

Datengestützte Entwicklung und Steuerung in Schulen und frühkindlicher Bildung

Gutachten der Ständigen Wissenschaftlichen Kommission der
Kultusministerkonferenz

Sperrfrist: 26.03.2026, 12:45 Uhr





Über die Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (SWK)

Die Ständige Wissenschaftliche Kommission (SWK) ist ein unabhängiges wissenschaftliches Beratungsgremium der Kultusministerkonferenz. Ihr gehören 16 Bildungsforscher:innen aus unterschiedlichen Disziplinen an. Die SWK berät die Länder zu bildungspolitischen Fragen. Sie identifiziert bestehende Herausforderungen und gibt evidenzbasierte Empfehlungen für deren Lösung. Dabei nimmt die Kommission eine interdisziplinäre, längerfristige und systemische Perspektive ein. Die SWK bindet externe Sachverständige in ihre Arbeit ein und hört Vertreter:innen aus Politik, Verwaltung, Bildungspraxis und Zivilgesellschaft an. Eine Geschäftsstelle unterstützt die Kommission bei ihrer Arbeit.



Inhaltsverzeichnis

Einleitung	6
Kapitelübergreifende Empfehlungslinien	11
1. Datengestützte Unterrichtsentwicklung und diagnosegeleitete Lernförderung	14
1.1 Situationsanalyse	16
1.1.1 Zentrale Lernstandserhebungen	16
1.1.2 Lernprozessbegleitende Diagnoseinstrumente	21
1.1.3 Learning Analytics	23
1.1.4 Daten zu überfachlichen Konstrukten	24
1.1.5 Instrumente zur Erfassung der Unterrichtsqualität	26
1.1.6 Darstellung der Daten in Form von Dashboards	27
1.2 Konzepte und Maßnahmen	27
1.2.1 Effekte und Bedingungen der Nutzung zentraler Lernstandserhebungen für Unterrichtsentwicklung	27
1.2.2 Effekte und Bedingungen der Nutzung diagnosebasierter Förderung mit lernprozessbegleitenden Diagnoseinstrumenten	31
1.2.3 Effekte und Bedingungen der Nutzung von Learning Analytics für Lernförderung	33
1.2.4 Effekte und Bedingungen der Nutzung von Daten zu überfachlichen Konstrukten	35
1.2.5 Effekte und Bedingungen der Nutzung von Instrumenten zur Erfassung der Unterrichtsqualität	38
1.3 Schlussfolgerungen und Empfehlungen	40
2. Datennutzung für Schulentwicklung im Rahmen erweiterter Gestaltungsspielräume	45
2.1 Situationsanalyse	46
2.1.1 Daten für Schulentwicklung aus interner und externer Evaluation sowie Lernstandserhebungen	46
2.1.2 Nutzung von Daten für Ziel- und Leistungsvereinbarungen mit der Schulaufsicht	50
2.1.3 Strukturen, Qualifikation und Unterstützungsmaßnahmen für datengestützte Schulentwicklung	50
2.2 Konzepte und Maßnahmen	52
2.2.1 Daten für strategische Schulentwicklung	53
2.2.2 Schulentwicklungspläne als Instrumente strategischer Schulentwicklung	55
2.2.3 Selbstevaluation als Instrument strategischer Schulentwicklung	57
2.2.4 Unterstützungsstrukturen für datengestützte Schulentwicklung	61
2.3 Schlussfolgerungen und Empfehlungen	62
3. Daten für Controlling und Unterstützung durch die Schulaufsicht	65
3.1 Situationsanalyse	66
3.1.1 Erfassung, Aufbereitung und Nutzung von Daten für schulaufsichtliches Handeln	66



3.1.2	Externe Evaluation und schulaufsichtliches Handeln	69
3.1.3	Schul- sowie Fachberatungen und schulaufsichtliches Handeln	71
3.2	Konzepte und Maßnahmen	72
3.2.1	Konzepte und Effekte datengestützter Rechenschaftslegung	73
3.2.2	Datenfeedback, Zielvereinbarungen und Unterstützung datengestützter Entwicklung	76
3.2.3	Qualifizierung der Schulaufsicht für datengestütztes Controlling und Unterstützung	79
3.3	Schlussfolgerungen und Empfehlungen	79
4.	Daten für Erziehungs- und Bildungspartnerschaften zwischen Eltern, Schüler:innen, Lehrkräften und pädagogischen Fachkräften	82
4.1	Situationsanalyse	83
4.1.1	Erziehungs- und Bildungspartnerschaften zwischen Schule, Eltern und außerschulischen Partnern	83
4.1.2	Nutzung und Bereitstellung von Daten für die Zusammenarbeit mit Eltern	86
4.1.3	Lernentwicklungs- und Förderplangespräche als Struktur für datengestützte Erziehungs- und Bildungspartnerschaften	92
4.2	Konzepte und Maßnahmen	96
4.2.1	Übergreifende Strategien für datengestützte Partnerschaften: internationale Beispiele ...	96
4.2.2	Umgang mit Herausforderungen datengestützter Partnerschaft	98
4.2.3	Gestaltung von Lernentwicklungsgesprächen	102
4.3	Schlussfolgerungen und Empfehlungen	103
5.	Daten für Systemmonitoring und -steuerung durch die Politik	106
5.1	Rückblick und Situationsanalyse	107
5.1.1	Die Gesamtstrategie der KMK zum Bildungsmonitoring	107
5.1.2	Weitere internationale Studien jenseits der Gesamtstrategie	112
5.1.3	Welche Informationen liefern die Large-Scale Assessments?	113
5.1.4	Welche Informationen liefert der Nationale Bildungsbericht?	115
5.1.5	Adressaten von Large-Scale Assessments und Nationalem Bildungsbericht	117
5.1.6	Berichtsformate der Large-Scale Assessments und des Nationalen Bildungsberichts	118
5.1.7	Federführung für Large-Scale Assessments: Institutionalisierung vs. wissenschaftliche Konsortien	119
5.1.8	Kosten der Large-Scale Assessments und der nationalen Bildungsberichterstattung	120
5.1.9	Der Blick ins Ausland	120
5.1.10	Reformen und Programme im schulischen Bildungssystem infolge des Bildungsmonitorings	122
5.1.11	Impulse für die Bildungsforschung	124
5.1.12	Die Rolle einer Bildungsverlaufsstatistik im Bildungsmonitoring	124
5.1.13	Potenziale künstlicher Intelligenz für Large-Scale Assessments	125
5.2	Schlussfolgerungen und Empfehlungen	126



6. Datengestützte Qualitätsentwicklung in der Frühen Bildung	129
6.1 Situationsanalyse.....	130
6.1.1 Monitoring und Controlling.....	130
6.1.2 Organisationsentwicklung und Ressourcenzuweisung	135
6.1.3 Individuelle Planung von Bildungs- und Förderangeboten sowie Übergang in die Grundschule	140
6.1.4 Erziehungs- und Bildungspartnerschaften.....	141
6.2 Konzepte und Maßnahmen	142
6.2.1 Weiterentwicklung des Systemmonitorings.....	142
6.2.2 Datengestützte Steuerung und organisationale Qualitätsentwicklung	145
6.2.3 Bildungs- und Förderangebote auf Basis von Entwicklungsstandsdaten und Übergang in die Grundschule	147
6.2.4 Datengestützte Erziehungs- und Bildungspartnerschaften	147
6.3 Schlussfolgerungen und Empfehlungen.....	149
Hinweise zur Umsetzung der Empfehlungen.....	153
Literatur.....	155
Anhänge	197
Anhang 1 Zum methodischen Vorgehen	197
Anhang 2 Weitere in den Ländern eingesetzte diagnostische bzw. formative Instrumente	198
Anhang 3 Übersicht über die in den Ländern eingesetzten Verfahren zur Evaluation von Unterricht im Rahmen interner Evaluation	200
Anhang 4 Externe Schulevaluationen der Länder	205
Anhang 5 Ziel- und Leistungsvereinbarungen in der Schulaufsicht	207
Anhang 6 Fachberatungen der Länder.....	210
Anhang 7 Mögliche Nutzung von Lernentwicklungsgesprächen als Ersatz und/oder in Ergänzung zum Halbjahreszeugnis	212
Abbildungsverzeichnis	215
Mitglieder der SWK.....	216
Mitwirkende.....	218
Impressum	220



Einleitung

Vor knapp 20 Jahren hat die Kultusministerkonferenz (KMK) ihre erste Version einer Gesamtstrategie zum Bildungsmonitoring vorgelegt, die sie selbst in der überarbeiteten Auflage aus dem Jahr 2015 als „Ausdruck einer neuen Steuerungsphilosophie, die sich an den Ergebnissen von Bildungsprozessen ausrichtet“ bezeichnet (KMK, 2015, S. 3). Diese Gesamtstrategie wurde zu Recht als „historische Leistung“ (Maritzen, 2014, S. 400) beschrieben, weil eine Verständigung auf erwartbare Ergebnisse schulischer Bildung in Form von Bildungsstandards gelang (Klieme et al., 2003; Stanat et al., 2025) und die Überprüfung der von Schüler:innen erreichten Kompetenzen durch Kompetenztests im Rahmen nationaler und internationaler Schulleistungsstudien sowie Vergleichsarbeiten verabredet wurde. Dazu hatte die KMK bereits im Jahr 2004 das Institut zur Qualitätsentwicklung im Bildungswesen (IOB) eingerichtet. Neben der Herstellung von Transparenz über zentrale Kompetenzziele, sind die Verbindung der Standardsicherung mit einer kontinuierlichen Qualitätsentwicklung auf allen Ebenen sowie die Verbesserung der länderübergreifenden Zusammenarbeit zentrale Ziele der Strategie.

Nahezu gleichzeitig erfolgte eine „Erweiterung schulischer Selbstständigkeit“ (Avenarius, 2007, S. 873), die sich in den Schulgesetzen aller deutschen Länder niederschlug. Diese eröffneten den Schulen neue curriculare, personelle und finanzielle Gestaltungsspielräume und stärkten Schulleitungen durch die teilweise Übertragung von Dienstvorgesetztenfunktionen. Wenngleich die Entscheidungsspielräume deutscher Schulen hinter denen anderer Länder der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) deutlich zurückbleiben (Avenarius et al., 2003; Brauckmann, 2012), erscheint die Charakterisierung der Reform als „substantielle Neuerung“ (Avenarius et al., 2003, S. 158), so die Autoren des ersten deutschen *Bildungsberichts*, durchaus angemessen. Diese Reform ging mit einer beträchtlichen Veränderung der Rollen und Funktionen unterschiedlicher Akteur:innen einher: Während Schulen und Lehrkräfte mehr Entscheidungskompetenzen erhielten, sollten sich „Schulpolitik und Schulaufsicht ... auf die Zielsetzung für die Schulen, die Gewährleistung der Handlungsvoraussetzungen eigenverantwortlicher Schulen unter Einschluss der Personal- und Finanzausstattung, die Beratung der Schulen sowie die Evaluation der Leistungen und Kontrolle der Zielerreichung beschränken“ (I. Richter, 1994, S. 5–7). „Vertrauen in die professionelle Selbstorganisationsfähigkeit der Schulen“ (Deutscher Bundestag – Wissenschaftliche Dienste, 2020, S. 10) kann aufgrund der staatlichen Verantwortung für ein qualitativ hochwertiges und zuverlässiges Angebot allgemeiner Bildung allerdings nicht bedingungslos gewährt werden. Für die Autoren des ersten Bildungsberichts war deshalb unstrittig, dass die Erweiterung schulischer Gestaltungsspielräume mit „Rechenschaftslegung und begleitenden Maßnahmen zur Qualitätssicherung einhergehen muss“ (Avenarius et al., 2003, S. 163).

Daten haben entsprechend im neuen Steuerungsmodell mehrere Funktionen:

- Zunächst sind Daten in den Händen der Lehrkräfte ein unverzichtbares Instrument der diagnosegeleiteten Lernförderung und Unterrichtsentwicklung. Eine wichtige Funktion für die Unterrichtsentwicklung haben neben Abschlussprüfungen Vergleichsarbeiten bzw. andere zentrale Lernstandserhebungen. Diese informieren Lehrkräfte darüber, inwiefern die in den Bildungsstandards definierten Kompetenzen erreicht wurden, und zeigen ggf. die Notwendigkeit einer Veränderung



der Unterrichtspraxis an (Hawlitschek et al., 2025). Zentrale Lernstandserhebungen eignen sich zudem dazu, Lehrkräfte über die Lernausgangslagen der Schüler:innen zu Beginn der Primarstufe und der Sekundarstufe zu informieren und die Wirksamkeit getroffener Maßnahmen zu überprüfen. Eine lernprozessbegleitende Diagnostik ermöglicht darüber hinausgehend eine enge Verschränkung von Diagnose und Förderung in kürzeren Abständen und fachlich größerer Ausdifferenzierung. Neben Schülerleistungen geben Daten aus Unterrichtsbefragungen von Schüler:innen oder aus Unterrichtsbeobachtungen von Kolleg:innen Aufschluss über die Qualität von Unterrichtsprozessen und identifizieren Ansatzpunkte für deren Weiterentwicklung.

- Auf der Schulebene ist die Erfassung von Daten eng verknüpft mit dem Schulprogramm (Bonsen & Frey, 2024). Schulprogrammarbeit erfolgt in enger Verbindung mit interner Evaluation; diese dient einerseits der Feststellung von Entwicklungsbedarf bei Curricula, klassenübergreifenden fachdidaktischen und pädagogischen Maßnahmen, Fortbildungsbedarfen oder dem Schulmanagement. Andererseits dient interne Evaluation der Überprüfung der im Rahmen erweiterter Gestaltungsspielräume implementierten Maßnahmen. Für eine kontinuierliche Schulentwicklung sollen Schulen auch zentrale Lernstandserhebungen und – wo vorhanden – Rückmeldungen externer Evaluation für die Schulentwicklung systematisch nutzen.
- Daten, insbesondere zur Ergebnisqualität, haben auch die Funktion des Controllings bzw. der Rechenschaftslegung. Sie informieren Schulaufsichten über die Arbeit der Schulen und bilden im Idealfall eine Grundlage für Zielvereinbarungen. Auch Prozessdaten aus der externen Evaluation werden teilweise für Zwecke der Rechenschaftslegung genutzt. Die Länder unterscheiden sich allerdings erheblich, was die Nutzung von Daten aus Vergleichsarbeiten sowie externer Evaluation für Zwecke der Rechenschaftslegung betrifft (Tarkian, Lankes & Thiel, 2019; Tarkian, Maritzen et al., 2019).
- Insbesondere Daten aus Vergleichsarbeiten und anderen Lernstandserhebungen dienen in einigen Ländern auch der Information der Eltern bzw. Schüler:innen bezüglich individueller Lernstände und -verläufe sowie der Förder- und Entwicklungsplanung für Schüler:innen mit sonderpädagogischem Förderbedarf (Tarkian, Maritzen et al., 2019). Internationale Ansätze datengestützter Elternarbeit begreifen die Datenrückmeldungen an Eltern als Grundlage für sogenannte Bildungspartnerschaften (Prabhakar et al., 2023). Häufig werden weitere pädagogische Fachkräfte in Bildungspartnerschaften einbezogen. In Deutschland sind Instrumente und Verfahren datengestützter Bildungspartnerschaften bislang noch wenig etabliert.
- Auf der Systemebene haben insbesondere Daten aus nationalen und internationalen Schulleistungsstudien die Funktion, Transparenz über die Leistungen und Leistungsfähigkeit des deutschen Schulsystems herzustellen und Ansatzpunkte für Qualitätsentwicklung zu identifizieren. Schulpolitik auf regionaler-, Länder- oder Bundesebene benötigt für Systemsteuerung und -entwicklung aggregierte Daten, die einen schulübergreifenden Handlungsbedarf anzeigen. Dieser betrifft beispielsweise die Definition von Rahmencurricula und -vorgaben, die (Re-)Allokation von Ressourcen oder die forschungsbasierte Entwicklung und Implementation von Programmen.

Diese unterschiedlichen Perspektiven und Funktionen sowie die Aktivitäten unterschiedlicher Akteure müssen ineinandergreifen, wenn von einer systemischen Strategie gesprochen werden soll. Datenerfassung, Datenrückmeldung und Datennutzung müssen im Mehrebenensystem der Bildung abgestimmt sein.



Bildungsstandards und die darauf beruhenden Kompetenzmodelle spielen für das Alignment über die verschiedenen Ebenen hinweg eine zentrale Rolle (Henschel & Stanat, 2019), wenngleich sie nicht den gesamten Bildungsauftrag der Schule abbilden.

Neben den oben genannten Funktionen spielen Daten auch eine zunehmend wichtigere Rolle bei der bedarfsorientierten Ressourcensteuerung. In Deutschland wurden im Rahmen des *Startchancen-Programms* in allen Ländern Sozialindizes entwickelt (Sendzik et al., 2025), auf deren Grundlage Schulen in herausfordernden Lagen zusätzliche Ressourcen durch Schulbehörden und Kommunen zugeteilt werden. Im vorliegenden Gutachten wird dieser Aspekt datengestützter Steuerung nur insofern berücksichtigt, als Sozialindizes auch für die Adjustierung von Ergebnissen standardbasierter Kompetenztests genutzt werden können. Was die bedarfsorientierte Ressourcenzuweisung betrifft, verweisen wir auf die entsprechenden Empfehlungen des SWK-Gutachtens zur Grundschule (2022a).

Die Implementation der Gesamtstrategie war von Beginn an durch intensive Debatten begleitet. Ergänzungsbedarf bzw. ein Bedarf an Nachjustierung wurde immer wieder artikuliert. Insbesondere drei Aspekte werden hinsichtlich einer kohärenten Strategie der Datennutzung im Mehrebenensystem gegenwärtig diskutiert:

- Lernstandserhebungen genießen zwar inzwischen bei vielen Lehrkräften Akzeptanz, insbesondere hinsichtlich der Identifikation eines Bedarfs an Förderung und Unterrichtsentwicklung vor Übergangsschwellen (Wurster, 2022). Allerdings wurde die Beschränkung der Datenerhebung bei Vergleichsarbeiten auf jeweils einen Messzeitpunkt in der Primar- und der Sekundarstufe bereits früh kritisch kommentiert (U. Maier, 2010). Einige Länder haben deshalb zusätzliche Lernstandserhebungen eingeführt. Die Verabredung der ländergemeinschaftlichen Pilotierung der Lernausgangslagen- und Lernentwicklungsuntersuchung *StarS (Stark in die Grundschule starten)* am Anfang der Grundschulzeit stellt einen großen Schritt dar. Kontrovers wird diskutiert, ob eine darüber hinausgehende längsschnittliche Erfassung von Schülerleistungen, wie sie bislang nur in Hamburg implementiert wurde, notwendig ist. Längsschnittliche Daten, die auch in die Bildungsberichterstattung eingehen, liegen im *Nationalen Bildungspanel* vor. Grundsätzlich ist die Einführung eines Bildungsverlaufsregisters bzw. einer Bildungsverlaufsstatistik, wie sie in anderen Ländern bereits seit Langem implementiert ist, erforderlich. Über zentrale Lernstandserhebungen hinaus benötigen Lehrkräfte für eine adaptive Unterrichtsentwicklung sowie Lernförderung, die der steigenden Heterogenität von Lernausgangslagen sowie spezifischen Förderbedarfen im inklusiven Unterricht Rechnung trägt, thematisch ausdifferenzierende, lernprozessbegleitende Diagnosen, die Lernverläufe über längere Zeiträume (sechs bis zwölf Monate) abbilden und die möglichst unmittelbar verknüpft sind mit Materialien zur Förderung und Unterrichtsentwicklung (Black & Wiliam, 2009; Souvignier & Hasselhorn, 2018). Learning-Analytics und andere KI-Anwendungen eröffnen neue Potenziale, erzeugen aber auch Professionalisierungsbedarfe und werfen neue datenschutzrechtliche Fragen auf.
- Lehrkräfte und Schulleitungen verfügen inzwischen über zahlreiche Daten. Häufig bleibt aber nicht nur unklar, welche Daten von besonderer Relevanz und wie diese zu interpretieren sind, sondern auch, welche Maßnahmen aus der Interpretation der Daten abgeleitet werden können (Wurster, 2022). Alle Länder haben zwar Unterstützungssysteme in Form von Schulberatungen eingeführt



(Dedering et al., 2022), bei diesen handelt es sich allerdings in der Regel nicht um Fachberatungen, sondern um Prozess- bzw. Organisationsberater:innen. Aus diesem Grund erfolgt im Anschluss an eine Evaluation häufig der Verweis auf schulinterne professionelle Lerngemeinschaften, die aus den Daten Maßnahmen eigenständig ableiten sollen (Hofbauer & Bernhard, 2023). Insbesondere bei Fachkollegien, die zu einem großen Teil aus Quereinsteigenden oder fachfremd unterrichtenden Lehrkräften bestehen, besteht das Risiko, dass sie durch den „hohen Grad an Selbstbestimmung“ und „fehlende Impulse von außen“ (Lipowsky & Rzejak, 2012, S. 243) überfordert sind. International etablierte und erfolgreich evaluierte Expertencoachings wie *fachdidaktisches Unterrichtacoaching* (L. West & Staub, 2003) oder *MyTeachingPartner* (Allen et al., 2011), die Lehrkräfte bei der Unterrichtsentwicklung unterstützen, sind im Unterschied zu Schulberatungen in Deutschland bislang nur wenig etabliert.

- Die Länder unterscheiden sich deutlich dahingehend, welche Daten aus Vergleichsarbeiten und externer Evaluation welchen Akteur:innen zur Verfügung gestellt werden (Tarkian, Lankes & Thiel, 2019; Tarkian, Maritzen et al., 2019). Während einige Länder die Daten aus Vergleichsarbeiten ausschließlich an Lehrkräfte rückmelden, nutzen andere die Daten auch für Controlling oder Monitoringzwecke. Im Unterschied zu der von Earl und Katz (2006, S. 12) vertretenen Auffassung, dass Rechenschaftslegung ohne Entwicklung leere Rhetorik und umgekehrt Entwicklung ohne Rechenschaftslegung orientierungsloser Aktivismus sei, ist in Deutschland strittig, ob die Entwicklungsfunktion mit der Funktion der Rechenschaftslegung vereinbar ist. Ähnliches gilt für die externe Evaluation. In einigen Ländern ist sie Teil der Schulaufsicht, in anderen Ländern wird sie inzwischen nur noch anlassbezogen durchgeführt und dort vor allem als Entwicklungsinstrument betrachtet, das eng mit Beratung verknüpft sein sollte. Die Fragen, wie sich Entwicklung zu Rechenschaftslegung im Allgemeinen und wie sich externe Evaluation zur Schulaufsicht im Besonderen verhält, werden angesichts der inzwischen von einigen Ländern eingeführten Datenblätter, die wiederum eine Grundlage für Zielvereinbarungen darstellen, noch einmal verschärft.

Das vorliegende Gutachten nimmt diese Diskussionen auf und bearbeitet mit Bezug auf die eingangs identifizierten Funktionen von Daten für unterschiedliche Akteur:innen die folgenden Fragen:

- Welche Rolle spielen Daten für die diagnosegeleitete Lernförderung von Schüler:innen und für Unterrichtsentwicklung durch Lehrkräfte?
- Wie können Daten Schulentwicklungsprozesse durch Schulleitungen und Kollegien, insbesondere in Verbindung mit Schulentwicklungsplänen, wirksam unterstützen?
- Welche Bedeutung haben Daten für Controlling und Unterstützung der Schulen durch die Schulaufsicht?
- Welchen Beitrag können Daten für den Auf- und Ausbau von Bildungspartnerschaften zwischen Lehrkräften, Eltern, Schüler:innen und weiteren Pädagog:innen leisten?
- Welche Rolle spielen Daten aus dem Systemmonitoring für die Systemsteuerung und -entwicklung durch Politik auf Bundes- und Landesebene?

In jedem Kapitel wird darauf eingegangen, welche Daten erhoben und den unterschiedlichen Akteur:innen rückgemeldet werden sollten, wie diese Daten verständlich aufbereitet werden können und wie eine effektive Nutzung der Daten für wirksame Maßnahmen unterstützt werden kann.



Dabei orientiert sich das Gutachten an folgenden grundlegenden Prinzipien:

- Der hier verwendete Begriff der Daten schließt neben Daten für Systemmonitoring, Schul- und Unterrichtsentwicklung auch diagnostische Informationen ein. Wir lehnen uns an die Arbeiten zum *formativen Assessment* an (Black & Wiliam, 2009; Souvignier & Hasselhorn, 2018). Diagnostische Informationen werden hier als Daten zum Lern- und Entwicklungsstand von Schüler:innen verstanden. Auf ihrer Grundlage erfolgt ein individualisiertes Feedback und werden Entscheidungen für die „Optimierung des Lehr-Lern-Prozesses“ (Souvignier & Hasselhorn, 2018, S. 694) getroffen. Dafür ist eine (möglichst digitale) lernprozessbegleitende Diagnostik erforderlich (*Diagnosefunktion von Daten*).
- Die Erfassung, Interpretation und Nutzung von Daten im Bildungssystem ist kein Selbstzweck. Sie ist einer effektiven Lern- und Entwicklungsförderung aller Kinder und Jugendlichen verpflichtet. Entsprechend müssen Erfassung, Interpretation und Nutzung von Daten fokussiert erfolgen und auf klare Ziele ausgerichtet sein. Eine Datafizierung aller schulischen Prozesse sowie eine Überforderung schulischer Akteur:innen im Umgang mit Daten sind zu vermeiden (*Datensparsamkeit*).
- Schule hat einen breiten Bildungsauftrag. Im Zentrum steht das fachliche Lernen. Schule stellt aber auch eine wichtige Lern- und Entwicklungsumwelt für den Erwerb überfachlicher Kompetenzen dar. Dieser breite Bildungsauftrag muss sich auch in der Definition und Operationalisierung von Bildungszielen sowie in der Erfassung von Ergebnissen schulischer Bildung niederschlagen (*Fokus auf fachliche und überfachliche Bildungsergebnisse*).
- Für die Bereitstellung von Daten für unterschiedliche Akteur:innen ist eine datenschutzkonforme und nutzerorientierte technische Infrastruktur erforderlich. Dies gilt auch für die Bereitstellung evidenzbasierter Fördermaterialien und digitaler Tools. Hier sind erhebliche Entwicklungsanstrengungen erforderlich, was ein ländergemeinsames Vorgehen mit Unterstützung des Bundes nahelegt (*Gewährleistung technischer und rechtlicher Voraussetzungen*).
- Akteur:innen auf allen Ebenen des Bildungssystems benötigen Unterstützung bei der Interpretation und Nutzung von Daten. Dies betrifft einerseits adressatenspezifische Qualifikationen und verständliche Rückmeldesysteme, andererseits müssen Schnittstellen zu (Fach-)Beratungs- und Unterstützungsangeboten für eine gezielte, auf Daten sowie wissenschaftliche Evidenz gestützte Unterrichts- und Schulentwicklung präzise definiert werden; Maßnahmen zur Unterrichtsentwicklung und Lernförderung kommt dabei eine besondere Bedeutung zu (*Unterstützung der Datennutzung*).
- Die Akteur:innen im Mehrebenensystem der Bildung haben unterschiedliche Funktionen und benötigen unterschiedliche Daten. Im Zentrum stehen Lehrkräfte als professionelle Expert:innen für eine Lernförderung und Unterrichtsgestaltung, die auf Diagnose und Evaluation beruht. Datengestützte Schulentwicklung sorgt für Kooperation zwischen und Unterstützung von Lehrkräften. Rechenschaftslegung und Controlling der Schule werden durch die Schulaufsicht sichergestellt. Politik definiert dafür die Rahmenvorgaben. Rahmenvorgaben und Controlling dürfen professionelle Gestaltungsspielräume nicht unzulässig einschränken (*Gestaltungsspielräume und Rechenschaftskaskade*).
- Die Einigung der Länder auf die Einführung und Überprüfung von Bildungsstandards sowie auf zentrale Abschlussprüfungen schafft einen verbindlichen Rahmen für unterschiedliche landesspezifische Steuerungsstrategien im Sinne des Wettbewerbsföderalismus. Großes Potenzial für eine



stärkere Kooperation zwischen den Ländern hat die gemeinsame Entwicklung von diagnostischen Instrumenten und Programmen zur diagnosebasierten Lernförderung und Unterrichtsentwicklung auf der Grundlage wissenschaftlicher Evidenz (z. B. *QuaMath*). Der Bund kann diese Kooperation durch Infrastrukturmaßnahmen und die Finanzierung evidenzbasierter, mit den Ländern gut abgestimmter Programme wie das Startchancen-Programm unterstützen (*Outputsteuerung und Evidenzorientierung*).

Mit dem Startchancen-Programm wurde das Bekenntnis zur datengestützten Schul- und Unterrichtsentwicklung gestärkt. Das vorliegende Gutachten will nicht nur einen Beitrag leisten zur kritischen Sichtung und Anpassung der im Anschluss an die Veröffentlichung der Gesamtstrategie der KMK in den letzten zwanzig Jahren implementierten Datenerhebungen und Instrumente. Das Gutachten gibt auch für das Startchancen-Programm konkrete Hinweise zur Implementation datengestützter Entwicklung und Steuerung von Schulen.

