle 3.2: Skala Gesamtbewertung

Items	Antwortkategorien
<ul style="list-style-type: none"> • Die Schüler profitieren von der Durchführung der Thüringer Kompetenztests. • Von der Durchführung der Thüringer Kompetenztests sind wichtige Impulse für die pädagogische Diskussion in meiner Schule ausgegangen. • Die Kompetenztests leisten einen Beitrag zur langfristigen Qualitätsentwicklung an meiner Schule. • Die landesweite Durchführung der Kompetenztests nehme ich als eine Chance wahr, über den eigenen Tellerrand zu blicken. • Die Kompetenztests nehme ich als Baustein in einem Gesamtkonzept zur Schulentwicklung wahr. • Ich empfinde die Thüringer Kompetenztests insgesamt als eine sinnvolle Maßnahme. 	6-stufige Skala von [1] „stimme voll zu“ bis [6] „stimme gar nicht zu“

² Weitere Informationen zur Entwicklung der hier verwendeten Skalen können dem Landesbericht zu den Thüringer Kompetenztests 2011 entnommen werden. Dieser kann unter www.kompetenztest.de/downloads/kompetenztests/archiv heruntergeladen werden.

Tabelle 3.3: Skala Auswirkungen auf die Unterrichtsentwicklung

Items	Antwortkategorien
<ul style="list-style-type: none"> • Anpassung der Unterrichtsinhalte (Übungsschwerpunkte, Aufgaben, ...) • Anpassung der Unterrichtsgestaltung (Didaktik, ...) • Binnendifferenzierung (Abstimmung des Unterrichts auf einzelne Schüler) • Intensivere Kommunikation mit einzelnen Schülern 	6-stufige Skala von [1] „stark“ bis [6] „gar nicht“

Tabelle 3.4: Skala Auswirkungen auf die Schulentwicklung

Items	Antwortkategorien
<ul style="list-style-type: none"> • Intensivere Kommunikation mit Fachkollegen an anderen Schulen • Diskussion über Qualitätsentwicklung an meiner Schule • Entwicklung von Maßnahmen zur Qualitätsentwicklung an meiner Schule • Auf- und Ausbau von Kontakten zu externen Schulberatern (Fachberater, Schulämter, ...) • Impuls zum Einsatz innovativer Unterrichtsmethoden • Präzisierung des Bedarfs an (externer) Weiter- / Fortbildung 	6-stufige Skala von [1] „stark“ bis [6] „gar nicht“

In den nachfolgenden Abschnitten werden die Ergebnisse der diesjährigen Schulleiter- und Lehrkräftebefragung in zusammengefasster Form dargestellt. Dazu werden zunächst die Einschätzungen der Fachlehrkräfte denen der Schulleiter gegenüber gestellt und dann die Einschätzungen zwischen den Schularten. Abschließend wird betrachtet, welche Auswirkung die Durchführungsart der Kompetenztests (Papiertest vs. Online-Tests) auf das Nutzungsverhalten und die Akzeptanz der Kompetenztests hat.

An dieser Stelle sei darauf hingewiesen, dass es sich bei den Skalen zur Gesamtbewertung, Unterrichtsentwicklung und Schulentwicklung um 6-stufige Skalen handelt. Dies hat zur Folge, dass der 'neutrale' Skalenmittelwert bei diesen drei Skalen bei 3,5 liegt.

3.3.1 Vergleich der Einschätzungen von Schulleitern und Fachlehrkräften

Die Ergebnisse der Evaluationsbefragung sind in den Tabellen 3.5 und 3.6 getrennt für Schulleiter und für Fachlehrkräfte sowie für die Klassenstufen 6 und 8 und Klasse 3 dargestellt. Auf allen drei Dimensionen bewerten die Schulleiter die Kompetenztests besser als die Fachlehrkräfte. Dieser Unterschied in der Bewertung zeigte sich auch in vergangenen Jahren. Statistisch signifikant sind die unterschiedlichen Bewertungen bezüglich der Schulentwicklung. Der Effekt ist in den Klassenstufen 6 und 8 klein und in Klassenstufe 3 mittelgroß.