Obwohl das Gutachten einen Schwerpunkt auf die Datennutzung in Schulen legt, wird in einem eigenen Kapitel auch ein Blick auf die frühkindliche Bildung geworfen. Hier müssen allerdings besondere Bedingungen der Organisation und Steuerung beachtet werden. Diese bestehen nicht nur darin, dass frühkindliche Bildung in der Kinder- und Jugendhilfe verortet ist und daher große Verantwortung und Entscheidungsspielraum bei den Kommunen und Trägern liegen, sondern auch darin, dass der Bildungsauftrag der frühkindlichen Bildung ganzheitlich ist und die Vorbereitung auf schulisches Lernen nur einen Teilaspekt des Bildungsauftrags erfasst. Auch was die Erfassung und Nutzung von Daten betrifft, müssen Besonderheiten der frühkindlichen Bildung in Rechnung gestellt werden. Bislang spielen Erfassungen des Entwicklungsstands auf Basis standardisierter Verfahren eine deutlich kleinere Rolle als in den Schulen. Daten und Dokumentationen auf Basis der Beobachtung des pädagogischen Alltags spielen allerdings eine große Rolle.

Obwohl Unterschiede in den organisationalen und institutionellen Bedingungen zwischen Schule und frühkindlicher Bildung bestehen, haben Daten in beiden Feldern gleichermaßen hohes Potenzial. Unter Berücksichtigung der besonderen Bedingungen der frühkindlichen Bildung werden in einem eigenständigen Kapitel Fragen des Systemmonitorings und Controllings, der Organisationsentwicklung, der Verbesserung der Qualität der pädagogischen Angebote und der Förderung der kindlichen Entwicklung sowie der Erziehungs- und Bildungspartnerschaften adressiert.

Kapitelübergreifende Empfehlungslinien

Datengestützte Steuerung und Entwicklung ist nur dann erfolgreich, wenn über die verschiedenen Ebenen und Funktionen hinweg Ziele klar definiert sind. Nur auf dieser Grundlage kann Kohärenz zwischen unterschiedlichen Maßnahmen hergestellt werden. Fachliche und überfachliche Lern- und Entwicklungsprozesse von Schüler:innen müssen dabei – auch als Voraussetzung gesellschaftlicher Teilhabe – ins Zentrum der Zielformulierung gerückt werden. Was die fachlichen Leistungen betrifft, existieren mit den Bildungsstandards länderübergreifend festgelegte klare Erwartungen, wenngleich die Verfahren zur Anpassung der Bildungsstandards bislang noch etwas schwerfällig sind und Bildungsstandards nicht für alle Fächer vorliegen. Für überfachliche Bildungsziele wurden in Deutschland im Unterschied zu anderen OECD-Ländern bislang noch keine gemeinsamen Rahmenmodelle definiert.



Was die Bereitstellung von Daten sowie von Materialien zur datengestützten Schul- und Unterrichtsentwicklung betrifft, wurden in den Ländern und teilweise auch durch Stiftungen unterschiedliche Plattformen sowie (digitale) Tools entwickelt, die allerdings nur teilweise kompatibel sind. Neben konzeptionellen Überlegungen zur gezielten Auswahl und Aufbereitung relevanter Daten, ist eine flexible technische Infrastruktur erforderlich, die eine Rückmeldung von Daten an unterschiedliche Akteur:innen (z. B. in Dashboards) sowie die Bereitstellung von Materialien für datengestützte Entwicklung ermöglicht. Erste Entwicklungsschritte wurden von unterschiedlichen Ländern hier bereits unternommen, und eine Vernetzung zwischen einigen Ländern wurde initiiert.

Datengestützte Entwicklung und Steuerung erfordert spezifische Kompetenzen der Datennutzung, die bislang nicht in Aus-, Fort- und Weiterbildung von Lehrkräften systematisch vermittelt werden. Grundlegend sind für alle Akteur:innen im Mehrebenensystem Kompetenzen für das Verständnis und die Interpretation von Daten (*Data Literacy*), die Kommunikation über Daten (Daten-Feedback) sowie die Ableitung evidenzgestützter Maßnahmen zu den jeweiligen fachlichen und überfachlichen Lernzielen. Übergreifend sind zudem die Themen Datenethik, Datenschutz und klare Regeln zur Zweckbindung und Löschung von Daten zu berücksichtigen. Darüber hinausgehend besteht bezogen auf die jeweiligen Funktionen auch ein spezifischer Qualifikationsbedarf. In allen drei Phasen der Lehrkräftebildung müssen Kompetenzen der Datennutzung verankert werden. Dazu muss eine systematische Entwicklung von Qualifikationszielen sowie -angeboten erfolgen. Die Kooperation zwischen Landesinstituten und einschlägiger Wissenschaft ist eine Erfolgsbedingung für wirksame Qualifikation.

Kapitelübergreifende Empfehlungslinien sind daher:

Empfehlung 1) Ziele datengestützter Entwicklung, die Lern- und Entwicklungsprozesse der Schüler:innen fokussieren, klar und eindeutig definieren, fortschreiben und implementieren; dabei

1. die Nutzung der vorliegenden Bildungsstandards als ländergemeinsame Grundlage der Qualitätsentwicklung stärken und prüfen, für welche weiteren Fächer Bildungsstandards entwickelt werden sollten,
2. einen agilen Prozess für die Weiterentwicklung der Bildungsstandards entwickeln und umsetzen, um veränderten Anforderungen (z. B. im Bereich Bildung in der digitalen Welt) Rechnung tragen zu können,
3. die Überarbeitung der Kerncurricula und Lehrpläne der Länder an die Weiterentwicklung der Bildungsstandards zeitlich anpassen,
4. ein Rahmenmodell für überfachliche Kompetenzen in Anlehnung an einschlägige internationale Modelle ländergemeinsam entwickeln.



Empfehlung 2) Eine nutzungsfreundliche technische Infrastruktur zur adressatenbezogenen Bereitstellung von Daten sowie Materialien für unterschiedliche Akteur:innen im Mehrebenensystem ländergemeinsam und in Kooperation mit dem Bund entwickeln; dabei

1. ein Konzept für eine sparsame und gezielte Auswahl und Aufbereitung von Daten für unterschiedliche Funktionen im Mehrebenensystem entwickeln,
2. eine sichere und datenschutzkonforme technische Infrastruktur als Grundlage für nutzerfreundliche, interaktiv nutzbare Datendashboards mit verständlichen Visualisierungsmöglichkeiten und Vorschlägen für mögliche Maßnahmen schaffen,
3. eine gemeinsame Plattform (alternativ: eine Metaplattform) entwickeln, die unterschiedlichen Akteur:innen mit definiertem Rollen- und Rechtemanagement einen niedrighwelligen Zugang zu evidenzbasierten Materialien zur datengestützten Entwicklung ermöglicht; dafür muss eine aussagekräftige Indexierung (Verschlagwortung) entwickelt werden,
4. Interoperabilität und offene Schnittstellen sicherstellen, um Daten aus unterschiedlichen Quellen integrieren zu können,
5. Barrierefreiheit und nutzungsfreundliches Design umsetzen (u. a. Orientierung an WCAG/BITV, responsives Design für unterschiedliche Endgeräte sowie adressatenspezifische Nutzeroberflächen für Schulleitungen, Lehrkräfte, Schulträger und Schulaufsicht sowie Eltern/Erziehungsbeauftragte und Schüler:innen),
6. Standards für Datenqualität und -transparenz etablieren (Angabe von Datenquellen, Stichtagen, Aggregationsebenen und Erläuterungen zu Kennzahlen direkt im Dashboard).

Empfehlung 3) Qualifikationsziele für datengestützte Entwicklung und Steuerung im Schulsystem für unterschiedliche Funktionen entwickeln in Kooperation von Landesinstituten und einschlägiger Wissenschaft; dabei

1. Kompetenzen datengestützter Entwicklung in der ersten, zweiten und dritten Phase der Lehrkräftebildung verankern,
2. datengestützte Entwicklung mit Schwerpunkt auf Schulentwicklung, Schulmanagement und Personalentwicklung in der Schulleitungsqualifikation verankern,
3. einen Qualifikationsrahmen für Schulaufsicht mit Schwerpunkt auf datengestütztes Controlling und Unterstützung entwickeln,
4. Qualifikationsangebote für Funktionsstellen in der Schulverwaltung mit einem Schwerpunkt auf datengestützter Steuerung entwickeln.



1. Datengestützte Unterrichtsentwicklung und diagnosegeleitete Lernförderung

Unterricht hat das zentrale Ziel, die in den Bildungsstandards und Lehrplänen festgelegten Kompetenzen Lernender zu fördern. Daten unterstützen Lehrkräfte bei der diagnosegeleiteten Lernförderung und der Entwicklung des Unterrichts, indem sie Ausgangslagen, Lernprozesse und -ergebnisse sowie Merkmale der Unterrichtsprozessqualität beschreiben. Parallel haben Lehrkräfte die Aufgabe, überfachliche Kompetenzen wie Selbstregulation oder soziale Kompetenzen im Unterricht zu fördern.

Wenn Daten mit effektiven Anschlussmaßnahmen verknüpft sind, regen sie einen kontinuierlichen Entwicklungszyklus an. Dazu werden Daten erhoben, die sich in ihrer Zielsetzung, dem zeitlichen Abstand der Erhebungen und ihrer Interpretation unterscheiden, auch wenn es Überlappungen gibt:

Daten zu fachbezogenen Lernausgangslagen, -fortschritten und -ergebnissen

- *Zentrale Lernstandserhebungen* helfen dabei, den Unterricht in einem längeren Zeithorizont zu planen und weiterzuentwickeln. Sie werden zentral erstellt (z. B. durch das IQB) und administriert (z. B. durch die Qualitätseinrichtungen der Länder) sowie in Verantwortung der Länder ausgewertet (z. B. VERA, StarS). Sie orientieren sich an den Bildungsstandards der KMK und werden oft an wichtigen Übergängen im Schulsystem eingesetzt. Zu Beginn eines neuen Abschnitts, insbesondere zu Beginn der Grundschule und der Sekundarstufe I, geben sie Auskunft über die Lernausgangslage der Lernenden. Gegen Ende eines Abschnitts weisen sie darauf hin, in welchen Bereichen verstärkte Anstrengungen unternommen werden sollten, um die jeweiligen Ziele zu erreichen. Zentrale Lernstandserhebungen können zudem Aufschluss geben über die Wirksamkeit ergriffener Maßnahmen.
- *Lernprozessbegleitende Diagnoseinstrumente* helfen Lehrkräften dabei, Unterricht in kürzeren Zeithorizonten zu planen und weiterzuentwickeln. Sie werden von den Schulen bzw. Lehrkräften eigenverantwortlich eingesetzt und ermöglichen es, Lernfortschritte in einem klar umgrenzten Bereich formativ zu begleiten (z. B. die orthografische Entwicklung oder die Entwicklung der Lesegeschwindigkeit in Deutsch, das Stellenwertverständnis in Mathematik oder das Hörverstehen alltagssprachlicher Dialoge im Englischunterricht). Dazu lässt sich ein breites Spektrum von informellen (z. B. Beobachtungsraster für Kommunikationsverhalten) bis hin zu standardisierten Instrumenten und Verfahren einsetzen (z. B. die Lesetests von *quop*). Diese Art von Diagnosen bezieht sich auf den konkreten Unterricht vor Ort und ist flexibel einsetzbar, im Sinne *formativer Assessments for Learning*, und mit anschließenden Fördermaßnahmen direkter verknüpfbar.

Diese Unterscheidung ist allerdings nicht trennscharf: Auch *zentrale Lernstandserhebungen* können z. B. durch häufigere Testzeitpunkte formativ genutzt werden (z. B. der *LIFT-Test (Lesen – individuelle Fertigkeiten testen)* in Hamburg). Und manche *lernprozessbegleitenden Diagnoseinstrumente* sind auch als zentrale Kompetenztests einsetzbar. Zu unterscheiden sind hier jeweils der Zweck der Testungen (geben Aufschluss über Lernausgangslage, Lernergebnis, Lernprozess, Lernverlauf), der Einsatzbereich



(Klassenstufen, Fächer, erfasste Kompetenzen etc.) und die Testzeitpunkte bzw. -zeiträume (z. B. jährlich, mehrfach im Schuljahr etc.).

Daten zu sozial-emotionalen Merkmalen

- *Instrumente zur Erfassung schulbezogenen Wohlbefindens, schulbezogener Zufriedenheit und sozialer Eingebundenheit* ermöglichen eine Erfassung von Daten, die Lehrkräften insbesondere hinsichtlich der Gestaltung des Klassenklimas und der konstruktiven Unterstützung von Schüler:innen Entwicklungsbedarf aufzeigen. Sozial-emotionales Lernen ist in unterschiedlicher Weise in den Bildungsplänen der Länder sowie in den Standards für die Lehrkräftebildung verankert. Allerdings sind Instrumente zur Erfassung dieser Merkmale in deutschen Schulen bislang kaum systematisch implementiert.
- *Instrumente zur Erfassung sozial-emotionaler Kompetenzen sowie Kompetenzen der Selbstregulation* helfen insbesondere bei Schüler:innen, die einen besonderen Unterstützungsbedarf haben, Ansatzpunkte für gezielte Maßnahmen zu identifizieren. Diese Instrumente sind bislang vor allem zur Förderung von Schüler:innen mit besonderem Förderbedarf implementiert. Im Rahmen von StarS werden Instrumente entwickelt, die über die basalen fachbezogenen Kompetenzen hinaus auch Aspekte der Selbstregulation und motivational-emotionale Orientierungen erfassen.

Daten zu Unterrichtsprozessen

- *Instrumente zur Erfassung von Unterrichtsqualität* durch Schülerbefragungen oder beobachtungsbasierte kollegiale Unterrichtsfeedbacks, die sich auf die zentralen lernwirksamen Basisdimensionen der Unterrichtsqualität beziehen, orientieren/unterstützen eine kontinuierliche Unterrichtsentwicklung durch Lehrkräfte. Inzwischen liegen einige validierte Instrumente zur Erfassung von Unterrichtsqualität vor, die explizit für den praktischen Einsatz im Unterricht entwickelt wurden.

Learning Analytics haben im Kontext von Unterricht das Potenzial, nahezu in Echtzeit statische und dynamische Daten über Aktivitäten von Lernenden in digitalen Lernumgebungen zu analysieren. Sie visualisieren nicht nur Lernergebnisse, sondern auch Lernaktivitäten und ermöglichen damit eine frühzeitige gezielte Unterstützung. *Classroom Analytics* ist ein Teilbereich, der speziell den Unterrichtskontext fokussiert und typischerweise auf zeitnahe Rückmeldungen für Lehrkräfte abzielt, z. B. zur Interaktion, Aufmerksamkeit und Beteiligung im Klassenraum. Learning Analytics und Classroom Analytics haben, was die datengestützte Lernförderung und Unterrichtsentwicklung betrifft, perspektivisch ein hohes Potenzial (Knoop-van Campen et al., 2023). Mittel- und langfristig werden formative Assessments wahrscheinlich überwiegend in Form von Learning Analytics erfolgen. Dazu ist allerdings eine stärkere Integration fachdidaktischer, psychometrischer und informatischer Forschung erforderlich (vgl. SWK-Gutachten zur Digitalisierung, 2022b).



1.1 Situationsanalyse

1.1.1 Zentrale Lernstandserhebungen

Ein Überblick über die in den Ländern verpflichtend und landesweit implementierten zentralen Lernstandserhebungen (Abbildung 1) zeigt, dass alle Länder in den Jahrgangsstufen 3 und 8 landesweite Vergleichsarbeiten (VERA) durchführen.¹

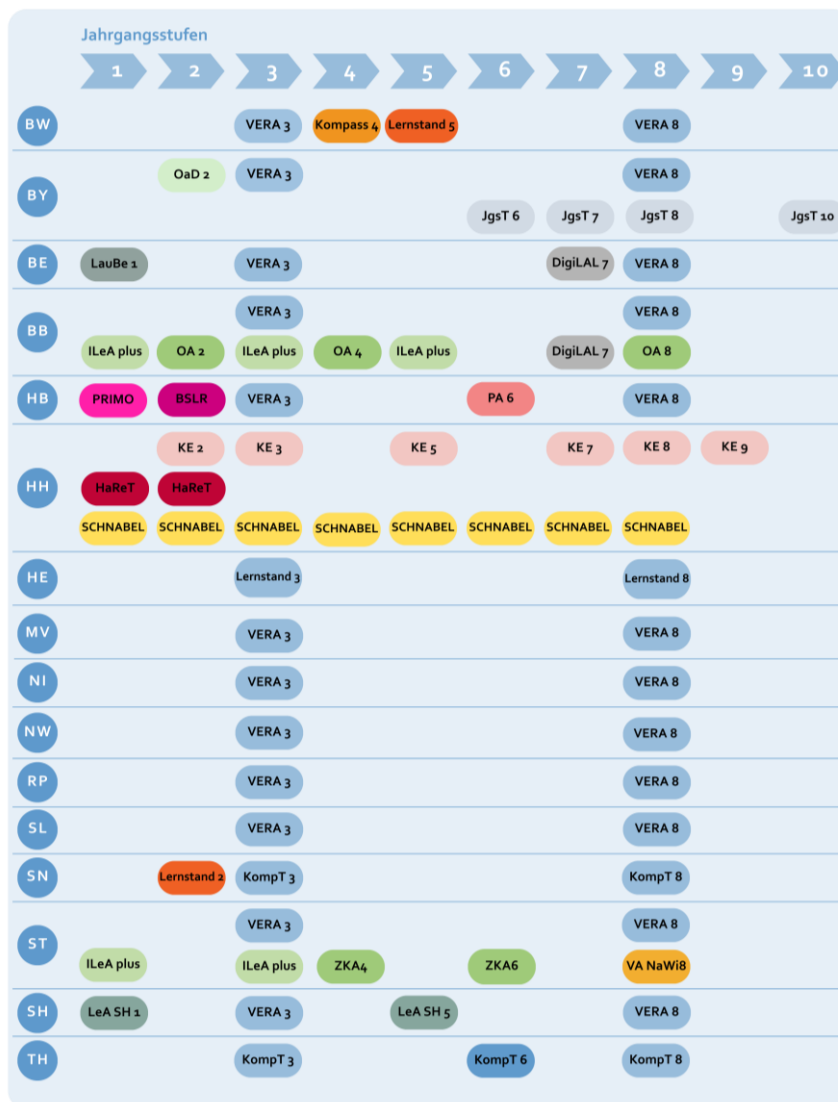


Abbildung 1: Zentrale Lernstandserhebungen in den Ländern

Anmerkungen: VERA – Vergleichsarbeiten | OaD – Orientierungsarbeit im Fach Deutsch | JgsT – Jahrgangsstufentest | LauBe – Lernausgangslage Berlin | DigiLAL – Digitale Lernausgangslage | ILeA – Individuelle Lernausgangslage | OA – Orientierungsarbeit | PRIMO – Sprachstandsfeststellung ein Jahr vor der Einschulung | BSLR – Bremer Screening für Lesen und Rechtschreiben | KE₂-KE₉ – Kompetenzen ermitteln | HaReT – Hamburger Rechentest | SCHNABEL – schreiben, nachdenken, anwenden, behalten, erfolgreich lernen | KompT – Kompetenztest | ZKA – zentrale Klassenarbeiten | VA NaWi – Vergleichsarbeit Naturwissenschaften | LeA SH – Lernausgangslage Schleswig-Holstein. Quelle: eigene Recherchen

¹ In Sachsen und Thüringen wird VERA unter der Bezeichnung Kompetenztests (KompT) und in Hessen unter der Bezeichnung Zentrale Lernstandserhebungen (Lernstand) durchgeführt.



Während sechs Länder sich auf die VERA-Erhebungen in den Jahrgängen 3 und 8 beschränken,² führen viele Länder weitere Erhebungen durch (z. B. Analysen der Lernausgangslage bei Schuleintritt oder Erhebungen zu einzelnen Kompetenzbereichen wie Rechtschreibung) und ergänzen damit VERA um zusätzliche Testzeitpunkte. VERA ist Bestandteil der Gesamtstrategie der KMK zum Bildungsmonitoring und gibt Lehrkräften eine Rückmeldung darüber, inwieweit ihre Schüler:innen bereits in der dritten bzw. achten Jahrgangsstufe über zentrale Kompetenzen verfügen, die mit den Bildungsstandards der KMK für das Ende der Primarstufe bzw. für den Ersten und Mittleren Schulabschluss definiert wurden. Das Instrument hat damit in den meisten Ländern ein Jahr vor dem Übergang in die weiterführende Schule³ bzw. die berufliche Bildung eine „Frühwarnfunktion“ mit Blick auf das Erreichen der Bildungsstandards. Anders als die *IQB-Bildungstrends* und die internationalen Schulleistungsstudien dient VERA nicht als Instrument des Systemmonitorings, sondern zielt vor allem auf die Schul- und Unterrichtsentwicklung ab. Mit entsprechenden Ergebnismeldungen erhalten Lehrkräfte Informationen darüber, wie sich die Lernenden in ihrer Klasse auf die Kompetenzstufen verteilen. In den meisten Ländern beinhalten die Rückmeldungen auch soziale Vergleiche, die Auskunft darüber geben, inwieweit ein ähnliches Niveau erreicht wird wie in anderen Klassen und Schulen insgesamt bzw. in anderen Klassen und Schulen mit ähnlicher Schülerschaft. Aus den Ergebnissen lassen sich Bedarfe für die Unterrichtsentwicklung und Förderung ableiten. Dazu werden im Rahmen von VERA Materialien, Handreichungen und Unterrichts Anregungen zur Verfügung gestellt, die Lehrkräfteteams in Schulen dabei unterstützen sollen, Ziele für die Unterrichtsentwicklung abzuleiten und Förderentscheidungen zu treffen.

Da der Anteil der Kinder, die am Ende der Primarstufe die Mindeststandards verfehlen, in den letzten Jahren zugenommen hat und deutlich zu hoch ist (vgl. IQB-Bildungstrend 2021; Stanat et al., 2022), entwickelt das IQB für die Länder zusätzlich zu VERA eine Reihe von Instrumenten zur Erfassung basaler Kompetenzen, die Kinder entwickeln müssen, um eine Chance zu haben, die Mindeststandards zu erreichen. Dies umfasst zum einen das Instrument StarS, das darauf abzielt, nach dem Übergang vom Elementar- zum Primarbereich die Lernausgangslage (Jahrgang 1) und die Lernentwicklung (Jahrgang 2) der Schüler:innen festzustellen. Der Fokus liegt auf basalen Kompetenzen in den Bereichen Sprache und Mathematik sowie auf Aspekten von Selbstregulation und motivationalen Orientierungen. StarS soll zum einen eine adaptive Unterrichtsgestaltung und Förderung der Kinder unterstützen und zum anderen für ein bundesweites Bildungsmonitoring auf Systemebene zu Beginn der Grundschulzeit (ähnlich wie bei den IQB-Bildungstrendstudien in den Jahrgangsstufen 4 und 9) genutzt werden. Die Tests für StarS werden ab 2027 den Ländern zur Verfügung gestellt, gemeinsam mit darauf bezogenen Fortbildungsmodulen und Hinweisen auf Fördermaßnahmen. Und auch für die Jahrgangsstufen 2 und 3 entwickelt das IQB Tests zur Erfassung basaler Kompetenzen in den Bereichen Sprache und Mathematik, die Lehrkräfte zu mehreren Erhebungszeitpunkten vor der Durchführung von VERA 3 einsetzen können (*Basiskompetenztests, BKT*). Diese Tests werden den Ländern bereits Ende 2026 für die Unterrichtsentwicklung zur Verfügung stehen.

² In Niedersachsen nur die am Startchancen-Programm teilnehmenden Schulen. Perspektivisch sollen VERA 3 und 8 in Niedersachsen wieder flächendeckend eingeführt werden, außerdem Lernstandserhebungen in den Klassen 1 und 5 (s. Bildung.Table, 21.01.2026 <https://table.mediabildung/news/niedersachsen-fuehrt-vergleichsarbeiten-wieder-ein> (aufgerufen am 20.02.2026)).

³ Insbesondere in Berlin und Brandenburg endet die Bildungsetappe im Primarbereich überwiegend erst nach der sechsten Jahrgangsstufe.



Mit diesen Instrumenten (StarS, BKT, VERA), die digital einsetzbar sind, haben die Länder gemeinsam mit dem IQB eine Grundlage für datengestützte Unterrichtsentwicklung geschaffen, die auf das Erreichen der Bildungsstandards ausgerichtet ist. Im Primarbereich wurde dabei der Fokus auf das Erreichen der Mindeststandards gestärkt, die für erfolgreiches Weiterlernen erforderlich sind. Die Instrumente werden auch in ländergemeinsamen Unterrichtsentwicklungsprojekten wie QuaMath und Startchancen integriert, um Lehrkräfte für die Entwicklung entsprechender Fördermaßnahmen zu professionalisieren.

In einzelnen Ländern werden über die ländergemeinsam entwickelten Instrumente hinaus weitere Verfahren eingesetzt. In Hamburg etwa besteht mit *KERMIT (Kompetenzen ermitteln)* eine weitgehend durchgängige Diagnostik in den Jahrgangsstufen 2, 3, 5, 7, 8 und 9, bei der die VERA-Tests (Jahrgangsstufen 3 und 8) und eigene Tests (Jahrgangsstufen 2, 5, 7, 9) eingesetzt werden. Um eine systematische Nutzung der Ergebnisse aus den Testungen zu unterstützen, wird dies flankiert durch ein Unterstützungssystem mit Fortbildungen und datengestützten Dialogen zwischen Schulaufsicht und Schulleitung sowie Schulleitung und Fachkonferenzen (vgl. Kapitel 2 und 3). Auch in Baden-Württemberg⁴ wird ergänzend zu VERA 3 und VERA 8 eine zentrale Lernstandserhebung in Klasse 5 für die Fächer Deutsch und Mathematik durchgeführt (*Lernstand 5*). Nordrhein-Westfalen plant, bis zum Jahr 2030 neben VERA 3 und VERA 8 weitere Lernstandserhebungen in den Jahrgangsstufen 2, 5 und 7 einzuführen (NRW-Schulkompass 2030, Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2025a, S. 4). Ebenso sieht Berlin⁵ im Rahmen seiner *Strategie zur Steigerung der Bildungsqualität* die Ermittlung des Lernverlaufs durch eine weitere Durchführung von VERA 3 und VERA 8 in der jeweils darauffolgenden Jahrgangsstufe vor. Und auch andere Länder setzen über die gemeinsam entwickelten Instrumente hinaus weitere Verfahren ein, z. B. Jahrgangsstufentests, zentrale Klassenarbeiten oder Lernausgangslagenuntersuchungen.

Zentral für die Unterrichtsentwicklung sind Rückmeldungen von Ergebnissen aus den zentralen Lernstandserhebungen auf Schul- und Klassenebene. Sie sollen Hinweise darauf geben, inwieweit die Förderung der Schüler:innen in den erfassten Kompetenzbereichen gut gelingt bzw. inwieweit Weiterentwicklungsbedarf besteht. Im Rahmen von VERA kommt zur Einordnung der Ergebnisse in einigen Ländern die *kriteriale Bezugsnorm der Bildungsstandards* zur Anwendung, d. h., Lehrkräfte erhalten eine Rückmeldung darüber, welcher Anteil ihrer Schüler:innen die Mindest-, Regel- und Optimalstandards erreicht. Dabei ist zu beachten, dass die Schüler:innen noch ein Jahr (VERA 3) bzw. ein bis zwei Jahre (VERA 8) Zeit haben, um die Bildungsstandards zu erreichen. Um dies zu unterstreichen und die Interpretation der teilweise recht stark besetzten unteren Kompetenzstufe zu unterstützen, wird diese in Hamburg in zwei Fähigkeitsniveaus unterteilt.

In fast allen Ländern werden zudem *soziale Vergleiche* vorgenommen, d. h., die in einer Schule oder Klasse erzielten Ergebnisse werden denen anderer Schulen oder Klassen gegenübergestellt. Damit erhalten die Adressat:innen eine Rückmeldung darüber, inwieweit in der betrachteten Gruppe insgesamt ähnliche, bessere oder schwächere Leistungen erzielt werden als in einer Vergleichsgruppe (z. B. die Vergleichsgruppe der Drittklässler:innen im jeweiligen Land insgesamt). Dies ist eine relevante Information, da sie die

⁴ <https://km.baden-wuerttemberg.de/de/schule/datengestuetzte-qualitaetsentwicklung/lernstandserhebungen-an-den-allgemein-bildenden-schulen> (aufgerufen am 20.02.2026).

⁵ <https://www.berlin.de/sen/bjf/bildungsqualitaet/> (aufgerufen am 20.02.2026).



Ergebnisse der Schule bzw. Klasse einordnet. Allerdings lassen sich die hierbei beobachteten Unterschiede in der Regel nicht allein auf Wirkungen der Schule bzw. des Unterrichts zurückführen, da die Kompetenzentwicklung von Kindern und Jugendlichen in erheblichem Maße auch durch andere Faktoren beeinflusst wird, die von Lehrkräften nicht beeinflussbar sind. Hierzu gehören insbesondere familiäre Bedingungen des Aufwachsens, die für den Lernerfolg relevant sind, wie etwa die in der Familie gesprochenen Sprachen und das sprachliche Anregungsniveau. Da sich Schulen teilweise erheblich darin unterscheiden, welche Lernvoraussetzungen ihre Schüler:innen mitbringen, ist es wichtig, dies bei der Interpretation von Ergebnissen aus Lernstandserhebungen zu berücksichtigen. Dies wird im Rahmen von VERA in einigen Ländern⁶ durch die Rückmeldung von Vergleichswerten unterstützt, die sich konzeptuell auf eine ähnlich zusammengesetzte Schülerschaft beziehen und zumeist als *faire* bzw. *adjustierte Vergleiche*⁷ bezeichnet werden. Die Vorgehensweisen für solche adjustierten Vergleiche unterscheiden sich zwischen den Ländern erheblich.

Ein grundlegender Unterschied besteht darin, wie die Vergleichswerte generiert werden. So werden in einigen Ländern die Ergebnisse der jeweiligen Schule mit den Ergebnissen einer unterschiedlich großen Gruppe anderer Schulen mit ähnlich zusammengesetzter Schülerschaft verglichen. In anderen Ländern hingegen werden mit regressions- bzw. mehrebenenanalytischen Verfahren sogenannte *Erwartungswerte* gebildet. Dabei wird anhand der empirisch vorgefundenen Zusammenhänge zwischen Hintergrundmerkmalen und Leistungen der Schüler:innen geschätzt, welche Ergebnisse für die jeweilige Schule aufgrund der Zusammensetzung ihrer Schülerschaft zu erwarten wären.

Das Vorgehen der Länder unterscheidet sich ferner darin, welche Hintergrundmerkmale der Schüler:innen bei den adjustierten Vergleichen herangezogen werden. In der Regel werden Indikatoren des sozioökonomischen Status der Familien sowie des Migrations- bzw. Sprachhintergrunds verwendet, die auf Individual- oder Schulebene vorliegen. Bei den Indikatoren auf Schulebene handelt es sich häufig um einen Index, in den mehrere Merkmale eingehen. Die Ausgangsleistungen der Schüler:innen werden bisher nur in wenigen Ländern, z. B. in Baden-Württemberg und Thüringen (Fiege, 2016), berücksichtigt, obwohl das fachliche Vorwissen einen der wichtigsten Prädiktoren für die Lernfortschritte darstellt (z. B. Geary et al., 2017). Daher ist die wiederholte Kompetenzmessung wichtig, um nicht nur adjustierte Vergleiche von Lernständen, sondern auch von Lernfortschritten zu ermöglichen.

Gemäß der *SWK-Länderabfrage Schule* (vgl. Anhang 1) erhalten Lehrkräfte in allen Ländern Rückmeldungen aus den VERA-Erhebungen auf Individualebene (d. h. bezogen auf die einzelne Schülerin oder den einzelnen Schüler) sowie auf Klassenebene. Darüber hinaus erfolgt in jeweils zwölf Ländern eine Rückmeldung auf Jahrgangsebene und auf Landesebene. Neun Länder geben an, adjustierte Vergleiche durchzuführen, ohne hierzu nähere Angaben zu machen. Die Rückmeldung erfolgt in sechs Ländern über ein Datenblatt

⁶ Nach einem Bericht zum Fortgang der Arbeiten zu VERA, der dem IQB vorliegt, wurden im Jahr 2024 für VERA 3 in acht Ländern Vergleichswerte auf Ebene von Kompetenzstufen rückgemeldet. Für VERA 8 meldeten sieben Länder Vergleichswerte auf Kompetenzstufenebene zurück.

⁷ In der internationalen Literatur wird diese Bezeichnung nicht verwendet, sondern – je nach Vorgehen – *value-added*, *contextualized* oder *adjusted* (vgl. z. B. Fiege, 2013). Die Bezeichnung *fair* ist aus verschiedenen Gründen nicht ganz unproblematisch. Zum einen impliziert sie, dass diese Vergleiche absolut betrachtet fair seien, wovon aufgrund einer Reihe von Einschränkungen nicht ausgegangen werden kann (Fiege (2013) hält daher die – allerdings etwas sperrige – Bezeichnung *faire Vergleiche* für angemessener). Zum anderen legt die Bezeichnung nahe, dass andere Vergleiche, die durchaus auch relevante Informationen enthalten, unfair und damit irrelevant seien. Aus Sicht der SWK wäre die Bezeichnung *adjustierter Vergleich* vorzuziehen, da sie präziser und international anschlussfähiger ist.



und in drei Ländern über ein digitales Dashboard. Einzelne Länder geben noch weitere Formate der Rückmeldung an, wie Ergebnisworkshops auf Anfrage, Onlineportale oder Diagramme. Ferner unterscheiden sich die Länder darin, inwieweit und in welcher Form sie die Lehrkräfte darin unterstützen, die Ergebnisse in fachdidaktisch tragfähige handlungsleitende Ansätze für den Unterricht zu überführen (z. B. im Hinblick auf Material, Fortbildungsangebote etc.).

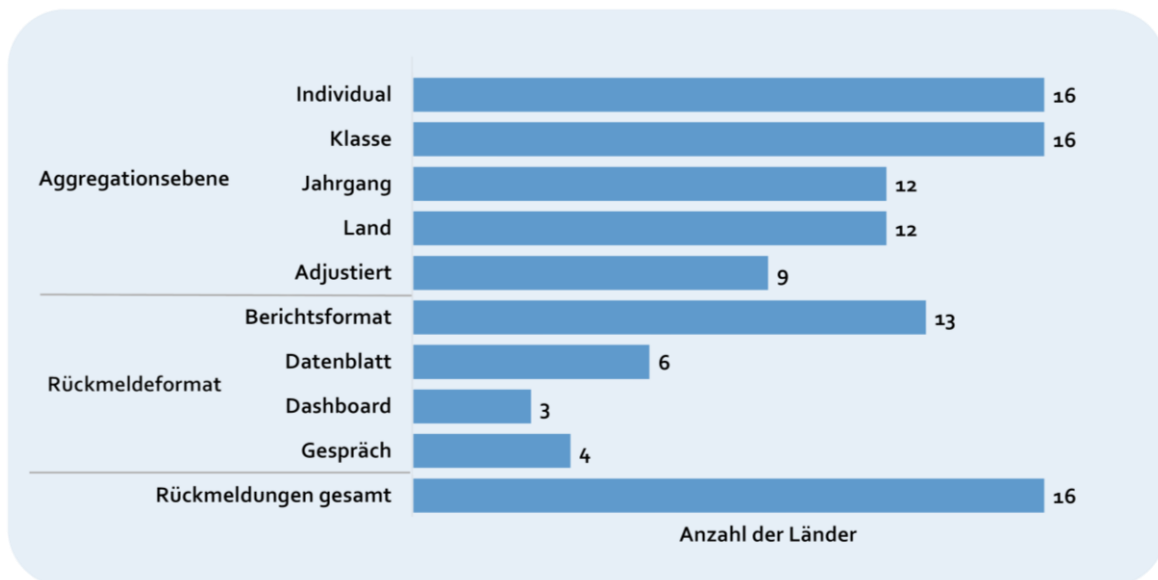


Abbildung 2: VERA-Rückmeldungen an Lehrkräfte
Quelle: SWK-Länderabfrage Schule

In der SWK-Länderabfrage Schule haben 13 Länder angegeben, dass die Rückmeldung zu den VERA-Ergebnissen mit Hinweisen für Lehrkräfte zur Unterrichtsentwicklung bzw. zur Lernförderung verknüpft ist. Für die hoch bedeutsamen Anschlussprozesse nutzen viele Länder die vom IQB bereitgestellten didaktischen Handreichungen für Schulleitungen und Lehrkräfte zum Umgang mit den Ergebnissen oder Leitfäden zur Ergebnisauswertung. Einige Länder stellen über ihre Landesinstitute bzw. Qualitätseinrichtungen digitale oder personelle Unterstützungsangebote zur Arbeit mit den Ergebnissen, kompetenzstufenspezifisches Material oder aufgabenbezogene didaktische Hinweise zur Verfügung. So bietet z. B. Berlin mit *VERA-Check* ein digitales Werkzeug für Lehrkräfte an, das passgenaue Trainingsmaterialien zu den Ergebnissen aus VERA digital bereitstellt. Vereinzelt werden Hinweise zur Weiterarbeit mit den Ergebnissen im Schulbericht bzw. auf dem Landesbildungsserver zur Verfügung gestellt. Im Rahmen des Startchancen-Programms erarbeiten die fachbezogenen Kompetenzzentren mit den Ländern zudem noch breitere fachdidaktische Unterstützung, um die Anschlussprozesse weiter zu verbessern.

Für Lernende mit einem sonderpädagogischen Förderbedarf, die zieldifferent unterrichtet werden, ist hinsichtlich der Bildungsstandards eine Orientierung an den Mindeststandards zumindest anzustreben, um ihnen eine weitestmögliche akademische und soziale Teilhabe zu ermöglichen (SWK, 2025). Dabei ist für Lernende mit den sonderpädagogischen Förderschwerpunkten Lernen und Geistige Entwicklung, die lernziel-different unterrichtet werden und teilweise deutlich unterhalb der Mindeststandards bleiben, vor allem die individuelle Bezugsnorm relevant.



1.1.2 Lernprozessbegleitende Diagnoseinstrumente

Lernprozessbegleitende Diagnoseinstrumente dienen dem *Formative Assessment for Learning* (Black & Wiliam, 2006), also der Erfassung individueller Lernausgangslagen und Lernfortschritte, insofern sie unmittelbar die unterrichtlichen Entscheidungen für diagnosegeleitete Lernförderung und sukzessive Lernprogressionen betreffen. Sie sind zusammen mit darauf abgestimmten Fördermaterialien wichtige fachdidaktische Werkzeuge, mit denen Unterricht und Förderung in Verantwortung der Lehrkräfte und Schulen datengestützt geplant und gesteuert wird, und umfassen auch viele informelle Vorgehensweisen (z. B. Sundermann & Selter, 2006).

In der SWK-Länderabfrage Schule zu bereits etablierten *lernprozessbegleitenden Diagnoseinstrumenten* zeigt sich, dass diese für unterschiedliche Fächer/Fachinhalte, Zwecke, Kompetenzbereiche und Altersgruppen (vgl. Anhang 2) angeboten und eingesetzt werden. Sie umfassen Instrumente, die sich auf ein breites Themenspektrum beziehen, bis hin zu sehr fokussierten Kurzdiagnosen, mit unterschiedlich konkreter Verknüpfung zu passenden Fördermaterialien. Allen Instrumenten gemeinsam ist, dass sie regelmäßig im Verlauf des Lernprozesses eingesetzt werden sollen. Einzelne der lernprozessbegleitenden Instrumente werden in mehreren Ländern eingesetzt. Drei häufig genutzte Instrumente werden hier beispielhaft genannt:

- *quop* ist eine computerbasierte Plattform, die für den Bereich Lesen (und einzelne Kompetenzbereiche in Mathematik in den Jahrgangsstufen 1 bis 6 sowie im Fach Englisch für die Jahrgangsstufen 5 und 6) eine kontinuierliche Lernverlaufsdagnostik ermöglicht und Lehrkräften, Eltern und Schüler:innen Rückmeldung über den Lernfortschritt der Kinder gibt. Die Wirksamkeit des regelmäßigen Einsatzes im Bereich Lesen ist empirisch nachgewiesen (Förster et al., 2024; Souvignier et al., 2014); für den Bereich Leseförderung wird quop auch im Rahmen des Bund-Länder-Programms *Leistung macht Schule (LemaS)* eingesetzt. In der Länderabfrage haben fünf Länder angegeben, quop zu nutzen, davon vier für den Bereich Lesen, drei auch für Mathematik und zwei für Englisch. Die Tests zur Mathematik werden derzeit überarbeitet.
- *ILeA plus (Individuelle Lernstands-Analyse)* ist ein digitales Instrument für formatives Assessment und eine Weiterentwicklung des 2005 entwickelten Instruments ILeA. Es ermöglicht kompetenzorientiert die regelmäßige Feststellung des Lernstands von Schüler:innen in den Fächern Deutsch und Mathematik vom Schuleintritt bis zur Jahrgangsstufe 6 und stellt die Entwicklung der Schüler:innen in ihrem individuellen Lernprozess dar. Gleichzeitig bietet das Verfahren ausführliche didaktische Hinweise zur individuellen Förderung (ILeA-plus-Handbuch, Landesinstitut für Schule und Medien Berlin-Brandenburg, 2021; Gottlebe et al., 2020; Liebers et al., 2020). In der Länderabfrage haben sechs Länder angegeben, ILeA plus zu nutzen.
- Mit *Mathe sicher können (MSK)* liegen Diagnose- und Fördermaterialien für die Sicherung arithmetischer Basiskompetenzen (insbesondere der Verstehensgrundlagen) in den Jahrgangsstufen 3 bis 7 vor, seit 2024 auch als digitale Diagnose im *MSK-Online-Check* mit handlungsleitenden Dashboards (Hankeln et al., 2025). Das diagnosegeleitete Förderkonzept wurde anhand fachdidaktischer Prinzipien entwickelt und ist verstehensorientiert und kommunikationsfördernd. Die Wirksamkeit ist sowohl für Lernzuwächse bei Lernenden als auch für die fachdidaktischen Förderkompetenzen der Lehrkräfte durch das Fortbildungskonzept wissenschaftlich nachgewiesen (Prediger



et al., 2019; Prediger et al., 2023). In der Länderabfrage haben fünf Länder angegeben, MSK zu nutzen, in weiteren Ländern wird MSK über die Programme QuaMath und Startchancen implementiert.

Weitere Instrumente sind in Anhang 2 aufgeführt. Insgesamt ergibt die Länderabfrage ein heterogenes Bild des Einsatzes lernprozessbegleitender Diagnoseinstrumente, mit vielen Instrumenten für unterschiedliche Kompetenzbereiche. Sie sind nur begrenzt mit Fördermaterialien und Fortbildungsangeboten verknüpft, und es gibt noch wenige ländergemeinsame Angebote.⁸ Der Eindruck der Heterogenität wird dadurch verstärkt, dass aus fachdidaktischen Zusammenhängen auch eine Reihe von Beurteilungs- und Bewertungsrastern auf dem Markt sind, die nicht durchweg evidenzbasiert oder empirisch erprobt sind (Behrens et al., 2014). Die Länder unterscheiden sich zudem darin, inwiefern Leitfäden zur kollegialen Unterrichtsentwicklung mit datengestützten Gesprächen etabliert sind. Dabei können auch die lernprozessbegleitenden formellen und informellen Diagnoseinstrumente einen Beitrag leisten, wenn sie entsprechend eingebunden werden (vgl. Kapitel 4.1.3).

Zur Gewährleistung einer hochwertigen inklusiven Bildung für Lernende mit einem sonderpädagogischen Förderbedarf oder anderen umfangreichen Unterstützungsbedarfen, die sehr unterschiedliche Lernausgangslagen mitbringen, ist eine individualisierte Lernprozessbegleitung und regelmäßige Überprüfung der Fortschritte anhand lernprozessbegleitender Diagnoseinstrumente im Zuge der Förderplanung verbindlich in allen Ländern vorgesehen (Gebhardt, 2023; Bernasconi et al., 2025). Instrumente und Maßnahmen zur Lernprozessdiagnostik und -begleitung sind insbesondere auch für Lernende mit einem sonderpädagogischen Unterstützungsbedarf hoch relevant und wurden teilweise explizit vor dem Hintergrund der Anforderungen inklusiver Schulen entwickelt. Im Rahmen der Verfahren zur Feststellung eines sonderpädagogischen Förderbedarfs und der anschließenden, jährlich vorgesehenen Überprüfung des Förderbedarfs sollen individuelle Entwicklungen und Leistungsfortschritte bei Lernenden mit sonderpädagogischem Förderbedarf dokumentiert werden.⁹ Die Rückmeldung an Schulleitungen, Lehrkräfte, Eltern und Schüler:innen sowie die Abstimmung über anstehende individuelle Förderziele und -maßnahmen erfolgt teilweise in Förderplangesprächen. Durch eine solche lernprozessbegleitende Diagnostik und entsprechende adaptive Förderung soll der Entstehung oder Verfestigung (ggf. weiterer) sonderpädagogischer Förderbedarfe präventiv entgegengewirkt werden.

⁸ Für den Bereich der sprachlichen Bildung ist im Rahmen der Bund-Länder-Initiative *Bildung durch Sprache und Schrift (BiSS)* eine Datenbank entstanden, die einen Überblick über die in den Netzwerken und Teilprojekten vorhandenen Instrumente zur datengestützten Unterrichtsentwicklung im Bereich Sprachbildung bereitstellt. In der BiSS-Tool-datenbank sind derzeit rund 130 Tools zur deutschen Sprache dokumentiert, die in vier Hauptkategorien unterteilt sind. Das sind neben lernprozessbegleitenden Diagnoseinstrumenten (ca. 72) auch umfassende Förderkonzepte (ca. 26) und -instrumente (ca. 20) sowie Professionalisierungskonzepte (ca. 23), die hinsichtlich wissenschaftlicher Fundierung, Gütekriterien und Praxistauglichkeit überprüft wurden.

⁹ Die Verfahren zur Feststellung eines sonderpädagogischen Förderbedarfs stehen in diesem Gutachten nicht im Fokus, sondern müssen gesondert betrachtet werden. Die Verfahren sind in den Bundesländern sehr unterschiedlich geregelt (Sälzer et al., 2015), dabei ist in den meisten Ländern ein kontinuierlicher Anstieg festgestellter sonderpädagogischer Förderbedarfe festzustellen (Scheer und Melzer, 2020; Neumann und Lütje-Klose, 2020). Zugleich weisen Studienergebnisse darauf hin, dass eine daran anknüpfende, rechtlich vorgesehene individuelle Förderung nicht durchgängig auch auf der Grundlage evidenzbasierter Methoden umgesetzt wird (Casale et al., 2023). Dass sowohl bezüglich der Feststellungsverfahren als auch der anknüpfenden Lernprozessbegleitung ein deutlicher Entwicklungsbedarf besteht, zeigt sich u. a. in den Ergebnissen eines interdisziplinär erstellten Gutachtens aufgrund eines Prüfauftrags des Landes NRW zum dortigen Verfahren zur Feststellung sonderpädagogischen Förderbedarfs (MSB NRW, 2024). Auf der Grundlage einer empirischen Analyse von 420 sonderpädagogischen Fallakten zeigte sich u. a., dass die zugrunde gelegten Definitionen und Verfahren uneinheitlich sind und nur selten standardisierte Vorgehensweisen verwendet werden.



1.1.3 Learning Analytics

Learning Analytics werden datengestützte Lernförderung und Unterrichtsentwicklung nachhaltig verändern. Der Begriff bezeichnet die Erfassung, Analyse, Visualisierung und Nutzung statischer und dynamischer Daten über Aktivitäten von Lernenden in Lernumgebungen, optimalerweise in Echtzeit während des Lernprozesses, um Lehr-Lern-Prozesse zu modellieren und ggf. frühzeitig zu optimieren (Sergis & Sampson, 2017). Learning Analytics bringen Lernendenverhalten, wie z. B. welche Aufgabe bearbeitet wurde, wie lange Lernende brauchen, welche Schritte sie wählen oder zu welchen Zeiten sie (nicht) arbeiten, mit Lernergebnissen in Zusammenhang und können – wenn geeignete Vergleichsdaten vorliegen – Datengrundlagen bieten, um Lernpfade entsprechend anpassen, Empfehlungen erstellen, Lehrkräfte auf Förderbedarf aufmerksam machen und Risikofälle identifizieren zu können. Learning Analytics sind eine Basis für Adaptivität und Feedback. Internationale Forschung zeigt, dass beide Faktoren mittlere ($d = ,48$; Wisniewski et al., 2020) bis hohe ($d = ,70$ – $,79$; Hattie, 2008, 2009) positive Effekte auf Lernprozesse und -ergebnisse haben, besonders bei Lernenden mit geringem Vorwissen (Tetzlaff et al., 2025).

Intelligente tutorielle Systeme (ITS) sind digitale, ggf. KI-basierte Lernprogramme, die Learning Analytics nutzen. Basierend auf Wissens- und Instruktionsmodellen werten sie die Bearbeitung einzelner Aufgaben unmittelbar als diagnostische Information für Feedback (*Short Cycle*) und adaptive Förderentscheidungen aus. Die gemessenen Daten sind eng auf die Lernziele bezogen, und die adaptive Zuteilung von Aufgaben basiert auf didaktischen Prinzipien. ITS sind also top-down entwickelte, auf Theorie und Fachdidaktik fundierende Anwendungen, die in der Entwicklung sehr aufwendig sind. Die internationale Forschung bestätigt insgesamt positive Effekte von ITS auf die Lernleistung (W. Ma et al., 2014).

Vorrangig wurden ITS bislang für Wissensdomänen entwickelt, in denen es klar strukturierte Wissensmodelle gibt (z. B. Grundrechenarten als Basis für Bruchrechnung) und in denen Training ein substanzieller Bestandteil des Lernens ist (z. B. Vokabeln). Während ITS in den USA seit den 1990er-Jahren in formalen Bereichen wie Mathematik und Programmierkursen systematisch genutzt werden, kommen solche Systeme in Deutschland erst in jüngerer Zeit zum Einsatz (und eingegrenzt auf wenige Wissensarten, in Mathematik z. B. vorrangig auf Fertigkeiten; Thurm & Graewert, 2022). Verschiedene Länder stellen Lernenden z. B. das Lernprogramm *bettermarks* für den Mathematikunterricht zur Verfügung. *bettermarks* ist eine adaptive Online-Plattform für das Fach Mathematik mit interaktiven Trainings- und Wiederholungsmaterialien für die Jahrgangsstufen 4 bis 13. ITS können jedoch auch für komplexere Wissensarten fachdidaktisch entwickelt werden und effektiv einsetzbar sein (Alevin & McLaughlin, 2016).

Andere Anwendungen (z. B. *FeedBook* für Englisch, *dizlesen* für Lesen, *FÖDIMA* für basale mathematische Kompetenzen) zeigen in Forschungsprojekten erste Erfolge, sind aber nicht in der Fläche ausgerollt.

Während ITS Anwendungen sind, die – top-down – auf fachdidaktischer Expertise beruhen, generieren KI-Technologien, die auf *Machine Learning (ML)* basieren, bottom-up Muster aus großen Datenmengen. Ein Beispiel: Ein ML-System erkennt in den Bearbeitungsdaten eines Mathematikurses, dass viele Schüler:innen systematisch bei Aufgaben scheitern, in denen Bruchrechnung mit negativen Zahlen kombiniert wird. Obwohl dieses spezifische Fehlermuster nicht explizit von Expert:innen vorgegeben wurde, identifiziert das System es automatisch und kann darauf basierend gezielte Hinweise geben oder zusätzliches Training vorschlagen. Inwieweit Machine Learning reliable und valide Diagnosedaten bietet, ist eine Forschungsfrage



und hängt vermutlich stark von den Wissensarten ab. Potenziale bieten z. B. die automatische Kodierung schriftlicher Leistungen oder die Generierung von Wissensfragen durch KI (Köller, 2025; Jihyun Lee et al., 2024; Yan Liu et al., 2025; vgl. auch Kapitel 5.1.13). Ein Beispiel dazu entsteht gerade in Nordrhein-Westfalen: *LeOn (Leseraum Online)* ist ein digitales Werkzeug zur Leseförderung für die Jahrgänge 2 bis 6, mit dem Lautleseproben aufgenommen werden können. Ein KI-basiertes Diagnoseinstrument wertet ab 2026 die Leseflüssigkeit automatisiert aus und gibt Hinweise zur individuellen Förderung (Kunze et al., i. E.).

Die meisten bisherigen Anwendungen, die auf Machine Learning basieren, nutzen allerdings unspezifische Daten, die beiläufig z. B. in *Lern-Management-Systemen (LMS)* anfallen, wie z. B. Klickzahlen, Anzahl der Beiträge in Diskussionsforen, Zeit und Anzahl von Downloads oder Länge der Bearbeitungszeiten. Diese Daten sagen aber ggf. wenig aus über Lernprozesse, die für das Lernziel wichtig sind. Eine Überinterpretation solcher Daten, die sich nicht am Lernziel orientieren, kann leicht zu Fehleinschätzungen über den aktuellen und zukünftigen Lernfortschritt und damit zu Fehlentscheidungen führen. Lernprozessbegleitende Diagnostik im schulischen Kontext muss auf den jeweiligen Lerngegenstand bezogen sein und zum Lernziel passen. Solche Lernenden-Daten müssen auf der Basis großer, für die Altersgruppe repräsentativer Daten mit Lernergebnissen in Zusammenhang gebracht werden, um valide lernzieldienliche Lernprozesse zu erfassen und Schlussfolgerungen zu erlauben. Erste Forschungsprojekte (z. B. das *HILA-Programm*, Drachsler, 2023) entwickeln die Erfassung von Lernprozessdaten, die spezifisch für das Lernziel sind, bisher gibt es aber noch keine breite Anwendung.

1.1.4 Daten zu überfachlichen Konstrukten

Datengestützte Unterrichtsentwicklung bedeutet auch, Daten zu überfachlichen Kompetenzen, insbesondere sozial-emotionalen Merkmalen in den Blick zu nehmen. Überfachliche Kompetenzen sind sowohl in den Bildungsplänen der Länder als auch in den Standards für die Lehrerbildung als Unterrichtsziele beschrieben. Die Verortung in den Lehr- und Bildungsplänen sowie die konkrete Operationalisierung unterscheiden sich allerdings zwischen den Ländern. Während überfachliche Kompetenzen in einigen Ländern eher fächerübergreifend in Kernlehrplänen, Rahmenlehrplänen oder Schulentwicklungsrahmen verortet sind (z. B. *Leitperspektive Prävention und Gesundheitsförderung* in Baden-Württemberg), sind sie in anderen Ländern stärker fachgebunden (z. B. Lernstrategien in Sprachen) formuliert. Beispielsweise werden in den *Bildungs- und Erziehungsgrundsätzen für die allgemeinbildenden Schulen in NRW* soziale und motivationale Haltungen und Einstellungen als Aspekte überfachlicher Bildung benannt, wobei dies „Selbstkompetenz [...], Eigenschaften wie Selbstständigkeit, Kritikfähigkeit, Selbstvertrauen, Zuverlässigkeit, Verantwortungs- und Pflichtbewusstsein“ umfasst (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2024, S. 14). Der *Hamburger Bildungsplan* für Grundschulen, Stadtteilschulen und Gymnasien benennt als konkrete Facetten überfachlicher Kompetenzen „das Vertrauen in die eigene Wirksamkeit, die Fähigkeit zur Selbstreflexion, Leistungsbereitschaft, Engagement, Motivation, Konzentrationsfähigkeit und Ausdauer“ sowie bezogen auf die methodischen und sozialen Kompetenzen „insbesondere Kommunikationsfähigkeit, Kreativität, kritisches Denken und Problemlösefähigkeit sowie Kooperation“ (Behörde für Schule und Berufsbildung Hamburg, 2022, S. 17–18).



Carstensen und Klusmann (i. E.) haben die Bildungspläne der Länder hinsichtlich überfachlicher Kompetenzen analysiert. Häufig wird auf die Trias Personal-, Sozial- und Methodenkompetenzen Bezug genommen. Überfachliche Kompetenzen werden demnach breit beschrieben. Sie umfassen sowohl Überzeugungen und Einstellungen (wie z. B. Vertrauen in die eigene Wirksamkeit oder Anstrengungsbereitschaft) als auch Fähigkeiten und Fertigkeiten (wie z. B. Kommunikationsfähigkeit, kritisches Denken, Kreativität oder Kooperationsfähigkeit). Die Vermittlung dieser Überzeugungen und Fähigkeiten wird als eigenständiges Ziel schulischer Bildung sowie als Grundlage erfolgreichen Lernens in allen Jahrgangsstufen beschrieben.

Die Stellungnahme der Leopoldina zur Förderung der Selbstregulationskompetenzen von Kindern und Jugendlichen (Nationale Akademie der Wissenschaften Leopoldina, 2024, S. 70) weist darauf hin, dass die Bildungs- und Lehrpläne der Länder Hinweise auf die Förderung von Selbstregulationskompetenzen von Kindern und Jugendlichen enthalten, oft jedoch relativ unbestimmt sind. Die Leopoldina empfiehlt, die Förderung von Selbstregulation als Leitperspektive in die Bildungs- und Lehrpläne aufzunehmen. Die SWK hat Empfehlungen zur systematischen Förderung von Selbstregulation, Wohlbefinden und sozialer Einbindung in der Grundschule vorgelegt (SWK, 2022a) und für die Sekundarstufe I spezifisch die systematische Förderung des selbstregulierten Lernens, der Identitätsentwicklung und der Berufsorientierung adressiert (SWK, 2025).