Tabelle 3.5: Vergleich der Bewertungen von Schulleitern und Lehrenden in Klassenstufe 6 und 8

	Schulleiter			Lehrende		
	MW	SD	N	MW	SD	N
Gesamtbewertung	3,85	1,23	43	4,20	1,26	227
Unterrichtsentwicklung	3,62	0,95	43	3,94	1,27	229
Schulentwicklung	4,44	0,87	43	4,89	1,00	227

Tabelle 3.6: Vergleich der Bewertungen von Schulleitern und Lehrenden in Klassenstufe 3

	Schulleiter			Lehrende		
	MW	SD	N	MW	SD	N
Gesamtbewertung	3,25	1,25	38	3,63	1,36	72
Unterrichtsentwicklung	2,99	1,17	39	3,42	1,18	73
Schulentwicklung	3,77	1,06	39	4,54	1,13	72

3.3.2 Vergleich der Einschätzungen nach Schulform

In Tabelle 3.7 sind die Skalenmittelwerte der **Fachlehrkräfte** schulformspezifisch dargestellt. Grundschullehrkräfte bewerten die Kompetenztest in allen drei Skalen besser als Regelschullehrkräfte und diese wiederum besser als Gymnasiallehrkräfte. Die unterschiedlichen Bewertungen der Kompetenztests durch Grundschul- und Gymnasiallehrkräfte sind mit mittelgroßen Effekten (in Bezug auf Unterrichts- und Schulentwicklung) bzw. kleinem Effekt (in Bezug auf die Gesamtbewertung) statistisch signifikant. Im Vergleich zwischen Grundschul- und Regelschullehrkräften unterscheidet sich nur die Gesamtbewertung signifikant mit einem kleinen Effekt. Zwischen Regelschul- und Gymnasiallehrkräften unterscheiden sich die Bewertungen bezüglich der Unterrichts- und Schulentwicklung jeweils mit kleinem Effekt signifikant.

Tabelle 3.7: Vergleich der Bewertungen der Fachlehrkräfte von Grundschulen, Regelschulen und Gymnasien

	Grundschulen			Regelschulen			Gymnasien		
	MW	SD	N	MW	SD	N	MW	SD	N
Gesamtbewertung	3,63	1,36	72	4,18	1,30	117	4,22	1,22	110
Unterrichtsentwicklung	3,42	1,18	73	3,74	1,29	118	4,16	1,22	111
Schulentwicklung	4,54	1,13	73	4,71	1,04	117	5,08	0,93	110

Die Kompetenztests werden von **Schulleitungen** (Tabelle 3.8) weiterführender Schulen homogen bewertet. Zwischen den Bewertungen von Grundschulschulleitungen und Regelschulleitungen zeigen sich hingegen auf allen drei betrachteten Skalen signifikante Unterschiede. Die Schulleitungen von Grundschulen schätzen die Kompetenztests jeweils mit mittleren Effekten besser ein, als Schulleitungen von Regelschulen. Ähnliche Ergebnisse ergeben sich beim Vergleich zwischen Schulleitungen von Grundschulen und Gymnasien. Hier ist die Einschätzung der Grundschulschulleitungen in Bezug auf die Unterrichts- und Schulentwicklung signifikant (jeweils mit mittelgroßem Effekt) besser als bei Schulleitungen von Gymnasien.

Tabelle 3.8: Vergleich der Bewertungen der Schulleiter von Grund- und Regelschulen sowie Gymnasien

	Grundschulen			Regelschulen			Gymnasien		
	MW	SD	N	MW	SD	N	MW	SD	N
Gesamtbewertung	3,25	1,25	38	3,88	1,23	30	3,78	1,27	13
Unterrichtsentwicklung	2,99	1,17	39	3,59	1,01	30	3,69	0,84	13
Schulentwicklung	3,77	1,06	39	4,41	0,86	30	4,50	0,92	13

3.3.3 Vergleich der Einschätzungen nach Teilnahme am Online-Test versus Papiervariante

Der folgende Abschnitt beschäftigt sich mit den Einschätzungen der Lehrkräfte und der Schulleiter in Abhängigkeit davon, ob sie an der Online- oder Papiervariante des Kompetenztests teilnahmen.