In den *Standards für die Lehrerbildung* der KMK (2022) sind überfachliche Kompetenzen insbesondere im Kompetenzbereich *Erziehen* verankert, teils auch im Kompetenzbereich *Unterrichten* (vor allem bezüglich Lern- und Selbstmotivationsstrategien). Lehrkräfte sollen die sozialen, kulturellen und technologischen Lebensbedingungen von Schüler:innen kennen und im Rahmen der Schule Einfluss auf die individuelle Entwicklung der Schüler:innen nehmen, sie sollen Werte, Normen und eine wertschätzende Haltung zu Diversität vermitteln, selbstbestimmtes und reflektiertes Urteilen und Handeln von Schüler:innen unterstützen sowie alters- und entwicklungspsychologisch adäquate Lösungsansätze für Schwierigkeiten und Konflikte in Schule und Unterricht finden.

Sozial-emotionale Merkmale wie soziale Eingebundenheit sowie motivationale Merkmale werden bislang trotz ihrer Verankerung in Schulgesetzen bzw. Bildungsplänen meist gar nicht oder nicht systematisch erfasst. In der SWK-Länderabfrage Schule haben drei Länder angegeben, Daten zum Schulklima sowie zum Wohlbefinden durch Beobachtungen bzw. Befragungen zu erheben. In einzelnen Ländern gibt es hierfür Module in den Portalen für schulinterne Evaluation (z. B. *SEP*). In einem Land erfolgt eine Rückmeldung unter Einbindung regionaler Beratungs- und Unterstützungszentren. Motivation, Lernverhalten und Selbstregulation werden in zwei Ländern durch Beobachtungen und Befragungen erfasst. Die Rückmeldung an Schulleitungen, Lehrkräfte, Eltern und Schüler:innen erfolgt schriftlich bzw. im Rahmen von (Lernentwicklungs-)Gesprächen. Ein Land berichtet, die Daten würden in der Regel von Lehrkräften zur Individualdiagnostik genutzt und seien nicht Bestandteil von Qualitätsentwicklungsgesprächen auf Schulebene. Baden-Württemberg hat im Januar 2026 zentrale Erhebungen zur Schul- und Unterrichtsqualität für die öffentlichen allgemeinbildenden und beruflichen Schulen eingeführt, in denen u. a. auch sozial-emotionale



Merkmale erfasst werden.¹⁰ Handreichungen oder Qualifizierungsangebote zur Struktur und zur Förderung überfachlicher Kompetenzen sind in den Ländern bislang nur vereinzelt zu finden.

1.1.5 Instrumente zur Erfassung der Unterrichtsqualität

Eine wichtige Gelingensbedingung für das Erreichen fachlicher und überfachlicher Kompetenzziele ist die Unterrichtsqualität (Klieme, 2018; Praetorius et al., 2018). Für kollegiale Prozesse der Unterrichtsentwicklung sind daher nicht nur Daten über die Lernenden, sondern auch über die Unterrichtsqualität hoch relevant. Sie werden durch unterschiedliche Verfahren erfasst, die von informellen Hospitationen bis hin zu standardisierten Beobachtungen und Schülerbefragungen reichen. Die Fokussierung der Beurteilung auf valide Konstrukte bzw. die Nutzung valider Instrumente ist eine Voraussetzung für die Wirksamkeit der Evaluation von Unterricht (vom Hofe et al., 2002).

In der SWK-Länderabfrage Schule haben 14 Länder angegeben, dass Schulen verpflichtet sind, schulinterne Evaluationen zur Unterrichts- und Schulqualität regelmäßig durchzuführen. In zwei Ländern besteht keine solche Verpflichtung (vgl. Kapitel 2). In keinem Land gibt es verbindliche Vorgaben dazu, welche Instrumente zur Erfassung von Schul- und Unterrichtsqualität eingesetzt werden sollen, aber die Landesinstitute bzw. Qualitätseinrichtungen der Länder stellen Schulen entsprechende Instrumente zur Verfügung. Die Angebotsstruktur ist hierbei eher unübersichtlich und die Validität der Instrumente oft nicht hinreichend untersucht. Einen eigenen Weg geht Baden-Württemberg: Hier werden Daten zur Unterrichtsqualität sowie zu Wohlbefinden über zentrale, derzeit freiwillige Erhebungen erfasst und interessierten Schulen zurückgemeldet. Einen Überblick ohne Anspruch auf Vollständigkeit über die in den Ländern eingesetzten, teils selbst entwickelten Instrumente zur Erhebung von Unterrichtsqualität im Rahmen interner Evaluationen bietet die Tabelle in Anhang 3.

Wer welche Rückmeldungen zu Merkmalen der Unterrichtsqualität erhält, unterscheidet sich zwischen den Ländern. Instrumente zur Erfassung von Schülerfeedback und zur Erfassung von Einschätzungen externer Beobachter:innen werden in den meisten Ländern mittlerweile digital bereitgestellt. Die Rückmeldungen gehen hier durchgehend (digital aufbereitet) an die Lehrkräfte selbst, die auch (freiwillig) die Erhebungen durchführen und organisieren. Selbstberichte von Lehrkräften werden meist in Kombination mit Schülerfeedback erhoben, um einen Perspektivenvergleich zu ermöglichen. Beobachtungsinstrumente werden mit Abstand am häufigsten im Rahmen der *externen Evaluation* an Schulen (auch *Schulinspektion* oder *Qualitätsanalyse*) genutzt (vgl. Kapitel 3). Hier erhalten die Schulen einen Evaluationsbericht auf der Schulebene, wobei zum Teil auch die Verteilung innerhalb der Schule zurückgemeldet wird. Individuelles Feedback an die einzelnen Lehrpersonen ist im Rahmen der Schulevaluation kaum vorgesehen.

In drei Ländern werden darüber hinaus Beobachtungsinstrumente angeboten, die explizit für individuelles Feedback an Lehrkräfte vorgesehen sind. Die Rahmenbedingungen, unter denen ein formatives Feedback für Lehrkräfte aus Unterrichtsbeobachtungen gegeben wird, sind dabei heterogen und meist wenig verbindlich, die Anschlussprozesse zur Verbesserung der Unterrichtsqualität bislang wenig strukturiert.

¹⁰ <https://www.baden-wuerttemberg.de/de/service/gesetze-und-verordnungen/gesetzblatt/detail/2026-3> (aufgerufen am 20.02.2026).



1.1.6 Darstellung der Daten in Form von Dashboards

Dashboards sind visuelle Anzeigen von Daten mit dem Ziel, alle wesentlichen Informationen den Nutzenden möglichst übersichtlich rückzumelden. Sie können für die Rückmeldung der Ergebnisse zentraler Lernstandserhebungen und lernprozessbegleitender Diagnoseinstrumente sowie im Rahmen von Learning Analytics für die Überwachung der Aktivitäten der Lernenden in digitalen Lernumgebungen mit dem Ziel einer schnellen Unterstützung gestaltet werden (Koh & Tan, 2017).

In der SWK-Länderabfrage Schule haben auf die spezifische Frage nach Rückmeldeformaten zehn Länder angegeben, Dashboards für Rückmeldungen zu nutzen. Die Einsatz- und Nutzungszwecke sowie die Zielgruppen, die Rückmeldung über Dashboards erhalten, variieren hierbei stark. Einzelne Länder nutzen Dashboards für Rückmeldungen zu Fehlzeiten, Schulklima und Wohlbefinden. Vor allem werden Dashboards aber für Rückmeldungen aus Lernstandserhebungen oder lernprozessbegleitenden Diagnoseinstrumenten genutzt, etwa Rückmeldungen aus VERA, ILeA bzw. ILeA plus, quop, bettermarks oder Mathe sicher können. Mit Blick auf die unterschiedlichen Zielgruppen geben neun Länder an, dass Rückmeldungen über Dashboards an Lehrkräfte erfolgen, je sechs Länder nutzen Dashboards für Rückmeldungen an Schulleitungen (vgl. Kapitel 2), fünf Länder für Rückmeldungen an Schüler:innen sowie zwei Länder für Rückmeldungen an weiteres pädagogisches Personal. Jeweils ein Land gibt an, Dashboards für Rückmeldungen an Eltern, Schulaufsicht oder Berater:innen aus dem Unterstützungssystem zu nutzen.

1.2 Konzepte und Maßnahmen

1.2.1 Effekte und Bedingungen der Nutzung zentraler Lernstandserhebungen für Unterrichtsentwicklung

Ergebnisse aus Lernstandserhebungen können wertvolle Hinweise für die unterrichtliche Weiterarbeit liefern. Ihre Nutzung erfordert einen mehrschrittigen Prozesskreislauf der Rezeption, Reflexion und darauf bezogenen Maßnahmenplanung und -evaluation, wobei weitere Daten (z. B. aus Unterrichtsbeobachtungen oder Schülerbefragungen) zu berücksichtigen sind (Helmke & Hosenfeld, 2005; Mandinach & Schildkamp, 2021; Visscher, 2021). Dies ist ein anspruchsvoller Bestandteil professionellen Handelns von Schulleitungen und Lehrkräften, der gezielt unterstützt werden muss.

International zeigt sich, dass die Nutzung der Ergebnisse zentraler Lernstandserhebungen für die Unterrichtsentwicklung dann wirksam wird, wenn sich eine kollaborative und reflexive Auseinandersetzung anschließt (Visscher, 2021). Zentrale Bedingungen zielführender Nutzung lassen sich drei übergeordneten Merkmalsbereichen zuordnen (Schildkamp et al., 2017):

1. Individuelle Merkmale der Datennutzenden beziehen sich auf persönliche Voraussetzungen und professionelle Dispositionen, die die effektive Arbeit mit Daten unterstützen. Professionelle Handlungskompetenzen umfassen insbesondere eine breite Wissensbasis (Jihyun Lee et al., 2024) sowie positive Einstellungen und motivationale Orientierungen gegenüber Datennutzung (z. B. Selbstwirksamkeit, Interesse und Freude an der Arbeit mit Daten, wahrgenommene Nützlichkeit und Wichtigkeit der Erhebungen; Datnow & Hubbard, 2016; Hoogland et al., 2016).



2. Schulische Merkmale beziehen sich auf die Qualität innerschulischer Strukturen, Prozesse und kollektiver Praktiken, die die Nutzung von Daten aus Lernstandserhebungen ermöglichen und rahmen. Dazu zählen unterstützendes Schulleitungshandeln (Schildkamp et al., 2019), gemeinsam getragene Organisations- und Kooperationspraktiken (z. B. feste Zeiten zur Datennutzung) und eine positive Datennutzungs-/Feedbackkultur im Kollegium (Datnow & Hubbard, 2016; Hawlitschek et al., 2024; Keuning et al., 2017).
3. Merkmale der Rückmeldesysteme beziehen sich auf die Qualität und Kohärenz schulexterner Rahmenbedingungen, die das datengestützte Arbeiten von Schulen unterstützen. Hierzu zählen vor allem die niedrigschwellige und zeitnahe Verfügbarkeit der Daten, die adressatengerechte Aufbereitung der Daten in Ergebnismeldungen (z. B. inhaltliche Interpretationshilfen) sowie die wissenschaftliche Fundierung und Qualität der Testkonstruktion (z. B. curriculare Validität; Schildkamp et al., 2017; Visscher, 2021).

Zur Bedeutung der genannten Bereiche liegen international zahlreiche empirische Studien vor (für einen Überblick siehe Hoogland et al., 2016; Visscher, 2021). In Deutschland wurden Studien zum Umgang mit VERA-Daten durchgeführt, die in eine ähnliche Richtung weisen und die Bedeutsamkeit individueller, schulischer und schulexterner Bedingungen unterstreichen (für einen Überblick siehe Altrichter & Maag Merki, 2016; Demski, 2021; Hawlitschek et al., 2025). Punktuell finden sich für Deutschland spezifische Befunde, etwa zu Unterschieden zwischen Schulformen (nicht-gymnasial vs. gymnasial), Ländern, Unterrichtsfächern (Deutsch vs. Mathematik), Akteursgruppen (Schulleitungen vs. Lehrkräfte) sowie zu zeitlichen Trends (z. B. Demski, 2021; Groß Ophoff et al., 2019; Hawlitschek et al., 2025; U. Maier, 2008; Wurster & Richter, 2016). Insgesamt dominiert dabei die Analyse der Akzeptanz und wahrgenommenen Nützlichkeit der VERA-Tests (exemplarisch: Detzel, 2024).

Zur Professionalisierung von Schulleitungen und Lehrkräften im Umgang mit VERA oder anderen Lernstandserhebungen in Deutschland liegen hingegen bislang kaum Befunde vor. Internationale Metaanalysen deuten darauf hin, dass die Teilnahme an Fortbildungen zu datengestützter Unterrichtsentwicklung positive Effekte auf das Wissen und die Motivation von Lehrkräften hat (Filderman et al., 2022; Gesel et al., 2021). Für den deutschsprachigen Raum fehlten hierzu lange Zeit entsprechende Befunde (Ausnahmen: Koch, 2011; Vogel et al., 2016). Neuere, von Bund und Ländern geförderte Projekte führen erstmals umfassendere, (quasi-)experimentelle Evaluationen von Fortbildungen zu VERA durch (Jost, 2025) und nutzen hierfür *Blended-Learning*-Ansätze (Kombination aus E-Learning, Workshops und ggf. begleitendem Coaching). Dazu gehören *VERA-BiSS* im Rahmen der Initiative *Bildung durch Sprache und Schrift* (Schnitzler et al., 2024) sowie ein Teilprojekt des Forschungsverbunds *Schule macht stark – SchuMaS* (Hawlitschek & Henschel, 2025). Weitere Projekte, die Unterstützungsangebote zu VERA oder zum Umgang mit Daten untersuchen, sind etwa *VERAcheck* (Gärtner, 2025), Lehrveranstaltungen für Studierende (Wurster et al., 2023) oder die Studie *WeSu* (Hosenfeld et al., 2024).

Es ist anzunehmen, dass evidenzbasierte Fortbildungen zum Datennutzungszyklus sowie zu den fachdidaktischen Hintergründen der jeweiligen Kompetenzbereiche die fachdidaktisch treffsichere Nutzung von Daten aus Lernstandserhebungen in der Unterrichtsgestaltung deutlich erhöhen. Eine Voraussetzung dafür ist aber, dass die Nutzung der Daten in ein kohärentes System datengestützter Qualitätsentwicklung eingebettet ist (Demski, 2017). Ein Beispiel hierfür ist Hamburg, wo die VERA-Tests in ein System der



Kompetenzermittlung und -förderung integriert sind, das weitere Verfahren und Datenquellen umfasst (Hahn et al., 2025; Schulte et al., 2016). Erfasst werden Kompetenzen in den Fächern Deutsch und Mathematik sowie ab Jahrgang 5 in Englisch und den Naturwissenschaften sowie ggf. Französisch in Jahrgang 8. Die insgesamt sechs Testzeitpunkte ermöglichen auch eine längsschnittliche Beobachtung von Lernfortschritten zwischen zwei oder drei Schuljahren.

Die Aussagekraft von Lernstandserhebungen kann durch eine solche längsschnittliche Betrachtung von Ergebnissen erhöht werden. Die Rückmeldungen enthalten dann nicht nur Informationen über den zu einem Zeitpunkt erreichten Leistungsstand der Schüler:innen, sondern auch über den Lernfortschritt über die Zeit. Nur so können Lehrkräfte und Schulen feststellen, ob die Maßnahmen, die sie aufgrund festgestellter Problembereiche ergriffen haben, die erwünschten Effekte hatten, z. B. ob der Anteil der Schüler:innen, die in VERA 3 den Mindeststandard in Mathematik verfehlt haben, reduziert werden konnte. Im Idealfall sind bei Längsschnitten Aussagen darüber möglich, wie sich die Kompetenzen der einzelnen Schüler:innen verändert haben. Allerdings sind die Bestimmung und Darstellung von Lernfortschritten voraussetzungsreich und komplex (vgl. z. B. Betebenner & Linn, 2009; von Eichhorn & Freunberger, 2024). Die Messungen müssen sich auf Konstrukte beziehen, die sich zwischen den Messzeitpunkten nicht grundlegend verändern, und sie müssen ausreichend präzise sein, um aussagekräftige Ergebnisse zu liefern. Zur Ermittlung von Lernfortschritten wurden verschiedene Verfahren entwickelt, die je nach Verwendungszweck und Datenlage mit verschiedenen Vor- und Nachteilen verbunden sind. Entsprechend kommen national und international unterschiedliche Vorgehensweisen zur Anwendung (z. B. Data Quality Campaign, 2019). In Luxemburg z. B. werden alle zwei Jahre (Vorschule, 1., 3., 5., 7. und 9. Klasse) Kompetenzmessungen in den Landesprachen und im Fach Mathematik durchgeführt, auf die Darstellung von Veränderungen auf Individualebene wird aber verzichtet. Nur auf Systemebene wird berichtet, wie sich die Anteile der Schüler:innen verändert haben, die verschiedene Kompetenzstufen erreichen. In Hamburg hingegen wird im Rahmen von KERMIT auch auf Individual-, Klassen- und Schulebene zurückgemeldet, inwieweit sich die erreichten Kompetenzen jeweils in aufeinanderfolgenden Erhebungen – also jeweils zwischen zwei Schuljahren – verändert haben. Ebenso erhalten Lehrkräfte und Schüler:innen in Österreich jährlich eine Rückmeldung darüber, wie sich die in den Fächern Deutsch, Mathematik und (nur in der Sekundarstufe) Englisch erfassten Kompetenzen von der dritten zur vierten Jahrgangsstufe bzw. von der siebten zur achten Jahrgangsstufe entwickelt haben. Dabei wird der individuelle Lernfortschritt mit dem Lernfortschritt verglichen, der bei der jeweiligen Ausgangsleistung zu erwarten ist (*Residual Gain Model*), ob also ein Kind, das in der dritten Jahrgangsstufe einen bestimmten Wert z. B. im Lesen erreicht hat, bis zur vierten Klasse einen geringeren, ähnlich hohen oder höheren Kompetenzzuwachs erzielt hat als andere Schüler:innen, die in der dritten Jahrgangsstufe auf vergleichbarem Niveau lesen konnten. In den Rückmeldungen für die Schüler:innen wird anhand einer Zeichnung eines Kindes im Weitsprung anschaulich dargestellt, ob es im vergangenen Jahr einen „kleinen“, einen „mittleren“ oder einen „großen“ Sprung im jeweiligen Kompetenzbereich gemacht hat, und auch die Lehrkräfte erhalten diese Information. Das Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen (IQS) hat sein Vorgehen nach gründlicher Abwägung verschiedener Vorgehensweisen gewählt (von Eichhorn & Freunberger, 2024) und sammelt damit aktuell erste Erfahrungen, die bei der Weiterentwicklung der Testsysteme in Deutschland berücksichtigt werden sollten.



Längsschnittliche Betrachtungen schulischer Kompetenzmessungen sind zudem durch die Notwendigkeit der Verknüpfung individueller Ergebnisse voraussetzungsvoll. Sie erfordern, dass für jede Schülerin bzw. jeden Schüler eine anonymisierte ID generiert wird, die keine Rückschlüsse auf ihre jeweiligen Namen oder Herkunft zulässt, aber gleichzeitig bei wiederholten Erhebungen die Verknüpfung der Leistungstests ermöglicht. In Hamburg beispielsweise wird aus Informationen der amtlichen Schulstatistik datenschutzkonform eine KERMIT-ID für alle Schüler:innen generiert. Auf Basis dieser ID, die mit den Testergebnissen abgespeichert wird, können die Daten anschließend längsschnittlich verknüpft werden, auch wenn Kinder oder Jugendliche die Schule oder Schulform wechseln. Für Adjustierungen können zudem Verknüpfungen mit Daten der Schulstatistik nützlich sein, die aber mit zusätzlichen Herausforderungen verbunden sind (zur Schüler- bzw. Bildung-ID vgl. Kapitel 5.1.12).

Rückmeldungen von Ergebnissen aus Lernstandserhebungen wie VERA müssen immer vor dem Hintergrund der Bedingungen interpretiert werden, unter denen sie entstanden sind. Hierbei können adjustierte Vergleiche (siehe Kapitel 1.2.1) hilfreich sein. Auch wenn sich adjustierte Vergleiche nicht kausal als Effekte des Unterrichts interpretieren lassen,¹¹ da u. a. nicht alle potenziell relevanten Merkmale in die Analysen einbezogen werden können, bieten sie eine deskriptive Annäherung (Fiege, 2013; Fiege et al., 2011) und bilden somit eine wichtige Ergänzung zu den unadjustierten Vergleichen. So sollte es beispielsweise für eine Grundschule mit einer stark benachteiligten Schülerschaft aufschlussreich sein, zu erfahren, dass sie nicht nur im Vergleich zu Grundschulen des Landes insgesamt schwächere Leistungen erzielt, sondern auch im Vergleich zu Grundschulen mit Schüler:innen, die über ähnliche Lernvoraussetzungen verfügen.

Zur Frage, wie adjustierte Vergleiche angelegt werden sollten, liegen vielfältige Analysen vor, von denen sich einige direkt auf VERA beziehen und verschiedene Strategien vergleichen (Fiege, 2013; Schult, 2022; für internationale Analysen vgl. z. B. Emslander et al., 2022; Koedel et al., 2015). Kurz zusammengefasst weist der Forschungsstand darauf hin, dass bei adjustierten Vergleichen nach Möglichkeit Indikatoren des fachspezifischen Vorwissens einbezogen werden sollten. Im Idealfall sollte dieses bei VERA 3 zu Beginn der ersten Jahrgangsstufe (z. B. mit StarS) und bei VERA 8 direkt nach dem Übergang in die Sekundarstufe I (bzw. zu Beginn des Einsetzens von Fachunterricht auf Sekundarschulniveau) erfasst werden. So können adjustierte Vergleiche Hinweise darauf geben, inwieweit die Förderung von Schüler:innen, die zu Beginn der Bildungsetappe über ähnliche Lernausgangslagen verfügten, in der jeweiligen Schule ähnlich erfolgreich war wie in anderen Schulen oder ob Unterschiede bestehen, die auf Weiterentwicklungsbedarf hinweisen. Werden Merkmale des familiären Hintergrunds in die Adjustierung einbezogen, sollte es sich im Sinne des Alignments um diejenigen Merkmale handeln, die auch zur Bildung des Sozialindex im jeweiligen Land verwendet werden.

Insgesamt gelingt eine datengestützte und fachdidaktisch treffsichere Unterrichtsentwicklung dann, wenn die jeweiligen Instrumente und Daten auf klar definierte Ziele und die Wege dorthin bezogen sind und für

¹¹ Insbesondere in den USA wird *Value-added Modeling* äußerst kontrovers diskutiert. Allerdings bezieht sich diese Diskussion nicht auf die Nutzung dieses Ansatzes für datengestützte Qualitätsentwicklung, sondern auf *High-Stakes Accountability* für Lehrkräfte. Dabei wird davon ausgegangen, dass adjustierte Leistungswerte kausal als Effekte des Unterrichts der jeweiligen Lehrkraft interpretierbar sind und diese so genau gemessen werden können, dass sie sich als (eine) Grundlage für sehr weitreichende Entscheidungen eignen (vgl. z. B. Darling-Hammond, 2015; Koedel et al., 2015). Um diese Art der Nutzung geht es im vorliegenden Gutachten nicht.



die Entwicklung von Unterricht themenbezogene Fördermaterialien und Fortbildungsangebote zur Verfügung stehen. Auf Ebene der Schulen erfordert eine zielführende Datennutzung zudem eine Leitungs- und Teamkultur, insbesondere in Fachteams, die offene und kollegiale Austauschprozesse und institutionalisierte Datenreflexion ermöglichen. Erforderlich ist ferner eine ausreichende Datenkompetenz der Fachteams, die in geeigneten Fortbildungen erworben werden kann. Für die systematische Auseinandersetzung mit den Daten und die Ableitung von Maßnahmen zur Unterrichtsgestaltung im Kollegium müssen schulinterne Arbeitsstrukturen und Zeitkorridore zur Verfügung stehen (z. B. regelmäßige Datenworkshops, Steuergruppen). Bei der Gestaltung dieser Prozesse und der Entwicklung einer konstruktiven und produktiven Feedbackkultur kommt den Schulleitungen und Fachgruppenleitungen eine zentrale Rolle zu.

1.2.2 Effekte und Bedingungen der Nutzung diagnosebasierter Förderung mit lernprozessbegleitenden Diagnoseinstrumenten

Anders als zentrale Kompetenztests dienen *lernprozessbegleitende Diagnosen und Tests* dazu, den Unterricht in kürzeren Zeithorizonten zu planen und weiterzuentwickeln (Black & Wiliam, 1998). Sie werden von den Lehrkräften eigenverantwortlich eingesetzt und für die diagnosegeleitete Lernförderung genutzt (Black & Wiliam, 2009; Ikemoto & Marsh, 2007; Karst et al., 2022, für Lesen; Prediger et al., 2025, für Mathematik; Selter, 2017) Dabei können neben standardisierten Tests auch informelle Ansätze zur Datenerhebung (z. B. reichhaltige Erkundungsaufgaben) oder fachdidaktisch valide Beobachtungsraster verwendet werden (Sundermann & Selter, 2006). Es werden kleinschrittige Ziele entlang fachdidaktisch fundierter Lernprogressionen in den Blick genommen, z. B. die Zone der nächsten Entwicklung im Schriftspracherwerb. Durch Abgleich von Kompetenzziele und -ständen werden nächste Lernziele ausgeschärft, individuell adaptive Fördermaßnahmen abgeleitet und durchgeführt (Confrey & Maloney, 2015; Prediger & Karst, i. E.). Die Wirkung der Fördermaßnahmen kann durch erneute Datenerhebung überprüft und Maßnahmen angepasst werden.

Lernprozessbegleitende Diagnostik bzw. formatives Assessment mit lernprozessbegleitenden Diagnoseinstrumenten ist im Unterricht noch wenig verbreitet (Fuchs et al., 2025). Vorwiegend wird nach wie vor anhand von Beobachtungen von Einzel- und Gruppenarbeiten, Klassengesprächen und Klassenarbeiten diagnostiziert. Reviews und Metaanalysen experimenteller und quasi-experimenteller Interventionsstudien zu Effekten formativen Assessments weisen aber darauf hin, dass lernprozessbegleitende Diagnostik wirksam ist, allerdings mit unterschiedlichen Effektstärken. Während einige Autor:innen mittlere bis große Effekte formativer Assessments berichten (Graham et al., 2015; Karaman, 2021), zeigen andere Analysen nur kleine Effekte (Kingston & Nash, 2011; H. Lee et al., 2020; Xuan et al., 2022; Yao et al., 2024). Die Effekte unterscheiden sich hinsichtlich der untersuchten Fächer, des Alters der Schüler:innen und der Art der Bereitstellung des Feedbacks (automatisiert vs. durch die Lehrkraft). Differenzierte Analysen zeigen, dass die Effekte formativer Assessments deutlich größer sind, wenn diese mit instruktionaler Unterstützung verknüpft sind (H. Lee et al., 2020; Xuan et al., 2022), wenn klare Guidelines und eine explizite Struktur der Implementierung kommuniziert werden (Aust et al., 2025) und wenn die Implementation durch eine fachdidaktische Professionalisierung der Lehrkräfte unterstützt wird (Kingston & Nash, 2011).



Vor dem Hintergrund dieser Befundlage lassen sich für *diagnosebasierte Förderung mit lernprozessbegleitenden Diagnoseinstrumenten* folgende fachdidaktische Gelingensbedingungen identifizieren, die die nachhaltige Lernwirksamkeit steigern (Black & Wiliam, 2009; Moser Opitz, 2022; Prediger & Karst, i. E.; Visscher, 2021).

Verfügbarkeit geeigneter lernprozessbegleitender Diagnoseinstrumente

Ergänzend zu zentralen Lernstandserhebungen wie VERA, die breitere Kompetenzbereiche abdecken, können lernprozessbegleitende Instrumente fokussierter Stärken und Schwächen von Lernenden zu ausgewählten Themengebieten erfassen (Burkhardt & Schoenfeld, 2018; Selter, 2017). Solche lernprozessbegleitenden Diagnoseinstrumente liegen allerdings erst für wenige Bereiche vor (SWK, 2022a, 2025; vgl. Kapitel 1.1.2). Zudem sind die vorhandenen Instrumente häufig nicht auf den Kern relevanter Kompetenzen bezogen. So erfassen etwa digitale Tools zu mathematischen Basiskompetenzen fast ausschließlich prozedurale Fertigkeiten (Thurm & Graewert, 2022), obwohl konzeptuelles Verständnis zum Weiterlernen ebenso relevant ist (Gersten et al., 2009). Die nur partiell fachdidaktisch treffsichere Auswahl wirkt sich auf Schwerpunktsetzungen des Unterrichts über Backwash-Effekte aus (Sabio et al., 2015), d. h., Lehrkräfte priorisieren die in Tests abgefragten Kompetenzen (*What you test is what you get*, Burkhardt & Schoenfeld, 2018). Durch fachdidaktisch treffsichere lernprozessbegleitende Instrumente kann dagegen eine Fokussierung von Lehrkräften auf relevante Kompetenzen gelingen (Stacey et al., 2018). Für diagnosegeleitete Lernförderung im Unterricht sollten die Instrumente zudem Lernfortschritte sichtbar machen (Black & Wiliam, 2009), wozu diese je nach Kompetenzstrukturen des Faches unterschiedlich angelegt sein müssen. Während z. B. im Deutschunterricht für die Leseflüssigkeit zur Lernverlaufsdagnostik dieselben Instrumente mehrfach eingesetzt werden (Förster & Souvignier, 2014), eignen sich im Mathematikunterricht eher kleine Vor- und Nachtests, die längsschnittlich curricular verknüpft werden, um den jeweiligen Lernprogressionen fachdidaktisch gerecht zu werden (Confrey & Maloney, 2015; Stacey et al., 2018).

Verknüpfung lernprozessbegleitender Diagnoseinstrumente mit fachdidaktisch fundiertem Fördermaterial

In vielen Kontexten werden Diagnoseinstrumente angeboten in der Annahme, Lehrkräfte könnten eigenständig dazu passende Förderansätze bzw. passendes Fördermaterial auffinden. Reviewstudien weisen jedoch darauf hin, dass dies häufig nicht gelingt: „teachers do find it difficult to translate student progress data into tailor-made instruction“ (Visscher, 2021, S. 1). Eine Metaanalyse zu Lernverlaufsdagnostiken in Deutsch und Mathematik zeigt zudem, dass höhere Wirksamkeit erzielt wird, wenn Lehrkräfte auch in Bezug auf die Förderung unterstützt werden (Fuchs et al., 2025). Dabei ist ein mehrschrittiger Aufbau der Teilkompetenzen entlang einer fachdidaktisch fundierten Lernprogression erforderlich. Ein systematischer Kompetenzaufbau wird ermöglicht durch iterativ entwickelte Fördermaterialien entlang forschungsfundierter Lernprogressionen (Confrey & Maloney, 2015). Solche Fördermaterialien sollten in Zukunft für alle relevanten Themenbereiche gemeinsam mit diagnostischen Tools entwickelt und bereitgestellt werden (SWK, 2022a, 2025).

Ausreichende Fortbildung zu Förderkonzepten

Reine Materialunterstützung reicht selten aus, damit sich Lehrkräfte in neue Förderkonzepte effektiv einbringen. Wenn sie dagegen mit hochwertigen, forschungsbasierten Fortbildungen kombiniert wird, kann nachweislich Wirksamkeit für die Lehrkräfteprofessionalität und den Kompetenzaufbau der Lernenden



erzielt werden (Lynch et al., 2019). Langfristige Begleitung durch qualifiziertes Personal ermöglicht zudem eine nachhaltige schulische Implementation (Karst et al., 2024). Dies umfasst die Arbeit von Fachlehrkräften in langfristig angelegten Fortbildungs-Fachnetzwerken (Lynch et al., 2019) ebenso wie gezieltes individuelles fachdidaktisches Unterrichtscoaching für Lehrkräfte, deren Klassen nicht genügend Lernfortschritte erzielen.

1.2.3 Effekte und Bedingungen der Nutzung von Learning Analytics für Lernförderung

Learning Analytics werden im schulischen Kontext momentan in Deutschland weder systematisch noch in der Breite genutzt. Mit dem Einsatz von KI in der Schule ist die Vision verbunden, dass durch die neuen technischen Möglichkeiten individuelles Feedback und Adaptivität in viel mehr Lernbereichen, für komplexeres Material und viel schneller, quasi *on the fly*, ermöglicht wird. Internationale Metaanalysen zu unterschiedlichen Learning-Analytics-Methoden, Interventionen, Anwendungskontexten und Inhaltsbereichen zeigen für kognitive Maße (z. B. Lernergebnisse) überwiegend kleine bis moderate Effekte (Kaliisa et al., 2024; Kokoç et al., 2025; Yan Liu et al., 2025), für motivationale Maße (z. B. selbstregulatives Lernen, Beteiligung) positive, aber variierende Effekte (Heikkinen et al., 2023; Kaliisa et al., 2024; Yu & Zhang, 2023) und für lernbezogene Einstellungen kleine bis vernachlässigbare Effekte (Kaliisa et al., 2024). Es dominieren allerdings Studien aus dem Hochschulbereich, den MINT-Fächern und Studien mit quasi-experimentellen Designs (Kaliisa et al., 2024). Damit gibt es zur Frage, inwiefern Learning Analytics wirksam sind, noch zu wenig systematische Evidenz (Kaliisa et al., 2024; Larsen et al., 2025; Sousa et al., 2021), schon gar nicht für die Schulpraxis in Deutschland. Ausgehend vom empirisch gesicherten hohen Potenzial von Feedback und Adaptivität lassen sich aber positive Effekte dann erwarten, wenn Lernangebote, Feedback und Adaptivität fachdidaktisch treffsicher gestaltet sind (Confrey & Maloney, 2015).

Learning Analytics können also ihr Potenzial nur erreichen, wenn sie hohen didaktischen Anforderungen entsprechen und eine Passung zwischen Lernziel, Lernmethoden, Diagnostik und Förderung besteht (*Constructive Alignment*). Diese Passung fehlt, wenn nur Maße, die durch digitale Medien leicht erfassbar sind (z. B. Klickzahlen), gemessen und zur Diagnostik des Lernprozesses genutzt werden. Stattdessen müssen Lernziele der Ausgangspunkt aller Überlegungen zur lernprozessbegleitenden Diagnostik auch in den digitalen Systemen sein. Damit die gewünschte Orientierung an Lerngegenständen erfolgen kann, ist eine verbindliche Konkretisierung der Curricula erforderlich, wie es die SWK regelmäßig empfiehlt (SWK, 2025). Learning Analytics sollten klar mit den curricularen Lernzielen verknüpft sein und relevante Lernprozesse (z. B. tiefe Verarbeitung der Inhalte) reliabel und valide erfassen. Die Kriterien für geeignete Daten sollten (wie bei allen diagnostischen Instrumenten) theoretische Fundierung, curriculare Validität, Konstruktvalidität, Testökologie, Adaptivität des Verfahrens, Gebrauchstauglichkeit, Testfairness und Testökonomie (Junger & Liebers, 2024) sein. Um die Kriterien bei neuen Anwendungen z. B. mit *Large Language Models (LLM)* zu gewährleisten, sind für das Training von KI große Datenmengen aus realen Anwendungen nötig. Nur authentische domänen- und altersspezifische Daten einer repräsentativen Stichprobe, die mit lernzieldienlichen Prozessen und Lernzielen verknüpft sind, können als *Goldstandard* für das Training einer KI verwendet werden. Das ist die Voraussetzung für eine valide Diagnose der Lernprozesse in Echtzeit und eine entsprechende fachdidaktisch fundierte Förderung und Adaptation. Außerdem setzt die Entwicklung von Learning Analytics Anschluss an Forschung zu Kompetenzmodellen, Erwerbsmodellen und



Förderkonzepten voraus. Dazu sind interdisziplinäre Kooperationen unentbehrlich. Nach der Entwicklung und Implementierung von Learning Analytics sind Wirksamkeitsstudien nötig. Langfristig sollten auch die gesellschaftlich viel diskutierten möglichen negativen Nebeneffekte auf Metakompetenzen – wie etwa der Verlust der Fähigkeit, selbstständig Lernprozesse zu beurteilen und zu organisieren, oder der Fähigkeit, Probleme selbstständig zu lösen, Frustration zu ertragen oder in Lernumgebungen zurechtzukommen, die nicht perfekt adaptiert sind – erfasst werden, um ggf. langfristiges *Deskilling* (d. h. den Verlust von Kompetenzen durch fehlende Übung) durch den Gebrauch von Learning Analytics zu erforschen.

Neben den didaktischen Bedingungen gibt es praktische Bedingungen für die Effizienz des Einsatzes von Learning Analytics: Nötig ist eine einheitliche digitale Infrastruktur, z. B. eine digitale Lernplattform mit Schnittstellen, die Daten aus unterschiedlichen Quellen zusammenführt (Single Sign-on; technische Interoperabilität). Die Rückmeldungen der Learning-Analytics-Systeme müssen außerdem so gestaltet sein, dass Lehrkräfte sie für konkrete pädagogische Entscheidungen nutzen können (Koh & Tan, 2017). Lehrkräfte müssen also zum einen die in den Dashboards rückgemeldeten Daten verstehen (d. h., die Daten dürfen nicht nur gesammelt, sondern müssen sinnvoll strukturiert und ausgewertet werden) und zum anderen in der Lage sein, entsprechende Maßnahmen auszuwählen zu können (Gross et al., 2025). Erst wenn Lernziele, Dateninterpretation und didaktische Handlung konsequent miteinander verknüpft werden, entfalten Learning-Analytics-Systeme ihr volles Potenzial (Drachler, 2023). Die Einbettung von Learning Analytics in ein didaktisches Gesamtkonzept ist daher zentral. Lehrkräfte müssen spürbar entlastet und durch hochwertige fachdidaktische Angebote für den Unterricht unterstützt werden. Dazu sollten Fachdidaktiken und Lehrkräfte gezielt von Anfang an in die Planung der Learning Analytics einbezogen werden, um didaktische Einbettung und Praxistauglichkeit zu gewährleisten und die Akzeptanz bei der Einführung zu erhöhen.

Für einen treffsicheren Einsatz der digitalen Lehr-Lern-Umgebungen mit Learning Analytics brauchen Lehrkräfte spezifische Kompetenzen, damit sie Informationen aus Learning Analytics wirksam nutzen können: dazu gehören die Fähigkeit zur Identifikation individueller Lernstände und potenzieller Hürden, *Feedback Literacy* und die Fähigkeit, die Erkenntnisse fachdidaktisch klug in pädagogisches Handeln umzusetzen. Lehrkräfte müssen entsprechend qualifiziert werden.

Die tatsächlich lernzieldienliche Nutzung von Learning Analytics hängt vermutlich – wie in anderen Bereichen auch – stark von der Akzeptanz und den Einstellungen der Lehrkräfte, Schüler:innen, Elternschaft und Leitungspersonen ab. Eine Implementation von Learning Analytics in der Praxis sollte durch eine gute Kommunikationsstrategie begleitet sein und alle Ebenen der Schule beteiligen. Ebenso stellen fehlende Infrastruktur und mangelnder Support, fehlende *Digital Literacy* der Beteiligten sowie fehlende Motivation und Selbstregulationskompetenz Barrieren dar, die ggf. dazu führen können, dass bestimmte Lehrpersonen und Lernende weniger oder gar nicht von der Implementation von Learning Analytics profitieren und ggf. den Anschluss verpassen. Im besten Fall verhelfen Learning Analytics zu mehr Bildungsgerechtigkeit, im schlechtesten Fall potenzieren sie die Ungleichheiten im System, wenn die Nützlichkeit von Learning Analytics an Ausgangsbedingungen geknüpft ist, die strukturell ungleich verteilt sind. Bei der Entwicklung sollte deshalb sichergestellt werden, dass sich die Bildungsschere nicht erweitert: Lernende mit schwierigeren Ausgangsbedingungen (z. B. geringe Digital Literacy, schwächere Leistungen in der Lerndomäne, geringe Selbstregulationskompetenz, geringe Identifikation mit dem Lernbereich oder mit Bildung) und



Lernende mit Risikofaktoren im Bildungssystem (z. B. bildungsferner familiärer Hintergrund, sonderpädagogischer Förderbedarf) dürfen nicht abgehängt werden.

Nach KI-Act der EU¹² handelt es sich bei der Bildung um einen Hochrisikobereich, in dem durch frühe Entscheidungen ganze Lebensläufe und Bildungschancen maßgeblich beeinflusst werden; (Datenschutz-) rechtliche Rahmenbedingungen zur Messung, zum Gebrauch und zur Sichtbarmachung von (ggf. sensiblen) Daten sollten zügig geklärt werden. Außerdem bedarf es einer fortlaufenden ethischen Diskussion dazu, (1) welche Datennutzung oder -Nicht-Nutzung im Rahmen von Learning Analytics ethisch fragwürdig sind, wenn dadurch Bildungschancen verwehrt werden, und (2) wie viel Adaptivität durch Learning Analytics wünschenswert ist bzw. ab wann eine Über-Adaptivität zu langfristigem De-skilling führen kann.

1.2.4 Effekte und Bedingungen der Nutzung von Daten zu überfachlichen Konstrukten

Überfachliche Fähigkeiten und Fertigkeiten (wie Kommunikationsfähigkeit, Problemlösen, Kooperation, Emotions- und Selbstregulation) sowie Überzeugungen und Einstellungen (wie das Vertrauen in die eigene Wirksamkeit, die Fähigkeit zur Selbstreflexion, Leistungsbereitschaft oder Engagement) stellen nicht nur eigenständige Bildungsziele dar (vgl. Kapitel 1.1.4), sie sind auch Grundlagen erfolgreichen Lernens (Durlak et al., 2022; SWK, 2022a). Der Zusammenhang zwischen Leistung und überfachlichen Aspekten wie Selbstregulation, sozialer Integration und Wohlbefinden ist wissenschaftlich gut belegt (Dent & Koenka, 2016; Donker et al., 2014; Marrone et al., 2024; Nationale Akademie der Wissenschaften Leopoldina, 2024; siehe auch SWK-Gutachten zur Sekundarstufe, 2025).

In vielen OECD-Staaten lässt sich in den letzten Jahren eine verstärkte Aufmerksamkeit für überfachliches Lernen feststellen. Dies ist nicht nur veränderten gesellschaftlichen und ökonomischen Kompetenzerwartungen geschuldet (OECD, 2021), sondern auch Befunden aus Surveystudien (Reiß et al., 2024) und Metaanalysen (Barkmann & Schulte-Markwort, 2012), die soziale und emotionale Probleme bei bis zu einem Fünftel der Kinder und Jugendlichen feststellen.

Die Bezeichnung *überfachliche Kompetenzen* wird als Sammelbegriff verwendet, der Dispositionen (Potenziale) beschreibt, die relevant sind für eine gelingende Lebensbewältigung in komplexen, insbesondere sozialen Kontexten (Fend, 1998; Grob et al., 2003). Häufig ist von sogenannten *21st Century Skills* die Rede (Trilling & Fadel, 2009). International werden derzeit vor allem zwei Frameworks diskutiert, die überfachliche Kompetenzen theoretisch beschreiben und operationalisieren: (1) *Collaborative for Academic, Social, and Emotional Learning (CASEL)*, das die Grundlage für einen systemischen Ansatz zur Förderung sozial-emotionaler Kompetenzen in Schulen darstellt (Dusenbury et al., 2015), sowie (2) *Social & Emotional Skills (SES)*, auf dessen Grundlage überfachliche Kompetenzen in OECD-Surveys erfasst werden (Chernyshenko et al., 2018; John et al., 2010).

¹² https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32024R1689#anx_III (aufgerufen am 20.02.2026). Laut Anhang 3 zum KI-Act gelten im Bereich allgemeine und berufliche Bildung solche Systeme als Hochrisiko-KI-Systeme, die (a) „zur Feststellung des Zugangs oder der Zulassung [zur] allgemeinen und beruflichen Bildung“, (b) „für die Bewertung von Lernergebnissen“, (c) [zur] Bewertung des [...] Bildungsniveaus“ sowie (d) „zur Überwachung und Erkennung von verbotenem Verhalten von Schülern bei Prüfungen“ verwendet werden sollen.



Beide Frameworks haben deutliche Überschneidungen mit dem Konstrukt *Selbstregulation*. Dieses Bündel kognitiver, sozialer und emotionaler Kompetenzen ermöglicht, individuelle Ziele zu entwickeln und effektiv zu verfolgen sowie Probleme zu lösen und Unsicherheit produktiv zu bewältigen (vgl. Stellungnahme Nationale Akademie der Wissenschaften Leopoldina, 2024).

Lehrkräfte haben einen großen Einfluss auf die sozial-emotionale Entwicklung von Schüler:innen (Schoenert-Reichl, 2017, 2019). Der Unterricht steht im Zentrum schulischer Ansätze zur Förderung sozial-emotionaler Kompetenzen, allein deshalb, weil Schüler:innen den Großteil ihrer Zeit im Klassenzimmer verbringen und dort mit leistungsbezogenen und sozialen Anforderungen konfrontiert sind, deren erfolgreiche Bewältigung einerseits sozial-emotionale Kompetenzen erfordert und die andererseits wiederum Gelegenheiten zum Erwerb sozial-emotionaler Kompetenzen darstellen (vgl. SWK, 2022a). Dabei beeinflusst das Unterrichtsverhalten der Lehrkräfte nicht nur die Lernergebnisse der Schüler:innen, sondern ist auch mit ihrem eigenen beruflichen Wohlbefinden assoziiert (Klusmann et al., 2022; Wartenberg et al., 2023).

Die Nutzung von Daten zu sozial-emotionalen Kompetenzen ermöglicht eine gezielte Förderung (OECD, 2021). International liegen einige Instrumente zur Erfassung sozial-emotionaler Kompetenzen für unterschiedliche Altersgruppen vor, allerdings vorwiegend für Forschungszwecke entwickelt. Häufig werden sozial-emotionale Kompetenzen über Selbstberichte erfasst (z. B. *SESAS – Social-Emotional Skills Assessment Scale*; *WCSD-SECA – Washoe County School District Social and Emotional Competency Assessment*; *SEC-Q – Social and Emotional Competencies Questionnaire*; *SEC-I-SR – Social-Emotional Competences Inventory* zur Selbsteinschätzung; Wolf et al., 2020). Teilweise existieren auch andere Testformate wie *Situational Judgement Tests* (*Zoo U*, ein gamebasiertes Assessment; *SEC-SJT – Social-Emotional Situational Judgement Test*) oder Ratingskalen für Lehrkräfte und Eltern (*SECQ – Social and Emotional Competences Evaluation Questionnaire*; *SEC-I-OR – Social-Emotional Competences Inventory* zur Fremdeinschätzung).

Eine valide Erfassung sozial-emotionaler Kompetenzen im Schulalltag ist ebenso wie die Erfassung anderer psychologischer Konstrukte mit methodischen Herausforderungen verbunden (Kautz et al., 2014). Zwar ist unstrittig, dass Selbstberichte prinzipiell internale psychische Zustände, z. B. ein Gefühl der Zugehörigkeit, valide erfassen können, bei anderen Konstrukten bestehen allerdings Risiken der Urteilsverzerrung (Duckworth & Yeager, 2015). Außerdem sind entwicklungsbedingte Veränderungen bzw. negative Entwicklungen im Verlauf der Schulzeit (wie etwa bei Selbstwirksamkeitserwartungen) in Rechnung zu stellen (Rimm-Kaufman et al., 2025). Deshalb sollte sowohl von einer Individualdiagnostik als auch von einer Datennutzung für Accountabilityzwecke aktuell Abstand genommen werden. Für eine Diagnose auffälliger Verhaltensweisen von Schüler:innen sind andere diagnostische Verfahren sowie multiprofessionelle Kooperationen erforderlich (vgl. dazu ausführlich SWK-Gutachten zur Grundschule, 2022a).

Instrumente zur Erfassung sozial-emotionaler Kompetenzen für den Einsatz in Schulen müssen sich an praktischen Anforderungen ausrichten (Duckworth & Yeager, 2015). Obwohl in den letzten Jahren eine ganze Reihe (teilweise auch digitale) Verfahren (im Überblick siehe McKown, 2019) entwickelt wurden, ist insbesondere im deutschsprachigen Raum weitere systematische Entwicklungsarbeit erforderlich. Für einen flächendeckenden Einsatz an Schulen eignen sich eher Instrumente, die schulbezogene sozial-emotionale Merkmale wie z. B. Wohlbefinden oder soziale Eingebundenheit mit einem vertretbaren Aufwand erfassen. Länder wie Kanada oder Österreich haben mittlerweile nicht nur Schülerbefragungen zu



überfachlichen Bildungszielen wie Wellbeing oder soziale Zugehörigkeit in ihr Monitoringsystem integriert (vgl. Kapitel 5), sie melden Lehrkräften diese Daten neben Daten zu fachlichen Leistungen auch regelmäßig zurück.

Für die gezielte Nutzung dieser Daten ist eine unmittelbare Verknüpfung mit evidenzbasierten Ansätzen der Förderung sozial-emotionaler Kompetenzen sehr sinnvoll. Die in den USA breit genutzte CASEL-Plattform³³ stellt beispielsweise ein Verzeichnis evidenzbasierter Programme zur Förderung sozial-emotionaler Kompetenzen bereit. Lehrkräfte und Schulen können jene entsprechend ihren Zielen in den Unterricht implementieren (Mahoney et al., 2021). In Deutschland kann über die *Grüne Liste Prävention*³⁴ auf evidenzbasierte Programme zur Förderung sozial-emotionaler Kompetenzen zugegriffen werden.

Es liegen zahlreiche Metaanalysen vor, die Effekte einer Implementation von Programmen zur Förderung sozial-emotionalen Lernens in Unterricht und Schule auf prosoziales Verhalten, psychische Anpassung und Leistungen belegen (Cipriano et al., 2023; Corcoran et al., 2018; Durlak et al., 2022).

Neben systemischen Programmen, in denen Diagnose und Intervention unmittelbar verknüpft sind, existieren eine ganze Reihe von Instrumenten, mit denen Aspekte wie Wellbeing, mentale Gesundheit, Lebensqualität oder soziale Zugehörigkeit valide erfasst werden können. Diese Instrumente wie der *Strengths and Difficulties Questionnaire* von Goodman (1997) oder einzelne Skalen aus *KIDSCREEN* (Ravens-Sieberer et al., 2005) werden derzeit vor allem in Surveystudien eingesetzt (z. B. *Deutsches Schulbarometer*, Bildungstrend, *PISA*). Sie eignen sich prinzipiell auch für den Einsatz im Klassenzimmer. Allerdings müssen Unterstützungs- und Qualifikationsangebote bzw. Schnittstellen zu anderen Professionen (Schulpsychologie, Sozialarbeit) gewährleistet werden, wenn Schulen einen Handlungsbedarf feststellen, den sie mit den eigenen Ressourcen nicht hinreichend adressieren können. Ein gezieltes Training von Lehrkräften ermöglicht es Lehrkräften auch, die Förderung sozial-emotionaler Kompetenzen in ihren Unterricht einzubauen (Schonert-Reichl, 2017).

Es spricht viel dafür, die Förderung sozial-emotionaler Kompetenzen im inklusiven Unterricht an einem dreistufigen Modell zu orientieren (Casale et al., 2018; Gordon, 1987; SWK-Gutachten zur Grundschule, 2022a). Auf der Grundlage der regelmäßigen Erfassung zentraler Merkmale sozial-emotionalen Lernens wie Wellbeing, soziale Eingebundenheit, Gefühl der Zugehörigkeit, Empathie und Growth Mindset im Rahmen der Unterrichtsevaluation (vgl. Kap. 1.2.5) kann eine unterrichtsintegrierte Förderung sozial-emotionaler Kompetenzen aller Schüler:innen erfolgen. Bei Auffälligkeiten können dann gezielte diagnostische Instrumente zur Erfassung sozial-emotionaler Kompetenzen eingesetzt und Maßnahmen, wie z. B. explizite Trainings sozial-emotionaler Kompetenzen (vgl. Casale et al., 2018), durchgeführt werden. Für Schüler:innen, die einen besonderen Unterstützungsbedarf aufweisen, müssen eine professionelle Diagnostik und eine spezifische Förderung durch multiprofessionelle Teams erfolgen.

³³ <https://pg.casel.org/review-programs/> (aufgerufen am 20.02.2026).

³⁴ <https://www.gruene-liste-praevention.de/nano.cms/datenbank/information> (aufgerufen am 20.02.2026).



1.2.5 Effekte und Bedingungen der Nutzung von Instrumenten zur Erfassung der Unterrichtsqualität

Ein datengestütztes Unterrichtsfeedback kann neben Daten zu Lernständen der Schüler:innen ein wirksames Tool zur Weiterentwicklung des Unterrichts sein (Röhl, 2021). Wenn ein solches Feedback mit gezielten Unterstützungsangeboten zur Unterrichtsentwicklung verknüpft wird, zeigen sich – so zwei Metaanalysen – sogar mittlere bis hohe Effekte auf die Unterrichtsqualität (Penny & Coe, 2004; Röhl, 2021). Allerdings mündet die veränderte Unterrichtspraxis offensichtlich nicht notwendigerweise in eine Verbesserung von Schülerleistungen. So sind die Effekte auf fachliche Leistungen häufig schwach bzw. die Befundlage ist inkonsistent (Charalambous & Praetorius, 2025; White et al., 2025). Drei generelle methodische Herausforderungen der Erfassung von Unterrichtsqualität werden zur Erklärung dieser Befundlage genannt:

- die Konzeptualisierung des Konstrukts Unterrichtsqualität; hier ist neben der Diskussion um eine Erweiterung des Modells der drei Basisdimensionen (Klassenführung, konstruktive Unterstützung und kognitive Aktivierung; Klieme, 2018) auch die Frage relevant, inwiefern Qualitätsmerkmale generisch, fach- oder sogar inhaltspezifisch operationalisiert und ausdifferenziert werden sollten (Dreher & Leuders, 2021; Praetorius et al., 2025),
- die Identifikation von Outcome-Maßen, mit denen Effekte erfasst werden (Grützmaker et al., 2025; Kunter & Voss, 2013), sowie
- der Einbezug von Kontextfaktoren, insbesondere die Zusammensetzung der Klasse (Fauth et al., 2020; Praetorius et al., 2025).

Grundsätzlich lassen sich zwei Quellen des Unterrichtsfeedbacks unterscheiden: trainierte Beobachter:innen, das können auch Kolleg:innen oder Schulleitungen sein, und Schüler:innen. Zwar werden Beobachtungen in der Forschung häufig als Goldstandard betrachtet, sie erfordern aber ein intensives Beobachtertraining (Pianta et al., 2008) und beziehen sich nur auf einen vergleichsweise kleinen Ausschnitt des Unterrichts. Schülerbefragungen sind nicht nur deutlich kostengünstiger, sondern beziehen sich auch auf viele unterschiedliche Situationen und einen längeren Zeithorizont. Außerdem ist die Schülerperspektive für gelingende Lernprozesse und damit auch für Unterrichtsentwicklung von hoher Relevanz. Eine ganze Reihe von Studien bescheinigen Schülerfeedbacks zum Unterricht über alle Dimensionen hinweg eine ähnliche psychometrische Güte wie Beobachtungen (Clausen, 2002; Fauth et al., 2014; Göllner et al., 2016; Lüdtke et al., 2009). Allerdings bestehen, was die Güte des Schülerurteils betrifft, Unterschiede hinsichtlich der verschiedenen Dimensionen; insbesondere kognitive Aktivierung ist für Schüler:innen schwerer zu beurteilen (Wagner et al., 2013), für Lernwirksamkeit jedoch zentral.

Aus der aktuellen Befundlage zur Erfassung von Unterrichtsqualität zum Zweck der Weiterentwicklung von Unterricht lassen sich folgende Schlussfolgerungen ziehen:

- Die theoretische Konzeptualisierung dreier basaler Dimensionen der Unterrichtsqualität hat sich für den Zweck des Unterrichtsfeedbacks bewährt (Trautwein et al., 2022). Zwar werden in der Unterrichtsforschung inzwischen auch erweiterte Modelle diskutiert (Praetorius & Charalambous, 2023); sie erfüllen aber nicht in gleicher Weise die Kriterien einer sparsamen Modellierung, des Nachweises von Reliabilität und prädiktiver Validität sowie der praktischen Relevanz und Akzeptanz (Kunter & Trautwein, 2013; Trautwein et al., 2022). Basierend auf dreidimensionalen Modellen



liegen verschiedene validierte Schülerfragebögen sowie Beobachtungsbögen vor (Gärtner, 2021; Gärtner et al., 2021; J. L. Maier et al., 2021).

- Ein Feedback auf der Grundlage eines generischen Modells zur Unterrichtsqualität ist für die Entwicklung von Unterricht dann sinnvoll, wenn die Indikatoren eng an psychologischen Theorien des Lernens orientiert sind (Trautwein et al., 2022), d.h. basale kognitive und motivationale Prozesse abbilden. Ein solches Feedback kann die Funktion eines Screenings haben, das durch eine fachdidaktisch ausdifferenzierte Beurteilung vertieft werden kann. Dazu können, wo vorhanden, spezifische fachbezogene Instrumente der Unterrichtsevaluation genutzt werden, die sich in Unterrichtsentwicklungsprojekten als fachdidaktisch handlungsleitend bewährt haben (Schoenfeld, 2018).
- Die Erfassung von Unterrichtsqualität durch Schüler:innen ist prinzipiell valide (Rollett et al., 2021). Es handelt sich um ein besonders ökonomisches Verfahren, das in der Fläche ohne größeren Aufwand implementiert werden kann. Dazu werden bereits zwei Varianten erprobt: zentrale Erhebung online-basierter Rückmeldungen, die in der Verantwortung wissenschaftsnaher Institute durchgeführt und den Lehrkräften rückgemeldet wird (J. L. Maier et al., 2021), oder ein validiertes App-basiertes Feedback, das die Lehrkraft in eigener Verantwortung bei den Schüler:innen abrufen kann (Bijlsma et al., 2019).
- Unterrichtsfeedbacks von Kollegien (Prilop et al., 2019) oder Schulleitungen (Kellermann et al., 2023) können neben Schülerfeedbacks wichtige Impulse für die Weiterentwicklung von Unterricht setzen. Beim kollegialen Feedback kann auch der Forderung nach Einbezug fachdidaktischer Expertise (Dreher & Leuders, 2021) Rechnung getragen werden. Eine zentrale Voraussetzung für die Validität der Unterrichtsbeobachtungen ist allerdings, dass Beobachter:innen qualifiziert werden und dass Materialien, Manuale sowie Anleitungen zum Unterrichtsfeedback zur Verfügung stehen (Kellermann et al., 2023). Der Aufwand für ein kollegiales Feedback bzw. ein Feedback durch die Schulleitung ist deutlich höher als der für ein Schülerfeedback. Eine Kombination eines vertiefenden kollegialen Feedbacks bzw. Feedbacks durch die Schulleitung in der Verantwortung der Schule mit in einem verpflichtend durchzuführenden Schülerfeedback erscheint deshalb sinnvoll.