Generell schätzten Lehrkräfte, die die Online-Version durchführten, den Kompetenztest besser ein. Statistisch bedeutsame Unterschiede finden sich hier in Bezug auf die Gesamtbewertung (kleiner Effekt) (siehe 3.9).

Tabelle 3.9: Einschätzung der Fachlehrkräfte nach Teilnahme am Online-Test versus Papiervariante

	Online-Test			Papiervariante		
	MW	SD	N	MW	SD	N
Gesamtbewertung	3,86	1,31	62	4,35	1,20	161
Unterrichtsentwicklung	3,76	1,22	63	4,02	1,30	162
Schulentwicklung	4,77	0,97	62	4,94	1,02	161

Die Einschätzung der Schulleiter unterscheidet sich für die drei betrachteten Skalen nicht signifikant in Bezug auf die Durchführung der Kompetenztests als online- oder Papiervariante (siehe 3.10).

Tabelle 3.10: Einschätzungen der Schulleiter nach Teilnahme am Online-Test versus Papiervariante

	Online-Test			Papiervariante		
	MW	SD	N	MW	SD	N
Gesamtbewertung	3,52	1,15	23	4,22	1,23	20
Unterrichtsentwicklung	3,54	0,84	23	3,71	1,08	20
Schulentwicklung	4,26	0,83	23	4,64	0,89	20

4 Literatur

Bärsch, W. (1987). Die Familie muss es verarbeiten. Über die Auswirkungen des Sitzenbleibens auf Eltern und Kinder. Westermanns Pädagogische Beiträge, 4, 28-26.

Cohen, J. (1988): Statistical Power Analysis for the Behavioral Sciences. Hillsdale: Erlbaum.

Hattie, J. (2015). Lernen sichtbar machen. Schneider Verlag GmbH.

Holmes, C. T. (1989). Grade level retention effects: A meta-analysis of research studies. Flunking grades: Research and policies on retention, 16, 33.

Ingenkamp, K. (1976). Zur Problematik der Jahrgangsklasse. Eine empirische Untersuchung. Belz.

Jeske, W. (1981). Versetzen oder Sitzenbleiben - Das Problem der Leistungsversager. Handbuch Schule und Unterricht, 1, 361-377.

Kemmler, L. (1976). Schulerfolg und Schulversagen: Eine Längsschnittuntersuchung vom ersten bis zum fünften Schulbesuchsjahr. Verlag für Psychologie, Hogrefe.

Kemnade, I. (1989). Schullaufbahnen und nDurchlässigkeit in der Sekundarstufe I: empirische Untersuchung von Schülerkarrieren in der Stadt Bremen. Lang.

Nachtigall, C. (Hrsg.) (2018): Landesbericht Thüringer Kompetenztests 2016.
www.kompetenztest.de/downloads

Nachtigall, C. & Wirtz, M. (2008): Wahrscheinlichkeitsrechnung und Inferenzstatistik. Weinheim: Juventa.

Palowski, M., Boller, S. Müller, M. (2016). Der Diskurs des Versagens. Springer Fachmedien Wiesbaden.

Rust, J. O. Wallace, K. A. (1993). Effects of grade level retention for four years. Journal of Instructional Psychology, 20 (2), 162.

Statistisches Bundesamt. (2018). Wiederholer nach Geschlecht und Bundesland.
www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Schulen/Tabellen/AllgemeinBildendeWiederholerBundeslaender

Yamamoto, K. (1980). Children under stress: The causes and cures. Family Weekly, Ogden Standard Examiner, 6-8.

5 Anhang

Statistische Auswertungen und deren Interpretation

Mittelwerte werden berechnet, indem die Summe aller Werte durch die Anzahl aller Werte geteilt wird. Ergänzend zu diesem Wert wird die **Standardabweichung** angegeben. Die Standardabweichung ist ein Maß der Streuung und weist darauf hin, wie stark die Werte im Durchschnitt vom Mittelwert abweichen. Sie gibt also Auskunft über die Unterschiedlichkeit einer Stichprobe. Um die Ergebnisse unterschiedlicher Gruppen mit unterschiedlichen Mittelwerten vergleichen zu können, kann man die **Effektstärke** d^1 berechnen. Die Effektstärke d gibt an, ob und wie stark sich die Gruppen unterscheiden. Dabei gilt nach Cohen (1988) eine Effektstärke ab $d = 0,2$ als geringer, ab $d = 0,5$ als mittelstarker und ab $d = 0,8$ als starker Effekt. Effektstärken unter $d = 0,2$ gelten als unbedeutend.

Beispiel: Gruppenunterschiede mit Effektstärke angeben. Aus den Mittelwerten und Standardabweichungen von zwei Schülergruppen (z. B. Mädchen und Jungen) errechnet sich bei einem Test eine Effektstärke $d = 0,28$. Diese zeigt zunächst an, dass sich die beiden Gruppen in ihren Leistungen voneinander unterscheiden. Es liegt jedoch nur ein kleiner Effekt vor. Das Vorzeichen gibt an, dass die erste Gruppe (hier: die Mädchen) im Durchschnitt bessere Leistungen erzielt. Ein negatives Vorzeichen würde anzeigen, dass die zweite Gruppe (hier: die Jungen) besser abschnidet.

Korrelationen geben an, inwieweit zwischen zwei Merkmalen ein Zusammenhang besteht und wie ausgeprägt dieser ist. Die Stärke des Zusammenhangs definiert sich durch einen Wert zwischen -1 und 1 . Dabei ist nach Nachtigall und Wirtz (2008) der Betrag einer Korrelation bis $0,50$ als gering, bis $0,70$ als mittel und ab $0,70$ als hoch einzustufen. Beispielsweise deutet die bei früheren Kompetenztests gefundene Korrelation zwischen den mathematischen und den sprachlichen Kompetenzen von $r \approx 0,62$ auf eine mittlere Stärke des Zusammenhanges hin.

Effektstärke, Korrelationen und Kausalität: Der Begriff des „Effektes“ legt nahe, dass es dabei um Kausalität geht. Dies muss allerdings keineswegs der Fall sein. Untersucht man z. B. die Leistungen von Gymnasiasten und Regelschülern beim Kompetenztest, Klassenstufe 6, so schneiden die Gymnasiasten im Schnitt deutlich besser ab. Es findet sich statistisch ein bedeutsamer Effekt der Schulart. Hierbei handelt es sich jedoch vor allem um einen Selektionseffekt, da die leistungsstärkeren Schüler tendenziell eher auf das Gymnasium gehen und die weniger leistungsstarken eher auf die Regelschule. Daher liegen bereits zu Beginn des 5. Schuljahres Leistungsunterschiede zwischen den Schularten vor, noch bevor die Schulart selbst irgendeinen kausalen Effekt haben kann.

Bei allen in diesem Bericht analysierten Effekten handelt es sich zunächst nur um Beschreibungen der Größe von Gruppenunterschieden. Diese Unterschiede können, müssen aber keineswegs kausal sein. Bei der Prüfung, ob ein Effekt wirklich kausal ist, muss immer überlegt werden, ob er nicht durch andere Einflussfaktoren zustande kommen könnte. Nur wenn dies ausgeschlossen werden kann, ist eine kausale Interpretation gerechtfertigt. Ähnlich ist es bei Korrelationen. Auch diese werden häufig vorschnell als kausale Zusammenhänge interpretiert. Auch hier ist zu beachten, dass Korrelationen zunächst nur das gemeinsame Auftreten von

¹ Konkret wird in diesem Bericht für die Berechnung der Effektstärken der Kennwert Hedges g verwendet.

Merkmale beschreiben, ohne dass ein Merkmal das andere Merkmal ursächlich beeinflussen muss. Ein illustratives Beispiel dafür ist der regional nachweisbare Zusammenhang zwischen der Geburtenrate und der Anzahl der Störche².

² vgl. http://www.zeit.de/2006/25/Stimmt-s_P-25_xml

Impressum:

Dieser Bericht wurde vom Projekt *kompetenztest.de* der Friedrich-Schiller-Universität Jena in Kooperation mit dem Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien (Thillm) im Auftrag des Thüringer Ministeriums für Bildung, Jugend und Sport (TMBJS) erstellt.

Dr. Christof Nachtigall
Dipl.-Psych. Isabell Storbeck

Bei Fragen wenden Sie sich bitte an: kontakt@kompetenztest.de

Erstellungsdatum: 11. November 2019