1.3 Schlussfolgerungen und Empfehlungen

Die kontinuierliche Weiterentwicklung des Unterrichts auf der Basis einer systematischen Nutzung von Daten bildet eine zentrale Voraussetzung für die gezielte Förderung von Lernprozessen und Bildungsgerechtigkeit. Um Lehr- und Lernprozesse systematisch zu verbessern, bedarf es einer regelmäßigen Rückmeldung zu Lernausgangslagen, Lernergebnissen und Lernentwicklungen zu fachlichen und überfachlichen Kompetenzen. In allen Ländern sollte deshalb ein kohärentes Testsystem zur Unterrichtsentwicklung mit den erforderlichen Unterstützungsmaßnahmen umgesetzt werden, das folgende Elemente umfasst:

1. verbindliche *Überprüfung der Lernausgangslagen* zu Beginn der Primarstufe und zu Beginn der Sekundarstufe I,
2. verbindliche *Überprüfung des auf die Bildungsstandards bezogenen Lernerfolgs* als Frühwarnsystem, um Fokusbereiche für die weitere Unterrichtsplanung zu identifizieren; hierbei sollte die *Frühwarnfunktion* von VERA weiterentwickelt werden, z. B. anhand angepasster Kompetenzstufenmodelle, die eine für Lehrkräfte gut interpretierbare Standortbestimmung im Verhältnis zu den Bildungsstandards ermöglichen,
3. *Überprüfung der Wirksamkeit getroffener Maßnahmen*, die sich aus (1) bzw. (2) ergeben haben.

Dabei handelt es sich um ein Minimalprogramm für die Fächer Deutsch, Mathematik und (in der Sekundarstufe I) Englisch, wofür überwiegend bereits ländergemeinsame Instrumente vorliegen bzw. aktuell entwickelt werden (vgl. Abbildung 3). Eine Übertragung der Prinzipien auf andere Fächer, für die Bildungsstandards vorliegen (aktuell die Fächer Biologie, Chemie und Physik in der Sekundarstufe I) bzw. zukünftig entwickelt werden, könnte in Betracht gezogen werden. Eine längsschnittliche Verknüpfung zur Abbildung von Lernentwicklungen wäre nach aktuellem Stand zwischen Klasse 1 und 2 (StarS) sowie zwischen Klasse 2 und 3 (BKT)³⁵, außerdem (z. B. durch Testwiederholung) zwischen Klasse 3 (VERA) und 4 sowie zwischen Klasse 8 (VERA) und 9 möglich, was ein geeignetes ID-Management bzw. eine Verknüpfbarkeit der Daten erfordert.

³⁵ Für einzelne Kompetenzbereiche möglicherweise auch zwischen 1 und 3 (StarS und BKT); das wird vom IQB geprüft.



Klasse/Zeitpunkt	Erfasste Kompetenzbereiche*, Funktionen	Verbindlichkeit	Akt. Bezeichnung bzw. Arbeitstitel**
1 Anfang	basale Kompetenzen Lernausgangslage	verbindlich	StarS 1
2 Anfang	basale Kompetenzen Überprüfung Wirksamkeit von Maßnahmen	verbindlich	StarS 2
2 Mitte	basale Kompetenzen Überprüfung Wirksamkeit von Maßnahmen	optional in Abh. von Ergebnissen Anfang Klasse 2	BKT 2
3 Mitte/Ende	basale Kompetenzen Überprüfung Wirksamkeit von Maßnahmen	optional in Abh. von Ergebnissen Klasse 2	BKT 3
3 Mitte/Ende	Erreichen Bildungsstandards Frühwarnsystem	verbindlich	VERA 3
4 Mitte/Ende	Erreichen Bildungsstandards Überprüfung Wirksamkeit von Maßnahmen	verbindlich	
5 Anfang	Erreichen Bildungsstandards Lernausgangslage	verbindlich	
6			
7			
8	Erreichen Bildungsstandards Frühwarnsystem	verbindlich	VERA 8
9	Erreichen Bildungsstandards Überprüfung Wirksamkeit von Maßnahmen	verbindlich	

Farbig unterlegte Zeilen: vorhandene bzw. bereits in der Entwicklung befindliche ländergemeinsame Instrumente

*Bezogen auf die Fächer Deutsch, Mathematik (Grundschule und Sek. I) sowie Englisch (Sek. I). Eine Übertragung der Prinzipien auf andere Fächer, für die Bildungsstandards vorliegen (aktuell die Fächer Biologie, Chemie und Physik in der Sek. I) bzw. zukünftig entwickelt werden, könnte in Betracht gezogen werden.

**Um die Kohärenz der Instrumente deutlich zu machen, sollte eine gemeinsame Bezeichnung gefunden werden. Für die Jahrgangsstufen 3/4/5 und 8/9 sollten jeweils dieselben Instrumente eingesetzt werden und Lernverläufe sichtbar gemacht werden.

Abbildung 3: Zentrale Lernstandserhebungen (vgl. Empfehlung 4)

Darüber hinaus wäre zu prüfen, ob perspektivisch (a) die basalen Kompetenztests für Klasse 1 und (insbesondere) Klasse 2 so weiterentwickelt werden sollen, dass sie das gesamte Kompetenzspektrum abdecken, sowie (b) ob für weitere Klassenstufen in der Sekundarstufe I Testungen vorgesehen werden sollen (Klasse 6 und Klasse 7), um Lehrkräften in jeder Klassenstufe Rückmeldungen über den erreichten Leistungsstand ihrer Schüler:innen zu geben und die Unterrichtsentwicklung dadurch fortlaufend zu orientieren, wie es beispielsweise in Kanada der Fall ist.



Eine jährliche (Grundschule), mindestens jedoch zweijährliche (Sekundarstufe) Durchführung und Auswertung standardisierter Lernstandserhebungen schafft hierfür eine verlässliche Datengrundlage. Entscheidend ist dabei, dass die Rückmeldungen nicht nur kriteriale und soziale, sondern auch adjustierte Vergleiche beinhalten, die die unterschiedlichen Zusammensetzungen der Schülerschaft berücksichtigen. Auf diese Weise können Schulen fundierte Aussagen über Lernstände treffen und Entwicklungsverläufe aussagekräftig interpretieren.

Die Analyse der Ergebnisse sollte verpflichtend in die schulinterne Unterrichtsentwicklung einfließen. Lehrkräfte und Fachkonferenzen benötigen hierfür gezielte Fortbildungen und Unterstützungsangebote, die sie befähigen, Daten kompetent auszuwerten und didaktisch wirksam zu nutzen.

Neben zentralen Erhebungen kommt der lernprozessbegleitenden Diagnostik eine große Bedeutung zu, die eng an lernwirksame Förderansätze geknüpft ist. Diagnostische Instrumente, die direkt in den Unterricht integriert und auf spezifische Unterrichtsinhalte bezogen sind, ermöglichen es, auf individuelle Lernverläufe zeitnah zu reagieren und adaptive Fördermaßnahmen abzuleiten. Hierzu bedarf es fachdidaktisch treffsicherer Materialien, die Diagnose und Förderung eng miteinander verknüpfen. Insbesondere digitale Lernumgebungen bieten hier wertvolle Potenziale. Damit Lehrkräfte diese Materialien systematisch nutzen können, sollten Fortbildungen zur Interpretation und Nutzung solcher Daten ein fester Bestandteil des Professionalisierungsprozesses sein. Dazu besteht ein erheblicher weiterer Entwicklungsbedarf, der idealerweise ländergemeinsam erfolgen sollte.

Schließlich sollte die Perspektive der Schüler:innen selbst in die datengestützte Unterrichtsentwicklung integriert werden. Valide Instrumente zur Erfassung der Unterrichtsqualität aus Schülersicht sowie zur Erhebung sozial-emotionaler Merkmale – wie Wohlbefinden, schulische Zufriedenheit und soziale Eingebundenheit – liefern wichtige Hinweise auf die Wirksamkeit und Akzeptanz pädagogischer Maßnahmen. In der Situationsanalyse zeigte sich allerdings, dass sozial-emotionales Lernen zwar als Bildungsziel überall festgeschrieben ist, bislang aber nur in wenigen Ländern systematisch oder gar flächendeckend Daten dazu erfasst und Fördermaßnahmen angeboten werden.

Eine umfassende Datengrundlage eröffnet Schulen die Möglichkeit, Unterricht und Förderung noch gezielter an den Bedürfnissen der Lernenden auszurichten und damit die Qualität schulischer Bildung insgesamt zu steigern. Vor diesem Hintergrund empfiehlt die SWK:

Empfehlung 4) Ergebnisse zentraler Lernstandserhebungen rückmelden, möglichst jährlich, für eine durchgängige, datengestützte Unterrichtsentwicklung und Förderung der Schüler:innen, mindestens in den Fächern, für die Bildungsstandards vorliegen; dabei

1. ein Minimalprogramm von Rückmeldungen zu erreichten Kompetenzen verbindlich umsetzen und für die Unterrichtsentwicklung nutzen. Das Minimalprogramm sollte folgende Elemente umfassen (vgl. Abbildung 3):
 - (1) Überprüfung der Lernausgangslagen zu Beginn der Primarstufe und zu Beginn der Sekundarstufe I;
 - (2) Überprüfung des auf die Bildungsstandards bezogenen Lernerfolgs als Frühwarnsystem (Jgst. 3 und 8);
 - (3) Überprüfung der Wirksamkeit getroffener Maßnahmen, die sich aus (1) bzw. (2) ergeben haben (Jgst. 4 und 9)



- und dabei die bereits ländergemeinsam entwickelten Verfahren (Tests zur Erfassung basaler Kompetenzen 1 bis 3 und VERA) konsequent nutzen und weiterentwickeln,
2. weitere Tests entwickeln, um möglichst in jeder Jahrgangsstufe (Grundschule), mindestens aber im Abstand von zwei Jahrgangsstufen (Sekundarstufe), eine Rückmeldung zu Lernentwicklungen zu geben, sofern dies für den jeweiligen Kompetenzbereich sinnvoll ist (in der Regel über einen Zeitraum von zwei Jahren),
 3. prüfen, ob weitere Tests ländergemeinsam entwickelt werden sollen (z. B. für Rückmeldungen zu höheren Kompetenzniveaus in Jahrgang 1 und 2, für Testungen in Jahrgang 6 und 7), um Lehrkräfte umfassend und fortlaufend zu orientieren,
 4. etablierte nutzerorientierte Rückmeldeformate nutzen und ggf. ausweiten, die neben kriterialen und sozialen auch adjustierte Vergleiche beinhalten, um die Zusammensetzung der Schülerschaft zu berücksichtigen,
 5. für adjustierte Vergleiche nach Möglichkeit Ausgangsleistungen der Schüler:innen zu Beginn der jeweiligen Bildungsetappe verwenden bzw. diejenigen Merkmale, die in den Sozialindex des jeweiligen Landes eingehen,
 6. Ergebnisse zentraler Lernstandserhebungen verbindlich für die Unterrichtsentwicklung durch Schulen und Lehrkräfte nutzen, wozu es der Bereitstellung evidenzbasierter diagnose- und fördergeleiteter Beratungs- und Begleitmaterialien bedarf,
 7. Lehrkräfte und Fachkonferenzleitungen/Fachteams mit verbindlichen Fortbildungen und Unterstützungsangeboten für datengestützte Unterrichtsentwicklung qualifizieren.

Empfehlung 5) Lernprozessbegleitende Diagnoseinstrumente zur adaptiven Lernförderung einsetzen; dabei

1. für zentrale Fächer und lehrplanrelevante Themen diagnosegeleitete Fördermaterialien implementieren, die fachdidaktisch treffsichere lernprozessbegleitende Diagnoseinstrumente mit geeignetem Unterrichtsmaterial verknüpfen; diese diagnosegeleiteten Fördermaterialien in digitalen Zentren (weiter-)entwickeln und bereitstellen (vgl. SWK-Gutachten zur Digitalisierung, 2022b),
2. adaptive, möglichst barrierearme digitale Lernumgebungen und Learning Analytics zur Gewinnung von Daten zu Lernprozessen der Schüler:innen bereitstellen und implementieren; Fortbildungsangebote für Lehrkräfte zum Umgang mit diesen Daten sind integraler Bestandteil,
3. diagnosegeleitete Lernförderung und Unterrichtsentwicklung systematisch unterstützen durch Bereitstellung von evidenzbasierten Materialien, Fachberatung und Fortbildungsangeboten.



Empfehlung 6) Valide Instrumente zur Erfassung von Unterrichtsqualität aus der Schülerperspektive einerseits und schulbezogenen sozial-emotionalen Aspekten (insbesondere Wohlbefinden, schulische Zufriedenheit und soziale Eingebundenheit) andererseits implementieren, die verpflichtende und wählbare Komponenten umfassen; dabei

1. verpflichtende, wenn möglich auf Länderebene zentral administrierte Schülerbefragungen zu zentralen Dimensionen der Unterrichtsqualität in jedem Schuljahr durchführen, die auch Konstrukte zum Wohlbefinden, zu schulischer Zufriedenheit und sozialer Eingebundenheit umfassen,
2. existierende valide Unterrichtsfeedback-Instrumente ländergemeinsam weiterentwickeln als digitale Tools für die verpflichtenden Befragungen (Apps, Plattformen für zentrale Erhebungen), inklusive Manualen und verständlicher Rückmeldeformate für Lehrkräfte,
3. Instrumente, die über die basale verpflichtende Befragung hinausgehen (z. B. zu sozial-emotionalen Merkmalen, zu fachlichem Interesse, zu kooperativem Lernen), weiterentwickeln und für einen anlassbezogenen Einsatz in der Verantwortung der Lehrkräfte bereitstellen,
4. zur Unterrichtsentwicklung qualitativ hochwertige Unterstützungsangebote wie Unterrichtscoachings bereitstellen, die im optimalen Fall Leistungsdaten und Daten zur Prozessqualität im Zusammenhang nutzen und den fachdidaktischen Fokus stärken.



2. Datennutzung für Schulentwicklung im Rahmen erweiterter Gestaltungsspielräume

Während Daten auf der Klassenebene von Lehrkräften zur Unterrichtsentwicklung und Lernförderung genutzt werden (vgl. Kapitel 1), geht es bei der Datennutzung auf der Ebene der Schulorganisation um die Abstimmung gemeinsamer Ziele, die Identifikation gesamtschulischen Entwicklungs- und Fortbildungsbedarfs sowie um die Implementation entsprechender Maßnahmen. In Deutschland ist die Schulentwicklungsarbeit eng verbunden mit dem seit Beginn der 2000er-Jahre in fast allen Ländern eingeführten Schulprogramm, in dem Schulen ihre Entwicklungsziele dokumentieren und Maßnahmen zur Qualitätssicherung und -entwicklung beschreiben (Tarkian, Riecke-Baulecke & Thiel, 2019). Das Schulprogramm wird entsprechend auch als „schulinternes Strategieinstrument“ bezeichnet (Bonsen & Frey, 2024, S. 13). Schulprogrammarbeit erfolgt in enger Verbindung mit interner Evaluation bzw. Selbstevaluation, deren Funktion darin besteht, die Umsetzung der Maßnahmen zu dokumentieren, ihre Wirksamkeit kritisch zu reflektieren und Aufschluss zu geben über die Prozessqualität schulischer Arbeit. Schulprogramm und interne Evaluation gelten als zentrale Instrumente schulischer Selbststeuerung. Sie sind Ausdruck der Gestaltungsspielräume einer professionellen Organisation (Lange, 2003a).

Die Maßstäbe für datengestützte Schulentwicklung und damit die Grundlage für Schulprogramme und interne Evaluation definieren sogenannte Qualitätsrahmen, die im Idealfall Qualitätsindikatoren auf der Grundlage wissenschaftlicher Evidenz so operationalisieren, dass sie Freiräume für die Entwicklung konkreter Maßnahmen vor Ort lassen. Die Bedeutung schulische Gestaltungsspielräume wird durch empirische Befunde gestützt, die zeigen, dass schulische Autonomie in Verbindung mit einer konsequenten Rechenschaftslegung nicht nur mit höheren Leistungen der Schüler:innen einhergeht (Eyles & Machin, 2019; Loeb & Strunk, 2007; Wößmann, 2007a), sondern auch mit einer höheren Motivation und Arbeitszufriedenheit von Schulleitungen (Horwood et al., 2022).

Die Wirksamkeit datengestützter Schulentwicklung ist nicht allein davon abhängig, dass Ziele treffend operationalisiert und evidenzbasierte Maßnahmen umgesetzt werden, sondern auch davon, dass für interne Evaluation valide Daten zu relevanten schulischen Prozessen und Ergebnissen genutzt werden (Schildkamp et al., 2017). Die Länder bieten Schulen deshalb vielfach Instrumente zur Erfassung von Daten zur Unterrichts- und Schulqualität an, stellen mit Vergleichsarbeiten sowie anderen Lernstandserhebungen Daten zum Lernerfolg von Schüler:innen bereit und unterstützen mit Maßnahmen, um fachdidaktische Konsequenzen zu ziehen (vgl. Kapitel 1). Neben interner Evaluation und Lernstandserhebungen sollen in einigen Ländern auch Daten aus externen Evaluationen für Schulentwicklungsprozesse genutzt werden (vgl. Kapitel 3).

Entscheidend für die Nutzung der unterschiedlichen Daten für eine gezielte Schulentwicklung sind einerseits die verständliche Aufbereitung der Daten bzw. die Rückmeldeformate, andererseits die Unterstützungsprozesse, die sich an die Dateninterpretation anschließen. Hier sind insbesondere Beratungs- und Qualifikationsangebote zu nennen. Darüber hinaus tragen Strukturen wie Funktionsstellen für datengestützte Schulentwicklung oder Schulnetzwerke zu einer effektiven Datennutzung bei.



In der nachfolgenden Situationsanalyse wird dargestellt, welche Rolle interne Evaluation im Zusammenhang mit Schulprogrammen für eine datengestützte Schulentwicklung spielt und welche weiteren Daten (Lernstandserhebungen, externe Evaluation) für eine datengestützte Schulentwicklung in welcher Form erfasst und aufbereitet (z. B. digitale Rückmeldeformate) werden. Abschließend wird ein Überblick über die in den Ländern derzeit implementierten Strukturen (z. B. Funktionsstellen, Schulnetzwerke) und Anschlussprozesse (von Ziel- und Leistungsvereinbarungen bis hin zu Qualifikation und Beratung) für eine datengestützte Schulentwicklung gegeben.

2.1 Situationsanalyse

2.1.1 Daten für Schulentwicklung aus interner und externer Evaluation sowie Lernstandserhebungen

Datengestützte Schulentwicklung erfordert ein gezieltes Ineinandergreifen unterschiedlicher Prozesse. Im Unterschied zur Unterrichtsentwicklung steht dabei die Entwicklung der gesamten Organisation Schule im Zentrum. Insofern geht es um die Formulierung strategischer und operativer Ziele für die gesamte Schule auf der Grundlage einer Bestandsaufnahme, um die Abstimmung gemeinsamer Maßnahmen, um die Festlegung von Indikatoren, mit denen der Erfolg der Maßnahmen überprüft werden kann, sowie um die Etablierung von Strukturen, die eine kontinuierliche Entwicklung ermöglichen. In Deutschland werden vor allem folgende Instrumente für eine strategische Schulentwicklung genutzt: Schulprogramm, interne Evaluation sowie Daten aus externer Evaluation und Lernstandserhebungen, schulinterne Fortbildungen bzw. kollegiale Professionalisierung und Maßnahmen schulischer Personalentwicklung. Im Folgenden wird dargestellt, wie die maßgeblichen Instrumente der datengestützten Schulentwicklung in den Ländern implementiert sind. Die Angaben gehen auf die SWK-Länderabfrage Schule sowie eigene Recherchen zurück.

Schulprogramme

Ein zentrales Instrument für die datengestützte Schulentwicklung ist das Schulprogramm. Schulprogramme enthalten eine datengestützte Bestandsaufnahme, auf deren Grundlage idealerweise Herausforderungen und strategische Ziele identifiziert werden. Bezogen auf die Ziele erfolgt eine Entscheidung für die Implementation von Maßnahmen in einem bestimmten Zeithorizont, deren Wirksamkeit dann wiederum mit Bezug auf Erfolgsindikatoren überprüft werden soll. Interne Evaluation spielt für die Schul(programm)entwicklung eine zentrale Rolle. Schulprogramme wurden nach 1990 über einen längeren Zeitraum in fast allen Ländern eingeführt. Die Schulgesetze der meisten Länder sehen ein Schulprogramm, manchmal auch bezeichnet als *Schulentwicklungsprogramm* (Bayern), vor. In den Ländern Baden-Württemberg und Rheinland-Pfalz werden entsprechende Dokumente zwar nicht als Schulprogramm bezeichnet, es bestehen aber analoge Prozesse: In Baden-Württemberg erstellen Schulen eine Dokumentation des Schulentwicklungsprozesses, in der Ziele, Maßnahmen, Überprüfung und Anpassung kurz und knapp dokumentiert sind.¹⁶ In dem Land sind der Prozess der Zielvereinbarungen zwischen Schulaufsicht und Schule sowie

¹⁶ <https://referenzrahmen.kultus-bw.de/Lde/Startseite/Unterstuetzung/Schulentwicklungsprozess> (aufgerufen am 20.02.2026).



die schulische Entwicklung eng verknüpft. Die Zielvereinbarungen, in die sowohl die bildungspolitischen Ziele des Landes als auch die Ziele der Schule vor Ort einfließen, bilden die Grundlage für die Definition von Maßnahmen sowie deren Evaluation (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2024). In Rheinland-Pfalz besteht zwar keine explizite Verpflichtung, ein Schulprogramm zu erstellen, der vorgegebene Zyklus der Schulentwicklung (Vision entwickeln, Bestandsaufnahme durchführen, Handlungsfelder herausarbeiten, Ziele formulieren und Maßnahmen planen, Zielvereinbarung abschließen, Maßnahmen umsetzen und evaluieren)¹⁷ entspricht allerdings den Vorgaben anderer Länder für das Schulprogramm.

Daten zur Implementation von Schulprogrammen in Deutschland liegen aus Studien vor, die zwischen den Jahren 2000 und 2006 durchgeführt wurden. Die untersuchten Schulprogramme variierten erheblich nach Umfang und Qualität. Insgesamt kommen die Dokumentenanalysen zu dem Schluss, dass es in einer Vielzahl von Programmen nicht gelingt, Ziele eindeutig zu formulieren, zu operationalisieren und von Maßnahmen zu unterscheiden. Die meisten Programme enthalten zwar eine allgemeine Bestandsaufnahme, allerdings basiert diese meistens nicht auf einer differenzierten Analyse vorhandener Daten und mündet nicht in eine Problemanalyse. Einzelmaßnahmen werden oft nicht aus der Bestandsaufnahme abgeleitet und stehen unverbunden nebeneinander. Zur Evaluation der Maßnahmen finden sich häufig gar keine Aussagen oder nur vage Indikatoren. Insgesamt wird einer großen Zahl der Schulprogramme mangelnde Kohärenz bescheinigt (Hilbrich & Walter, 2003; Holtappels & Müller, 2004; Mohr, 2006). Die Befunde der Dokumentenanalysen wurden durch Befragungen von Schulleitungen (Hilbrich & Walter, 2003; Mohr, 2006) und Schulaufsichten (Burkard, 2004) bestätigt. Die im Rahmen der *Internationalen Grundschul-Lese-Untersuchung (IGLU)* durchgeführte Schulleitungsbefragung legte zudem eine deutliche Skepsis hinsichtlich der Nutzung von Daten für Bestandsaufnahme und Evaluation offen (Mohr, 2006).

Neuere Analysen zur Qualität von Schulprogrammen liegen nicht vor (vgl. zur Anzeige des Desiderats Alt-richter, 2019). Eine nicht repräsentative Sichtung von Schulprogrammen aus unterschiedlichen Ländern lässt allerdings vermuten, dass die in den zitierten Studien beschriebenen Mängel nach wie vor bestehen. Auffällig ist nicht nur, dass die Schulprogramme in ihrem Umfang erheblich variieren, sondern auch, dass in den gesichteten Programmen die Bestandsaufnahme in den wenigsten Fällen datengestützt erfolgt und dass keine eindeutigen Indikatoren für die Überprüfung der Maßnahmen definiert werden.

Daraus lässt sich schließen, dass das Potenzial von Schulprogrammen als Steuerungs- und Entwicklungsinstrument derzeit nicht ausgeschöpft wird und Schulprogramme in der gegenwärtigen Form ihre strategische Bedeutung für die schulische Qualitätsentwicklung nicht entfalten können.

Interne Evaluation

Zwischen interner Evaluation und der Arbeit am Schulprogramm besteht ein enger Zusammenhang, da die Überprüfung und Weiterentwicklung des Schulprogramms meist als (ein) maßgebliches Ziel interner Evaluationen definiert ist.

¹⁷ <https://bildung.rlp.de/schule/schulentwicklung/se-prozess> (aufgerufen am 20.02.2026).



In den meisten Ländern sind Schulen verpflichtet, schulinterne Evaluationen zur Unterrichts- und Schulqualität regelmäßig durchzuführen. Interne Evaluationen sind hierbei mit einer Entwicklungsfunktion verknüpft. Als Ziele interner Evaluation sind z. B. die Bewertung von Schul- und Unterrichtsqualität, die Qualitätssicherung und -entwicklung sowie die Überprüfung und ggf. Weiterentwicklung der pädagogischen Ziele und des Schulprogramms ausgewiesen. Die Verantwortung für interne Evaluationen liegt bei der einzelnen Schule, Vorgaben zu Art, Umfang oder zeitlichen Abständen sind meist nicht im Schulgesetz spezifiziert. Die Länder machen in der Regel keine verbindlichen Vorgaben zum Turnus der Evaluationen oder zu einzusetzenden Instrumenten.

Obwohl interne Evaluationen nicht in allen Ländern verpflichtend sind, stellen alle Länder ihren Schulen – in der Regel über die Landesinstitute – Instrumente wie Feedback- und Befragungstools, Fragebögen, Selbstreflexionsbögen, Apps bzw. Portale für die interne Evaluation zur Verfügung (vgl. Kapitel 1 und Anhang 3).¹⁸ Die meisten der den Schulen zur Verfügung gestellten Instrumente erfüllen die Kriterien einer systematischen Validierung nicht. Eine Ausnahme stellt Baden-Württemberg dar, wo ein Instrument zur Unterrichtsbeurteilung unter Einbezug einschlägiger Wissenschaftler:innen entwickelt und systematisch validiert wurde.

Seit 2018 besteht das länderübergreifende Kooperationsnetzwerk *Länderkooperation Interne Evaluation*, in dem mittlerweile alle Länder vertreten sind. Die Kooperation umfasst neben regelmäßigen Austauschtreffen auch den Austausch und die gemeinsame Entwicklung von Evaluationsinstrumenten (in der Regel Fragebögen), den Austausch zu Angeboten der Beratung und Unterstützung schulischer Akteure, Öffentlichkeitsarbeit sowie die gemeinsame Nutzung und Weiterentwicklung von Befragungsplattformen (z. B. basierend auf *SEP-SCHULE* und *SEP-Klassik* vom Institut für Schulqualität des Landes Berlin – ISQ). Einschlägige Wissenschaftler:innen sind an dem Netzwerk punktuell durch Vorträge etc. beteiligt. Zwei Länder gestalten im Rahmen des Netzwerks auch gemeinsame Qualifizierungsmaßnahmen zur internen Evaluation.

Häufig bieten die Landesinstitute zusätzlich auch Beratung zu Evaluationsmaßnahmen und -instrumenten, teils auch Fortbildungsangebote, an. Die Evaluationsportale und -materialien sind meist bausteinartig aufgebaut, sodass Schulen die Instrumente entsprechend ihrem Bedarf und Erkenntnisinteresse zusammenstellen können, etwa um Rückmeldung zu ausgewählten Aspekten von Schulqualität (z. B. Ganzttag, Inklusion, Gesundheit) zu erhalten. Die zur Verfügung gestellten Instrumente ermöglichen darüber hinaus auch Rückmeldungen zum eigenen professionellen Handeln für die verschiedenen an Schule beteiligten Gruppen (Schulleitungen, Lehrkräfte, Seminarleitungen, Schulaufsicht). Auch für die Zusammenarbeit mit Eltern können diese Daten genutzt werden. Zur Validität der Instrumente liegen allerdings nur vereinzelt Belege vor.

Nutzung von Lernstandserhebungen/VERA für die Schulentwicklung

Laut SWK-Länderabfrage Schule erhalten in allen 16 Ländern Schulleitungen Rückmeldungen aus den VERA-Erhebungen. Dabei erfolgt die Rückmeldung in 10 Ländern auf Individualebene, in 16 Ländern auf

¹⁸ In Thüringen werden die zur Verfügung stehenden Methoden und Instrumente laut Website derzeit (Stand: Februar 2026) überarbeitet und sind nicht abrufbar.



Klassenebene, in 12 Ländern auf Jahrgangsebene und in 13 Ländern auf Landesebene. 9 Länder geben an, adjustierte Vergleiche durchzuführen, ohne hierzu nähere Angaben zu machen. Die Rückmeldung erfolgt in 8 Ländern über ein Datenblatt, in 3 Ländern per Dashboard, in 13 Ländern über ein Berichtsformat und in 4 Ländern in Form eines Gesprächs. Einzelne Länder geben noch weitere Formate der Rückmeldung an wie Ergebnisworkshops (auf Anfrage) oder Diagramme. In den meisten Ländern ist die Rückmeldung zu den VERA-Ergebnissen mit Hinweisen zur Unterrichtsentwicklung bzw. zur Lernförderung verknüpft (vgl. Kapitel 1).

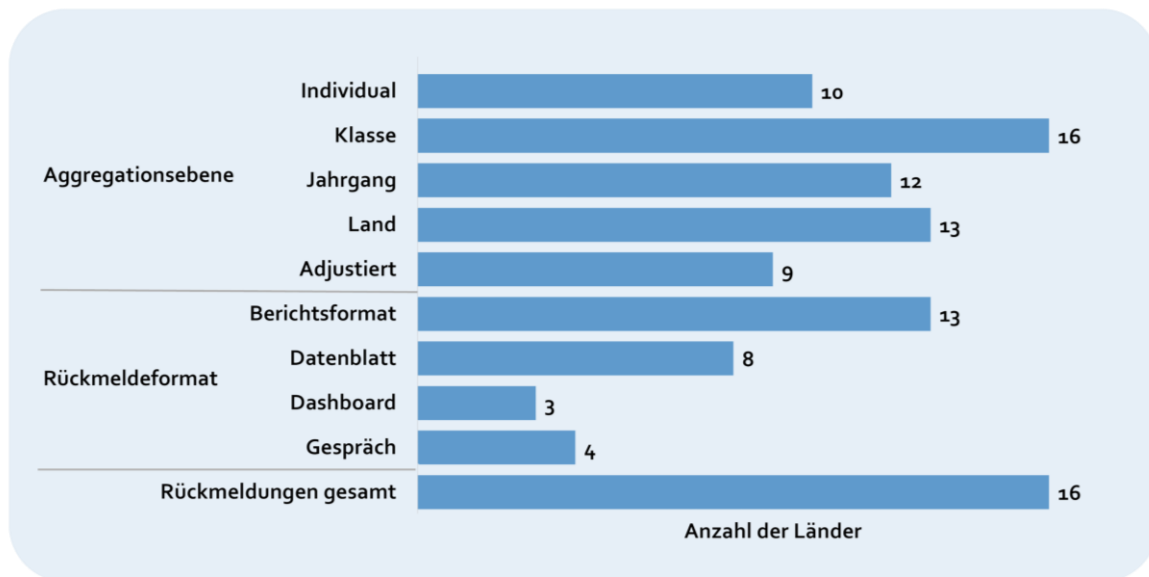


Abbildung 4: VERA-Rückmeldungen an Schulleitungen
Quelle: SWK-Länderabfrage Schule

Externe Evaluation und Schulentwicklung

In ihrer Gesamtstrategie zum Bildungsmonitoring hat die KMK (2015) externe Evaluation als ein Instrument der Schulentwicklung bezeichnet. Externe Evaluationen oder Schulinspektionen, die auf der Grundlage ländereigener Qualitätsrahmen die schulische Arbeit evaluieren, sollen Schulen vor allem mit Daten zu den Prozessen schulischer Arbeit versorgen (Tarkian, Lankes & Thiel, 2019). Dazu geben sie Rückmeldungen zu Stärken und Schwächen von Schulen und setzen idealerweise Impulse für deren weitere Entwicklung (Informationen zur Qualitätsanalyse, Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2025b). Allerdings wurde die externe Evaluation inzwischen in einer ganzen Reihe von Ländern vor allem aus Validitäts- und/oder Kostengründen wieder abgeschafft bzw. auf ein anlassbezogenes Verfahren umgestellt (vgl. Kapitel 3.1.2).

Bezüglich der Akzeptanz und Nützlichkeits einschätzung der externen Evaluation zeigt eine Evaluation der Berliner Inspektion von Preuße, Pohl und Gärtner (2019), dass 40 Prozent der befragten Schulleitungen die Aussage ablehnen, Schulinspektion generiere Wissen für die Schulentwicklung. 74 Prozent der Schulleitungen sind außerdem der Meinung, dass die Ergebnisse keine neuen Informationen für die Schulentwicklung böten. Obwohl aus Sicht der Schulleitung Schulinspektion zwar den schulinternen Austausch intensiviert, scheint sie eher als Instrument der Kontrolle und Rechenschaftslegung wahrgenommen zu werden (Preuße



et al., 2019). Dazu passt der Befund, dass häufig keine Unterstützungsmaßnahmen auf eine Rückmeldung durch die Schulinspektion folgen (Gärtner, 2021). Studien aus Berlin, Sachsen und Hessen verweisen außerdem gleichlautend darauf, dass aus Sicht der Schulen keine systematische Zusammenarbeit zwischen Schule und Schulaufsicht existiert (Drinck et al., 2013; Keune, 2014; Preuße et al., 2019).

Gärtner (2021) fasst den Forschungsstand zu Wirkungen der Schulinspektion wie folgt zusammen: „Für die große Mehrzahl der Schulen können [...] bisher keine nachhaltigen Entwicklungseffekte aufgrund einer Schulinspektion nachgewiesen werden“ (Gärtner, 2021, S. 183). Angesichts „der beträchtlichen finanziellen und personellen Ressourcen“ (Gärtner, 2016, S. 64), die für Schulinspektionen bereitgestellt werden, werfen die offensichtlich nur eingeschränkt wahrgenommene Nützlichkeit der Schulinspektion für schulische Entwicklung sowie fehlende Wirksamkeitsnachweise Fragen der Effektivität und Effizienz auf. Problematisch erscheint auch, dass weder Unterstützungsprozesse im Anschluss an eine Inspektion noch die Verknüpfung von Schulinspektion und Schulaufsicht (vgl. Kapitel 3) zufriedenstellend gelungen sind.

2.1.2 Nutzung von Daten für Ziel- und Leistungsvereinbarungen mit der Schulaufsicht

Einige Länder haben in den letzten Jahren datengestützte Dialoge und Zielvereinbarungen zwischen Schulen und Schulaufsichten auf der Grundlage von Daten aus der internen und externen Evaluation sowie Daten aus Lernstandserhebungen und weiteren Daten z. B. zu Abschlussquoten, Absentismus und Drop-out eingeführt. Diese Instrumente werden in Kapitel 3 ausführlich dargestellt.

2.1.3 Strukturen, Qualifikation und Unterstützungsmaßnahmen für datengestützte Schulentwicklung

Funktionsstellen für datengestützte Schulentwicklung

In allen Ländern bestehen – zumindest an weiterführenden Schulen¹⁹ – neben der Schulleitung weitere Positionen, die im Rahmen der erweiterten Schulleitung oder des mittleren Managements für besondere Funktionen eingerichtet wurden. Das gilt etwa für die stellvertretende Schulleitung oder die Leitung eines Fachbereichs bzw. einer Fachkonferenz. Bremen, Hamburg, Hessen, Niedersachsen, das Saarland und Schleswig-Holstein waren die ersten Länder, die die Möglichkeit einer funktionsbezogenen Ausschreibung als gezieltes Instrument der Personalentwicklung im Rahmen erweiterter schulischer Gestaltungsspielräume von Schulen einführten (Thiel et al., 2022). Inzwischen sind in allen Ländern Funktionsstellen unterhalb der Schulleitung mit spezifischen Funktionen vorgesehen. In Grundschulen bestehen allerdings deutlich weniger Funktionsstellen (teilweise gar keine außer der Schulleitung) als in weiterführenden Schulen (SWK, 2022a). Neben Funktionen wie Fachleitung oder Koordination schulischer Aufgaben wie Inklusion ist in Berlin auch explizit eine Funktionsstelle für Qualitätsentwicklung als „Studiendirektorin oder Studiendirektor bzw. Sekundarschulrektorin oder Sekundarschulrektor zur Koordinierung schulfachlicher Aufgaben

¹⁹ An sehr kleinen Grundschulen gibt es nicht immer eine stellvertretende Schulleitung; in einigen Ländern sind hierfür Schwellenwerte definiert (z. B. „ab 180 Schüler:innen“), in anderen bestehen keine festen Vorgaben.



(Schulqualität/Schulprogramm/Schulentwicklung/Evaluation – Qualitätsbeauftragte oder Qualitätsbeauftragter)²⁰ vorgesehen.

Unterstützungssysteme und Fachberatung

In allen Ländern existieren Unterstützungssysteme für die datengestützte Schul- und Unterrichtsentwicklung in Form von Fachberatung, Fortbildungs- und Qualifizierungsangeboten sowie Handreichungen. Die Fachberatung ist meist an den Landesinstituten und Qualitätseinrichtungen, teils auch an den Kultusministerien oder nachgeordneten Behörden (z. B. Bezirksregierung, Schulämter) angesiedelt. Ihre Aufgaben umfassen die Beratung von Fachschaften zur Unterrichtsentwicklung (vgl. Kapitel 1) in den jeweiligen Fächern sowie die Durchführung fachdidaktischer Fortbildungen. Weiterhin unterstützen die Fachberatungen zum Teil auch bei der Materialentwicklung sowie der Auswertung von Lernstandserhebungen. Ebenfalls sind in allen Ländern Angebote der Schulentwicklungsberatung eingerichtet (Dedering et al., 2022) und an den Landesinstituten und Qualitätseinrichtungen verankert (vgl. Kapitel 3.1.3). Häufig sind jedoch die Aufgaben von Fachberatung und Schulentwicklungsberatung nicht klar getrennt, und die Fachberatung bezieht sich oft auch auf fachübergreifende Themen wie z. B. Inklusion, Medienbildung, Ganzttag oder auch Themen der Schulentwicklung, wohingegen eine im engeren Sinne fachdidaktisch ausgerichtete Fachberatung zu spezifischen Fachinhalten nicht in hinreichendem Maße zur Verfügung zu stehen scheint. Insgesamt besteht eine ungleiche Verteilung der Ressourcen zwischen Schulentwicklungs- und Fachberatung (vgl. dazu ausführlicher Kapitel 3).

Qualifizierung und Fortbildung

Für die Zielgruppe der *Lehrkräfte* haben 14 Länder in der SWK-Länderabfrage Schule angegeben, dass Fortbildungen und Qualifikationen zur datengestützten Schulentwicklung angeboten werden. Eine Teilnahmeverpflichtung besteht hierbei in der Regel nicht. Zwei Länder haben angegeben, dass keine entsprechenden Fortbildungen oder Qualifikationsangebote vorgesehen sind.

Für *Schulleitungen* sind Fortbildungen und Qualifikationsangebote zur datengestützten Schulentwicklung in 14 Ländern vorgesehen. In fünf Ländern besteht für Schulleitungen eine Verpflichtung zur Teilnahme an entsprechenden Fortbildungen im Rahmen der Schulleitungsausbildung. Darüber hinaus sind die Angebote für Schulleitungen freiwillig.

Für *Steuergruppen, Qualitätsteams o. Ä.* sind in 13 Ländern Fortbildungsangebote zur datengestützten Schul- und Unterrichtsentwicklung vorgesehen. Eine Teilnahmeverpflichtung besteht für diese Zielgruppe in keinem Land.

²⁰ <https://www.berlin.de/sen/bildung/schule/rechtsvorschriften/verwaltungsvorschriften/vv-zuordnung.pdf?ts=1738768871> (aufgerufen am 20.02.2026).

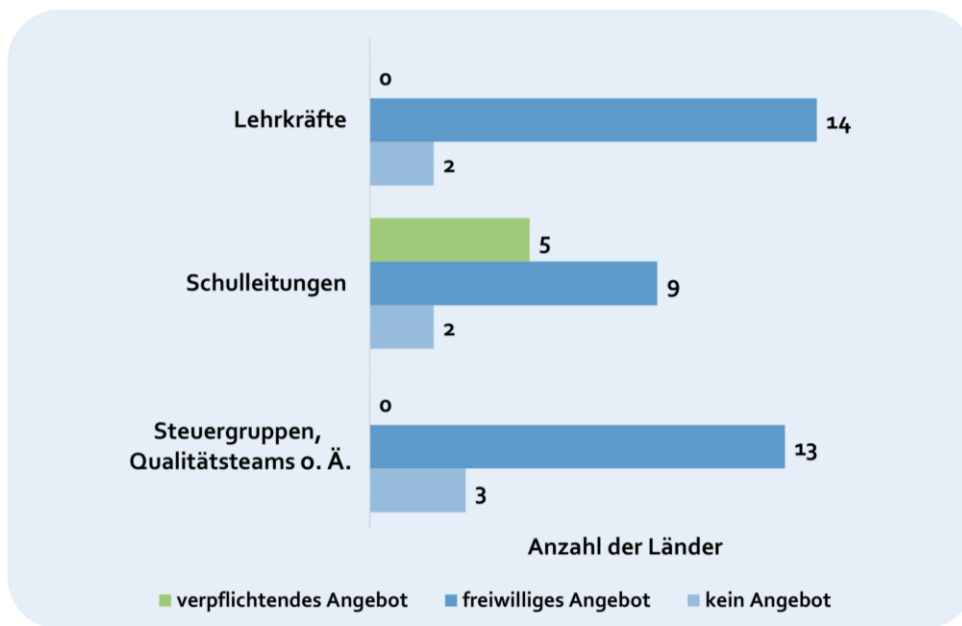


Abbildung 5: Angebote zur Fortbildung und Qualifikation zur datengestützten Schulentwicklung
Quelle: SWK-Länderabfrage Schule

2.2 Konzepte und Maßnahmen

Schulische Gestaltungsspielräume können dann in eine effektive Schulentwicklung und schließlich in Lernzuwächse von Schüler:innen münden, wenn Verantwortlichkeiten klar definiert und Schulleitungen qualifiziert sind. Führungs- und Managementkompetenzen sind dabei ausschlaggebend (Jackson, 2023). Diese Befunde werden durch eine Reihe von Metaanalysen gestützt, die mittlere (Hattie, 2009; Wu & Shen, 2022) bis hohe (Tan et al., 2022) Effekte schulischer Führung auf Schülerleistungen belegen. Datengestützte Entscheidungsfindung scheint dabei ein zentrales Merkmal erfolgreicher Schulleitung zu sein (Adanne, 2024; Schildkamp et al., 2013). Neben der Schulleitung sind kollaborative Strukturen von zentraler Bedeutung für eine datengestützte Schulentwicklung. Dabei geht es nicht nur um die Abstimmung schulweiter Ziele und Maßnahmen im Gesamtkollegium, sondern auch um die regelmäßige Abstimmung zwischen Fachlehrkräften oder Jahrgangsteams (Kelchtermans, 2006; Wurster & Richter, 2016).

Im Folgenden wird zunächst diskutiert, welche Daten für eine gezielte Schulentwicklung relevant sind (2.2.1). Vor diesem Hintergrund werden Befunde zu den beiden zentralen Instrumenten der datengestützten strategischen Schulentwicklung, Schulentwicklungsplänen (2.2.2) und Tools für Selbstevaluation (2.2.3), dargestellt. Abschließend wird kurz auf Unterstützungsstrukturen für eine datengestützte Schulentwicklung eingegangen (2.2.4); eine ausführliche Darstellung von Unterstützungsstrukturen findet sich in Kapitel 3.2.2.



2.2.1 Daten für strategische Schulentwicklung

Für eine strategische datengestützte Schulentwicklung durch Schulleitungen und Kollegien sind systematisch erfasste valide Daten erforderlich (Schildkamp, 2019). Sie bilden die Grundlage für eine Bestandsaufnahme und die Identifikation von Handlungsbedarf und ermöglichen die Überprüfung der Maßnahmen der Schulentwicklung. Diese Daten bilden das Verständnis von Bildungsqualität ab, an dem die Arbeit von Schulen gemessen wird. Ihre Erfassung und Nutzung für Schulentwicklung und Rechenschaftslegung (vgl. Kapitel 3) ist deshalb Gegenstand von Kontroversen.

Nutzung von Leistungsdaten für Schulentwicklung

In vielen OECD-Ländern wurden Schulen lange Zeit nahezu ausschließlich an den Lernzuwächsen ihrer Schüler:innen in Kernfächern gemessen. Studien zu Effekten einer systematischen Datennutzung durch Schulen liegen entsprechend vor allem für standardisierte Schulleistungstests vor. Diese Befunde geben Hinweise darauf, dass eine gezielte Nutzung von Daten durch die gesamte Schule Schülerleistungen verbessern kann. In den Niederlanden wurde die schulweite Nutzung von Testdaten auf der Grundlage eines Rahmenmodells implementiert und in mehreren schulbezogenen Interventionsstudien evaluiert (Visscher, 2021). Teilweise deutliche Effekte konnten für drei Kompetenzdomänen (Mathematik, Lesen, Rechtschreibung) gezeigt werden. Den größten Lernzuwachs hatten Schüler:innen mit niedrigem oder hohem sozioökonomischen Hintergrund. Effekte wurden in allen Jahrgangsstufen (1 bis 8) gemessen (Keuning et al., 2019; van der Scheer & Visscher, 2016; van Geel et al., 2016).

Drei Interventionsstudien zur Leseförderung, die nach einem ähnlichen Konzept in Neuseeland unter dem Titel *Learnings School Model* durchgeführt wurden, kamen zu ähnlichen Ergebnissen. Teilweise wurden drei bis vier Monate zusätzlicher Lernfortschritt pro Schuljahr gemessen (Lai et al., 2009; McNaughton et al., 2012).

In Deutschland liegen mit den Vergleichsarbeiten zentrale Lernstandserhebungen für Kernfächer vor, die auf Bildungsstandards beruhen und Kompetenzstufen ausweisen. Sie wurden auf der Grundlage fachdidaktischer und lernpsychologischer Theorien operationalisiert und explizit für Zwecke der Schul- und Unterrichtsentwicklung implementiert (vgl. Kapitel 1). Zwar liegen zahlreiche Publikationen zu Akzeptanz und Nutzung von Vergleichsarbeiten vor, mit eher ernüchternden Befunden, aber kaum Studien, die Effekte der Nutzung von Vergleichsarbeiten für systematische Schulentwicklung auf Schülerleistungen untersuchen. Eine Ausnahme bildet eine Interventionsstudie zur Nutzung von VERA-Daten zur Förderung von Lesekompetenz (Karst et al., 2022). Sie konnte zwar keinen generellen Effekt auf die Leseleistung zeigen. Allerdings weisen die Befunde darauf hin, dass die Arbeit mit Vergleichsarbeitsdaten Lehrkräfte darin unterstützt, mit den Leistungsunterschieden in der Klasse effektiv umzugehen und allen Schüler:innen im Leistungsspektrum Lernfortschritte zu ermöglichen.

Auch Analysen mit PISA-Daten zeigen für mehrere Länder, dass Schüler:innen in Schulen, die Leistungen ihrer Schüler:innen durch Lernstandserhebungen systematisch erfassen, höhere Leistungen im Lesen aufweisen (OECD, 2020).



Nutzung von Daten zu sozial-emotionalem Lernen für Schulentwicklung

Neben Fachleistungen werden inzwischen in einigen Ländern auch Daten zu sozial-emotionalen Merkmalen für die Schulentwicklung genutzt. Die Förderung sozial-emotionaler Kompetenzen (vgl. Kapitel 1) erfolgt in evidenzbasierten Programmen häufig als systemischer schulweiter Ansatz. Seitens der OECD (2023) wird entsprechend empfohlen, Schulen als *Hubs for Social and Emotional Learning* weiterzuentwickeln und soziale und emotionale Merkmale zu einem integralen Bestandteil schulischer Assessments zu machen. Allerdings zeigen OECD-Daten für viele Länder, dass bei der praktischen Verankerung nicht-kognitiver Bildungsziele in der Schule noch erheblicher Nachholbedarf besteht (OECD, 2021).

In den USA wurden die Bundesstaaten im Jahr 2015 durch den *Every Student Succeeds Act (ESSA)*²¹ dazu verpflichtet, Maße, die sozial-emotionales Lernen oder Schulkultur/Schulklima abbilden, in das schulbezogene Assessment einzubeziehen. Das California Department of Education formulierte 2016 entsprechend die Erwartung an die Schuldistrikte, Indikatoren und valide Messinstrumente für *Accountability* und *Improvement* zu entwickeln. Zu diesem Zweck haben sich mehrere Distrikte zum sogenannten *CORE-Projekt*²² zusammengeschlossen, das auf der Grundlage eines vier Konstrukte umfassenden Modells sozial-emotionaler Kompetenzen (*Growth-Mindset, Self-Efficacy, Self-Management, Social Awareness*) sowie eines ebenfalls vier Konstrukte umfassenden Modells für Schulkultur/-klima (*Support for Academic Learning, Sense of Belonging and School Connectedness, Knowledge and Perceived Fairness of Discipline Rules and Norms, Safety*) Instrumente in Kooperation mit Universitäten entwickelte und sorgfältig validierte (Hough et al., 2017). Die Konstrukte werden aus Schüler-, Eltern- und Lehrerperspektive erfasst.

Sozial-emotionale Kompetenzen variieren deutlich zwischen den Klassen einer Schule sowie zwischen Schulen (Loeb et al., 2019; R. H. Meyer et al., 2019). Die Erfassung sozial-emotionaler Kompetenzen ist allerdings mit einigen methodischen Herausforderungen verbunden (vgl. Kapitel 1.2.4), sodass von einer Integration dieser überfachlichen Maße in ein Accountabilitysystem derzeit Abstand genommen werden sollte (Hough et al., 2017). Das bedeutet allerdings nicht, dass Daten zu sozial-emotionalen Kompetenzen für die Schulentwicklung wertlos sind. Sie können für die Formulierung von Zielen, Prioritäten und die Entscheidung für bedarfsorientierte Maßnahmen sowie für die Ressourcenallokation im Rahmen der Schulentwicklung sehr nützlich sein (M. R. West et al., 2020). Sie sollten aber als „flashlight, not as hammer“ (Gehlbach & Hough, 2018) genutzt werden. Für Zwecke der Schulentwicklung ist es in der Regel sinnvoller, anstelle sozial-emotionaler Kompetenzen Merkmale sozial-emotionalen Lernens wie soziale Zugehörigkeit oder schulische Zufriedenheit zu erfassen. Das gilt auch für die Integration entsprechender Daten in Dashboards für Schulaufsichten (vgl. Kapitel 3).

²¹ <https://www.ed.gov/laws-and-policy/laws-preschool-grade-12-education/every-student-succeeds-act-essa> (aufgerufen am 20.02.2026).

²² <https://coredistricts.org/data-collaborative/> (aufgerufen am 20.02.2026).



Nutzung standardisiert erfasster Prozessdaten für Schulentwicklung

Neben Outcome-Daten benötigen Schulen für eine effektive Schulentwicklung auch Daten über die Qualität von Prozessen. Dies betrifft einerseits Daten zur Qualität des Unterrichts (vgl. Kapitel 1), andererseits Daten zur Schulqualität.

Was die Erfassung von Prozessen auf der Schulebene betrifft, hat Ditton bereits im Jahr 2000 den Forschungsstand zur Schul- und Organisationsentwicklung in einem gut begründeten theoretischen Rahmenmodell mit vier Dimensionen verdichtet: *Management/Führung*, *Professionalisierung/Personalentwicklung*, *Kooperation/Abstimmung* und *Schulkultur* (Ditton, 2000). Dieses Modell erlaubt eine Operationalisierung relevanter Prozessvariablen und hat Eingang in die meisten Qualitätsrahmen der Länder gefunden (Thiel & Tarkian, 2019). Was die Erfassung der einzelnen Konstrukte betrifft, gilt auch hier, dass die vorliegenden Skalen für Forschungszwecke entwickelt und validiert wurden (Gerecht et al., 2007) und sich für den Zweck der Schulentwicklung nur bedingt eignen. Ein explizit für die Selbstevaluation im Rahmen der Schulentwicklung entwickeltes Instrument, für das erste Validitätshinweise vorliegen (Gärtner & Brauckmann, 2014), wurde vom Berliner ISQ entwickelt. Das Instrument ist im Selbstevaluationsportal des ISQ²³ verfügbar und wird bereits in einigen anderen Ländern eingesetzt (vgl. Anhang 3).

Nutzung nicht-standardisierter Daten für Schulentwicklung

Hamilton et al. (2009) empfehlen in einer Leitlinie für das *What Works Clearinghouse* auch die kollaborative Nutzung von Daten aus Klassenarbeiten, Projektarbeiten oder Beobachtungen im Klassenzimmer für Zwecke der Schulentwicklung (vgl. auch Schildkamp, 2019). Diese Daten können einfach erfasst werden. Der Nachteil ist, dass Erwartungshorizonte und Bewertungsmaßstäbe der Lehrkräfte stark variieren. Diese Discrepanzen sollten deshalb explizit zum Gegenstand kollaborativer Reflexion gemacht werden, mit dem Ziel der Abstimmung von Standards und der Kalibrierung von Bewertungen. Kalibrierung bedeutet z. B., dass dieselben Arbeiten von Schüler:innen im Team bewertet und Abweichungen systematisch diskutiert oder dass Unterrichtsbeobachtungen systematisch verglichen werden (Wayman et al., 2017). Wenn dieselben Klassenarbeiten und Tests bei allen Lehrkräften der Jahrgangsstufe eingesetzt werden und eine sorgfältige Kalibrierung erfolgt, können diese Daten auch für die Schulentwicklung genutzt werden.

2.2.2 Schulentwicklungspläne als Instrumente strategischer Schulentwicklung

Datengestützte Schulentwicklung beruht wie datengestützte Unterrichtsentwicklung (vgl. Kapitel 1) auf einer kontinuierlichen zyklischen Abfolge von datengestützter Bestandsaufnahme, Ableitung konkreter Entwicklungsziele, Identifikation und Implementation von Maßnahmen sowie schließlich deren Überprüfung (van Geel et al., 2017).

²³ https://sep-klassik.isq.berlin/de_DE/start/index.html (aufgerufen am 20.02.2026).



Die zentralen Instrumente bzw. Tools für die datengestützte Schulentwicklung sind auf der einen Seite Schulentwicklungspläne und auf der anderen Seite entsprechende Daten bereitstellende Tools zur Bestandsaufnahme und Überprüfung der implementierten Maßnahmen (Selbstevaluation).

Datengestützte Schulentwicklung erfolgt als Prozess, an dessen Beginn eine Bestandsaufnahme steht, auf deren Grundlage spezifische, messbare, akzeptierte, realistische und terminierte Ziele formuliert werden, die wiederum die Entscheidung für – im Idealfall evidenzbasierte – Maßnahmen zur Zielerreichung orientieren. In diesem Sinne wurden Schulentwicklungspläne in vielen OECD-Ländern als strategisches Entwicklungstool im Zusammenhang mit den Bildungsreformen um das Jahr 2000 eingeführt (Dunaway et al., 2012). Über die Dokumentation datengestützter Ziele und evidenzbasierter Maßnahmen hinaus fixieren Schulentwicklungspläne auch Verantwortlichkeiten und beschreiben benötigte Ressourcen sowie Indikatoren, an denen die Zielerreichung überprüft werden kann. Schulentwicklungspläne werden in regelmäßigen Abständen datengestützt fortgeschrieben.

Der Prozess der Zielsetzung spielt für die Planung der Schulentwicklung eine zentrale Rolle (Mandinach et al., 2008; Schildkamp, 2019). Ohne klare, auf Grundlage einer datengestützten Bestandsaufnahme definierte Ziele kann weder eine begründete Entscheidung für Maßnahmen erfolgen, noch kann der Erfolg der Maßnahmen überprüft werden. In den Niederlanden werden Schulentwicklungsziele für ein Schuljahr formuliert (Meyers & van Gronigen, 2019).

Die internationale Befundlage zu Effekten von Schulentwicklungsplänen ist eher schmal und zudem widersprüchlich. Während in einigen Studien Zusammenhänge zwischen Schulentwicklungsplänen und der Qualität von Prozessen bzw. Leistungen von Schüler:innen gezeigt werden konnten (Ettinger, 2015; Fernandez, 2011; Reeves, 2011), beschreiben andere Autor:innen Schulentwicklungspläne als symbolische, weitgehend wirkungslos bleibende Reaktionen auf administrative Erwartungen (Kruse, 2001; Mintrop & MacLellan, 2002; O'Day, 2002). Offensichtlich ist, dass die Qualität der Entwicklungsplanung entscheidend ist für ihre Effekte (Caputo & Rastelli, 2014; Fernandez, 2011). In der Praxis weisen viele Pläne allerdings eine geringe Qualität auf und sind ungeeignet für strategische Entwicklung (Meyers & van Gronigen, 2019; Strunk et al., 2016; vgl. die Situationsanalyse zu Schulprogrammen in Deutschland in Kapitel 2.1.1).

Folgende Aspekte scheinen für eine effektive Nutzung von Schulentwicklungsplänen entscheidend zu sein:

- datengestützte Bestandsaufnahme von Handlungsbedarfen und deren Priorisierung (Escobar, 2019); vor diesem Hintergrund muss eine klare Zielformulierung erfolgen (Huber & Conway, 2015),
- Alignment der Schulentwicklungspläne mit übergeordneten bildungspolitischen Zielen bei gleichzeitiger Fokussierung auf schulspezifische Herausforderungen (Escobar, 2019; Klopsch & Sliwka, 2020),
- Herstellung von Commitment aller Lehrkräfte zum Schulentwicklungsplan durch breite Diskussion von Bestandsaufnahme, Zielsetzung, Maßnahmen und Erfolgsindikatoren im Kollegium (Dunaway et al., 2012),
- Entwicklung von Routinen der kontinuierlichen Schulentwicklung durch regelmäßige, mindestens jährliche, datengestützte Fortschreibung von Schulentwicklungsplänen (Dunaway et al., 2012; van Gronigen & Meyers, 2021),



- Unterstützung einer strukturierten und fokussierten Planung durch (digitale) Templates (Ettinger, 2015; Fernandez, 2011). Ein Beispiel für ein digitales Template für Schulentwicklungsplanung, das eine datengestützte Begründung und Priorisierung von Zielen und Maßnahmen unterstützt, wurde in Österreich entwickelt.²⁴ Ein anderes Beispiel stammt aus Alberta/Kanada.²⁵

2.2.3 Selbstevaluation als Instrument strategischer Schulentwicklung

Selbstevaluation leistet einerseits einen zentralen Beitrag zur Bestandsaufnahme für die Schulentwicklungsplanung und andererseits zur Überprüfung der im Schulentwicklungsplan dokumentierten Ziele. Entsprechend ist Selbstevaluation ein zentraler Bestandteil der Qualitätsentwicklung in der Verantwortung der Schule. Im Unterschied zu externer Evaluation (vgl. Kapitel 3) liegt die Verantwortung für Selbstevaluation in den Händen der Schulen. Schulische Selbstevaluation bzw. interne Evaluation mittels selbst ausgewählter oder selbst entwickelter Instrumente genießt bei Lehrkräften im Vergleich zu Testdaten eine höhere Akzeptanz (Wurster et al., 2013). Allerdings artikulieren einige Forschende teilweise erhebliche Bedenken hinsichtlich der Validität der Daten aus interner Evaluation (Cremers-van Wees et al., 1996; Faddar et al., 2017; Kyriakides & Campbell, 2004; Scriven, 1991).

Voraussetzungen wirksamer Selbstevaluation

Nach der Einführung von Selbstevaluation als Instrument der Qualitätssicherung in vielen Ländern (OECD, 2015) wurden seit Ende der 1990er-Jahre zahlreiche Studien durchgeführt, um den Impact von Selbstevaluation auf Schülerleistungen und die Qualität des Unterrichts, des Schulmanagements oder der Lehrkräftekooperation zu erfassen. Neben der Notwendigkeit einer systematischen Verschränkung von Zielformulierung, Planung und Evaluation (Caputo & Rastelli, 2014; Chapman & Sammons, 2013) können folgende Voraussetzungen für eine wirksame Selbstevaluation identifiziert werden:

- Implementation evidenzbasierter, auf Kernfaktoren fokussierter Qualitätsstandards bzw. Frameworks (Chapman & Sammons, 2013; Demetriou & Kyriakides, 2012; Hutchinson & Young, 2011), die für schulische Akteur:innen verständlich sind (van der Bij et al., 2016). Diese Rahmenmodelle sowie Standards dienen als Grundlage für Selbstevaluation und sind darüber hinaus eine Voraussetzung für ein Alignment von Maßnahmen innerhalb der Schule sowie über die Ebenen des Schulsystems hinweg.
- Verwendung validierter und gleichermaßen praxistauglicher Evaluationsinstrumente (Hendriks et al., 2001; Schildkamp et al., 2009; van der Bij et al., 2016). Der Einsatz standardisierter Instrumente kann die Belastung von Schulen durch interne Evaluation (Visscher, 2002) reduzieren und gleichzeitig die Validität der Daten erhöhen. Wenn nicht-standardisierte Daten für Schulentwicklung verwendet werden, muss eine kollaborative Kalibrierung erfolgen (Danzig et al., 2013).

²⁴ https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.qms.at%2Fimages%2FSchulentwicklungsplan_2-o.docx&wdOrigin=BROWSELINK (aufgerufen am 20.02.2026).

²⁵ <https://dalhousie.cbe.ab.ca/documents/24e14e5f-02b4-4264-bff0-8527e0591825/School-Development-Plan-2024-25.pdf> (aufgerufen am 20.02.2026).



- Aufbau schulischer Kapazitäten zur Erfassung, Auswertung und Interpretation von Daten (Timperley, 2013), einerseits durch eine stärkere Ausrichtung der Qualifizierung von Lehrkräften auf das professionelle Leitbild datengestützter Entscheidung, andererseits durch Ausweisung spezifischer Funktionen, Stellenprofile bzw. Positionen (Farrell, 2015; Gallimore et al., 2009). Dazu gehört auch ein entsprechendes Verständnis datengestützter Schulleitung (Chapman & Sammons, 2013; Nelson et al., 2015). In Deutschland spielen Fachkonferenzen eine wichtige Rolle (Wurster & Richter, 2016).
- Bereitstellung von Unterstützungsstrukturen innerhalb der Schule wie Zeitressourcen für kollegiale Datenauswertung und Reflexion (Hamilton et al., 2009; Hejtmanek et al., 2024; Jimerson et al., 2021; Lai & McNaughton, 2013; Marsh et al., 2010) sowie externe Unterstützung wie Coaching und Beratung. Dabei erscheint es wichtig, dass Unterstützungsangebote nicht auf Schulleitungen und Managementfunktionen beschränkt bleiben, sondern direkt bei Lehrkräften bzw. Fachkonferenzen ansetzen, wenn Unterricht wirksam verbessert werden soll (Caputo & Rastelli, 2014; Chapman & Sammons, 2013; Timperley et al., 2009; Visscher, 2021; vgl. Kapitel 1).
- Entwicklung einer schulischen Evaluationskultur (Keuning et al., 2017; Nelson et al., 2015), die dem viel bestätigten Befund Rechnung trägt, dass die wahrgenommene Nützlichkeit von Daten der wichtigste Prädiktor der Datennutzung ist (Wurster et al., 2013). Diese Kultur sollte vom Leitbild einer *lernenden Organisation* (einschließlich einer fehlerfreundlichen, wertschätzenden Feedbackkultur) geprägt sein, in der die Überprüfung der eigenen Arbeit im Sinne einer Routine ein selbstverständlicher Anlass für ihre Weiterentwicklung ist (Hofman et al., 2009) und Daten nicht defensiv abgewertet werden und in Schuldzuweisungen münden (Hejtmanek et al., 2024). Dazu ist es erforderlich, dass Daten aus Selbstevaluation nicht als Instrument der Kontrolle, sondern der Entwicklung verstanden werden (Brown et al., 2021; Demski, 2017).

Datenfeedbacksysteme bzw. Datendashboards für datengestützte Entwicklung

Durch digitale Datenfeedbacksysteme kann nicht nur schulische Selbstevaluation wirksam unterstützt werden, diese Systeme ermöglichen vielmehr jederzeit den unkomplizierten Rückgriff auf Daten für Entscheidungen und tragen dazu bei, datengestützte Schulentwicklung als kontinuierlichen Prozess im Schulalltag zu verankern. Sie stellen Schulen regelhaft erfasste Testdaten, administrative Daten wie Fehlzeiten oder Absentismus und Daten aus zentralen Befragungen zur Schulqualität für eine strategische Entwicklung zur Verfügung.

Neben den USA, UK und Australien wurden insbesondere in den Niederlanden, lange vor der Einführung von Datendashboards, papier- und später computerbasierte *School-Performance-Feedback-Systeme (SPFS)* entwickelt und implementiert (Visscher & Coe, 2003). In Deutschland wurde die Forschung und Entwicklung von SPFS maßgeblich von Andreas Breiter an der Universität Bremen in Kooperation mit der Universität Twente betrieben (Breiter & Light, 2006).

Aus dem Forschungsstand zu SPFS und den wenigen Studien zu schulbezogenen Datendashboards können vier grundlegende Anforderungen an ein effektives Datenfeedbacksystem abgeleitet werden:



- **Fokussierung auf zentrale Konstrukte anstelle einer Rückmeldung möglichst vieler Daten**
Breiter & Light (2006) haben in einem grundlegenden Beitrag zum Design von Informationssystemen Befunde aus internationalen Studien zum datengestützten Entscheiden aufgearbeitet und kommen u. a. zu dem Schluss, dass die Selektion relevanter Daten und der Ausschluss belangloser und irrelevanter Daten wesentliche Voraussetzungen für eine effektive Nutzung sind. Um eine von Entscheider:innen nicht mehr verarbeitbare Datenflut zu vermeiden, muss sorgfältig identifiziert werden, welche Informationen für Nutzer:innen relevant sind. Outcomes sollten entsprechend durch einen klar operationalisierten Begriff von Bildungsqualität definiert werden, der neben kognitiven auch nicht-kognitive Bildungsziele einschließt (vgl. 2.2.1). Das Kriterium für den Bericht von Prozessmerkmalen sollte ihr empirisch belegter Beitrag zur Erreichung dieser Bildungsziele sein (Hendriks et al., 2001).
- **Fokussierung auf aussagekräftige statische Kennwerte und aussagekräftige Vergleiche**
Was aussagekräftige Kennwerte betrifft, geht es um die Rückmeldung von Prozentwerten, Mittelwerten, Streuungsmaßen oder längsschnittlichen Daten. Daten aus Vergleichsarbeiten ermöglichen darüber hinaus die Rückmeldung mit Bezug auf Kompetenzstufen sowie die Erreichung der Bildungsstandards und haben dadurch eine Frühwarnfunktion, was besonderen Unterstützungsbedarf betrifft. Daneben können Vergleiche mit dem Landesmittelwert rückgemeldet werden (Koch, 2011). Der Adjustierung von Daten kommt eine besondere Bedeutung zu (vgl. die Ausführungen zur Adjustierung in Kapitel 1).
Wichtig ist auch, zu entscheiden, auf welcher Aggregationsebene Daten den einzelnen Akteur:innen zurückgemeldet werden. Während Lehrkräfte individuelle Schülerdaten benötigen, um gezielt zu fördern (vgl. Kapitel 1), benötigen Schulleitungen und Fachkonferenzen auf Klassenebene aggregierte Daten, um fachdidaktische und überfachliche Unterstützungsbedarfe bei einzelnen Lehrkräften zu identifizieren sowie Entscheidungen über Ressourcenverwendung zu treffen.
- **Beachtung der Kriterien für eine effektive konzeptionelle und technische Gestaltung eines interaktiven, digitalen Feedbacksystems**
Entscheidend für die technische Gestaltung ist ein Modell zu relevanten Parametern eines effektiven Datenfeedbacks (Pukrop, 2019; Visscher & Coe, 2003). Light et al. (2005) nennen sechs grundlegende Anforderungen an ein Datenfeedbacksystem: (1) einfache Zugänglichkeit; (2) kurze Feedback-Loops bzw. Aktualität des Feedbacks; (3) verständliche grafische Aufbereitung der Daten; (4) Möglichkeit der Datenexploration; (5) Fokussierung auf nutzerrelevante Daten; (6) Verknüpfung mit Hinweisen zur Unterrichtsentwicklung. Pukrop (2019) konnte zeigen, dass ein entsprechendes Feedbacksystem die Verständlichkeit von Daten aus der Sicht schulischer Akteur:innen signifikant beeinflusst. Er nennt folgende Aspekte, die bei einer interaktiven Visualisierung von Daten berücksichtigt werden sollten: Nutzer:innen sollten einen Überblick erhalten (*Overview*), Bereiche von besonderem Interesse in einer größeren Auflösung betrachten können (*Zoom*), (aktuell) nicht interessante Bereiche ausblenden können (*Filter*), einzelne Aspekte oder Gruppen auswählen können (*Details-on-Demand*), Beziehungen zwischen einzelnen Aspekten sichtbar machen können (*Relate*), Aktionen rückgängig machen oder wiederholen können (*History*) sowie Unterkategorien extrahieren können (*Extract*).



Die frühen Datenfeedbacksysteme wurden inzwischen zu online-basierten Dashboards mit interaktiven Displays weiterentwickelt. Dabei sind Dashboards für lernprozessbegleitende Diagnostik sowie Classroom Analytics (vgl. Kapitel 1) zu unterscheiden von *Systems-Level Dashboards*, die Daten aggregiert für Entscheidungen auf der Schulebene oder auch für die Schulaufsicht (vgl. Kapitel 3) zur Verfügung stellen (Curran et al., 2024). Bislang liegen nur wenige Studien zu Systems-Level Dashboards vor. In einem systematischen Review zeigen Curran et al. (2024), dass die meisten Dashboards inzwischen auch Daten zu überfachlichen Maßen sowie Absentismus, Schulklima, Partizipation von Eltern sowie administrative Daten zu Finanzen, Personal oder Stundenplan umfassen. Diese „comprehensive systems“ (Curran et al., 2024, S. 26) ermöglichen zwar eine Übersicht über viele relevante Indikatoren. Allerdings wird wie auch in älteren Studien zu Managementinformationssystemen auf die Gefahr der Überforderung durch Überflutung mit Daten hingewiesen.

- **Qualifizierung schulischer Akteur:innen für eine effektive Nutzung von Datenfeedbacksystemen**

Die Qualifizierung für die systematische Nutzung von Datenfeedbacksystemen sollte auf das Leitbild der *Professional Statistical Literacy* (Pierce & Chick, 2011) ausgerichtet sein, die an den drei Anforderungen *Reading Values*, *Comparing Values* und *Analyzing the Data Sets* orientiert und auf die Anforderungen der domänenspezifischen, hier schulischen, Datennutzung bezogen ist. Gal (2004) hält zwar mathematische Grundkompetenzen für erforderlich, vor allem die Kenntnis statistischer Grundkonzepte wie Mittelwerte, Prozentwerte und Standardabweichungen. Wichtig ist aber vor allem, dass Lehrkräfte Daten kontextualisieren sowie Schlussfolgerungen aus Daten ableiten können. Das heißt, sie müssen auch mit Grundkonzepten wie Standards, Kompetenzstufen und Bezugsnormen vertraut sein.

Entscheidend für die Qualität eines Datenfeedbacksystems, ob auf Papier oder in Form eines Dashboards, sind die zugrunde liegenden konzeptuellen Überlegungen zur Schulqualität, die die Auswahl der Daten bzw. Konstrukte begründen. In den Niederlanden wurde bereits vor 25 Jahren ein Datenfeedbacksystem für schulische Selbstevaluation mit dem Label *ZEBO* in Kooperation dreier wissenschaftlicher Einrichtungen entwickelt (Hendriks et al., 2001). Neben standardbasierten Tests umfasst das Instrument auch Prozessindikatoren auf der Schulebene, die aus Befunden der Schuleffektivitäts- und Schulentwicklungsforschung abgeleitet wurden. Die Operationalisierung schulischer Prozessindikatoren erfolgte sparsam, bezogen auf ihren Beitrag zur Erklärung von Lernzuwachs. Das System beruht neben zentral erhobenen Testdaten auf validen Instrumenten, mit denen unterschiedliche Perspektiven (Schüler:innen, Lehrkräfte, Eltern) erfasst werden. Es wurde nicht nur hinsichtlich der Validität der Daten, sondern auch bezüglich technischer und grafischer Aspekte sehr sorgfältig entwickelt (Visscher & Coe, 2003) und in mehreren Pilotierungsschleifen getestet. Zusätzlich zum Datenfeedback wurde eine Telefonhotline zur Unterstützung der Dateninterpretation eingerichtet.

Um Schulen bei einer reliablen, validen und gleichermaßen effizienten Evaluation zu unterstützen, orientierte sich die Entwicklung an folgenden Kriterien:

- Den Grundpfeiler bilden Leistungsdaten von Schüler:innen zur Abbildung von Lernfortschritten.
- Zum Zweck der Adjustierung dieser Daten werden Informationen zu Hintergrundmerkmalen der Schüler:innen wie sozioökonomischer Status oder Lernausgangslage erfasst.
- Prozessindikatoren sollten fokussiert sein auf schulische Faktoren, die Lernzuwächse erklären.



- Die Nutzung des Tools sollte keine zusätzliche Belastung der Schulen mit sich bringen. Wenn möglich sollten Daten genutzt werden, die Schulen nicht selbst erfassen müssen.
- Schulische Akteur:innen sollten ohne statistische Kenntnisse begründete datengestützte Entscheidungen treffen können.
- Das System sollte Schulen einen sinnvollen Vergleich mit anderen Schulen ermöglichen.

Das Tool wurde in niederländischen Grundschulen implementiert und evaluiert. Die Studien zeigen deutliche Leistungszuwächse der Schüler:innen im Vergleich zu Kontrollgruppen (Hendriks et al., 2001; Schildkamp & Visscher, 2009; Schildkamp et al., 2009; Schildkamp et al., 2012). Visscher und Coe (2003) berichten außerdem eine Steigerung der Qualität von Schulentwicklungsplänen infolge der Implementation von ZEBO. Die sorgfältige und systematische Konzeption, Implementation und Evaluation von ZEBO setzten Maßstäbe für die Entwicklung von Datendashboards für Schulen.

2.2.4 Unterstützungsstrukturen für datengestützte Schulentwicklung

Sowohl die Befunde zur Implementation von Schulentwicklungsplänen als auch die Befunde zur Nutzung von Datenfeedbacksystemen zeigen, dass es notwendig ist, begleitend zur Implementation der strategischen Steuerungsinstrumente Unterstützungsstrukturen zu implementieren. Folgende Unterstützungsangebote können unterschieden werden: Trainings für Data Literacy, Hotlines zur Unterstützung der Dateninterpretation, Coachings, Handreichungen, Materialien sowie Aufgabensammlungen zur kollegialen Schulentwicklung.

Data-Literacy-Trainings für Lehrkräfte haben laut einer aktuellen Metaanalyse auf der Basis von 33 Studien mittlere Effekte auf Datennutzungskompetenz und -motivation. Besonders erfolgreich sind kollaborative Formate (Filderman et al., 2022). Andere niedrigschwellige Formate sind Hotlines zur Unterstützung der Dateninterpretation, wie sie z. B. im Rahmen des oben dargestellten ZEBO-Systems in den Niederlanden implementiert wurden. Befunde aus einem aktuellen Forschungs- und Entwicklungsvorhaben zur KI-basierter Unterstützung datengestützter Entscheidungen in Schulen weisen darauf hin, dass Live-Agent Hotlines mittelfristig durch KI-basierte Chatbots ersetzt werden können (Jiwon Lee & Lee, 2025).

Unterstützung darf sich aber nicht nur auf die Interpretation der Daten im Sinne der Förderung von Data Literacy beziehen, sondern muss auch die Ableitung und Implementation von Maßnahmen umfassen. Grabarek und Kallemeyn (2020) ziehen aus einem Review von 39 Studien zur datengestützten Schulentwicklung den Schluss, dass die Interventionsstudien, die positive Effekte auf Schülerleistungen identifiziert haben, eine kontinuierliche professionelle Entwicklung unterstützten. Besonders effektiv sind unterrichtsbezogene Coachings, die auf der Ebene der Fachkonferenzen sowie einzelner Lehrkräfte ansetzen. Die fachliche bzw. fachdidaktische Expertise dieser Coachenden ist eine entscheidende Voraussetzung gelingender Entwicklung (vgl. Kapitel 3.2.2).

Eine zentrale Bedeutung für eine wirksame datengestützte Schulentwicklung hat die Bereitstellung evidenzbasierter, fachdidaktisch begründeter Förderkonzepte, Materialien sowie Aufgaben (Prediger & Karst, i. E.). Hier sind einerseits Handreichungen, Aufgabensammlungen (Tarkian, Lankes & Thiel, 2019) sowie Blended-Learning-Tools zur Unterstützung der Unterrichtsentwicklung mit Vergleichsarbeiten



(Hawlitschek et al., 2025) zu nennen, andererseits fachdidaktische Materialien, wie sie z. B. für Sprachförderung in BiSS (Becker-Mrotzek et al., 2022) und für Mathematik im QuaMath-Verbund (Prediger et al., 2024) entwickelt wurden. Die Notwendigkeit, Schulentwicklung durch evidenzbasierte Ansätze und Materialien zu unterstützen, gilt gleichermaßen für überfachliche Bildungsziele. Positive Effekte schulbasierter Interventionen zur Förderung sozial-emotionalen Lernens sind durch mehrere Reviews und Metaanalysen belegt (Cipriano et al., 2023; Payton et al., 2008; SWK, 2022a).

Professionalisierung sowie die Implementation von Förderkonzepten und Materialien können über die Einzelschule hinaus in Schulnetzwerken erfolgen (Herman et al., 2024; Klopsch & Sliwka, 2020). Diese Netzwerke ermöglichen nicht nur die Abstimmung von Zielen und Bildungsangeboten in einer Bildungsregion, sondern auch eine effiziente und nachhaltige Implementation von Qualifizierungs- und Coachingangeboten, z. B. für Schulleitungen, Daten-Teams oder Fach-Teams.

2.3 Schlussfolgerungen und Empfehlungen

Um die gesamte Schule strategisch zu entwickeln und Unterrichtsentwicklung sowie Lernförderung zwischen den einzelnen Lehrkräften einer Schule abzustimmen, benötigen Schulen zuverlässige Daten, die eine Bestandsaufnahme ermöglichen und Entwicklungen sichtbar machen. Dazu liegen mit den ländergemeinschaftlich entwickelten zentralen Lernstandserhebungen Daten vor (vgl. Kapitel 1). Wurde Bildungsqualität lange vor allem an Leistungsdaten gemessen, erfolgte international in den letzten Jahren eine Ergänzung um überfachliche Bildungsziele, die in schulischen Qualitätsrahmen vieler OECD-Länder verankert wurden. Dies ist in Deutschland erst in Ansätzen der Fall.

Schulen benötigen für eine strategische Schulentwicklung auch Daten zu Prozessen der Unterrichts- und Schulqualität. Für Unterrichtsevaluation existieren in Deutschland bislang nur wenige validierte Instrumente (vgl. Kapitel 1). Was Indikatoren der Schulqualität betrifft, zeigt sich in den Qualitätsrahmen der Länder eine hohe Übereinstimmung. Allerdings liegen auch hier bislang nur wenige validierte Instrumente vor.

Die wichtigsten strategischen Instrumente einer datengestützten Schulentwicklung sind Schulentwicklungspläne und Selbstevaluationstools. Im Unterschied zu fokussierten Schulentwicklungsplänen, die z. B. in Österreich oder Kanada mit digitalen Templates auf wenigen Seiten erstellt werden, sind Schulprogramme in Deutschland in der Regel häufig nicht nur sehr textlastig und umfangreich, sondern oft auch nicht auf klar formulierte und priorisierte Entwicklungsziele ausgerichtet.

Während sich in Deutschland Datenfeedbacks vor allem auf Vergleichsarbeiten beziehen und Schulen andere relevante Konstrukte in der Regel selbst erfassen müssen, wurden in anderen Staaten integrierte Datenfeedbacksysteme für Selbstevaluation implementiert, die Leistungsdaten, Kontextdaten und Prozessdaten zu Unterrichts- und Schulqualität sowie Daten zu überfachlichen Konstrukten zur Verfügung stellen. Die in Dashboards eingespeisten Daten liegen entweder als administrative Daten bereits vor oder werden zentral erfasst. So werden Schulen von der Datenerhebung entlastet. Entscheidend für eine wirksame Nutzung ist, dass den Dashboards ein auf zentrale Qualitätsmerkmale fokussiertes Modell zugrunde liegt.



Datenfeedbacks allein genügen nicht, um wirksame Schulentwicklung anzuregen. Es ist darüber hinaus notwendig, Entwicklung systematisch mit evidenzbasierten Materialien sowie Professionalisierungs- bzw. Fortbildungsangeboten zu unterstützen, die in der gesamten Schule sowie über die Einzelschule hinausgehend in Schulnetzwerken einer Bildungsregion implementiert werden.

Unverzichtbare Voraussetzung einer effektiven datengestützten Schulentwicklung ist ein professionelles Leitbild, in dem die datengestützte Diagnose bzw. Bestandsaufnahme sowie Überprüfung der eigenen Arbeit verankert sind. Dieses Leitbild muss Ausdruck finden in der Kultur jeder Schule.

Vor diesem Hintergrund empfiehlt die SWK:

Empfehlung 7) Strategische Schulentwicklungsplanung und Datennutzung in einem Qualitätsentwicklungszyklus systematisch verknüpfen; dabei

1. Daten aus zentralen Lernstandserhebungen klassen- und jahrgangsbezogen rückmelden, dabei auch Mittelwerte und Streuungen angeben sowie kriteriale, soziale und adjustierte²⁶ Vergleiche ausweisen (vgl. Kapitel 1); außerdem Daten zu zentralen Prüfungen, Abschlüssen, Übergängen/Wiederholungen, Unterrichtsausfall, Absentismus, Wohlbefinden, schulischer Zufriedenheit, sozialer Eingebundenheit rückmelden; im optimalen Fall erfolgt die Datenrückmeldung mittels Dashboards; Selbstevaluationsdaten sollten ins Dashboard integriert werden können,
2. (digitale) Templates für Schulentwicklungspläne anstelle von Schulprogrammen entwickeln und implementieren, die zentrale datengestützte fachbezogene und überfachliche Ziele, evidenzbasierte Maßnahmen, notwendige Ressourcen, Verantwortlichkeiten sowie Indikatoren zur Evaluation der Zielerreichung festhalten und jährlich fortschreiben,
3. Vorgaben für Selbstevaluation vereinfachen und Berichtspflichten reduzieren zugunsten einer kontinuierlichen datengestützten Entwicklung, die Schulentwicklungsziele und Maßnahmen regelmäßig fortschreibt und im Schulentwicklungsplan dokumentiert,
4. Selbstevaluation und anlassbezogene externe Evaluation (vgl. Kapitel 3) verknüpfen auf der Grundlage eines fokussierten, evidenzbasierten Qualitätsrahmens, der zentrale Outcomes abbildet, auf Kernindikatoren wirksamer Schule beschränkt ist und Lern- und Entwicklungsprozesse von Schüler:innen ins Zentrum stellt.

Empfehlung 8) Strukturelle und personelle Voraussetzungen schaffen für eine datengestützte Schulentwicklung; dabei

1. Verantwortlichkeiten für datengestützte Schulentwicklung in Schulleitungsteams, bei Fachkonferenzleitungen und entsprechenden Funktionsstellen definieren sowie aufgabenbezogene Fort- und Weiterbildungen anbieten,

²⁶ Sofern vorhanden, sollten für die adjustierte Datenrückmeldung die Lernausgangslagen der Schüler:innen bzw. die Sozialindizes der Länder genutzt werden.



2. Zeitfenster ausweisen für kollegiale datengestützte Entwicklung in Fachgruppen, Jahrgangsteams sowie im Gesamtkollegium, einschließlich der multiprofessionellen Teams im Ganztage, um eine verbindliche Abstimmung über Unterrichtsziele, Aufgaben, Lernmaterial, Klassenarbeiten sowie die Kalibrierung von Korrekturen zu erzielen,
3. Unterstützungssysteme für datengestützte Schulentwicklung mit einem Schwerpunkt auf fachbezogene²⁷ Unterrichtsentwicklung bereitstellen sowie Schulnetzwerke zur schulübergreifenden datengestützten Entwicklung implementieren, insbesondere bei kleineren Schulen in ländlichen Regionen (vgl. Kapitel 1, 3),
4. zentrale Erhebungen auf Länderebene zu Kernindikatoren der Schul- und Unterrichtsqualität inklusive Wohlbefinden, schulischer Zufriedenheit sowie sozialer Eingebundenheit (vgl. Kapitel 1) implementieren, die den Schulen über das Datendashboard zur Verfügung gestellt werden können; Schaffung der datenschutzrechtlichen Voraussetzung für die Nutzung administrativer Daten.

²⁷ Bzw. für die berufliche Bildung: lernfeldbezogene Unterrichtsentwicklung.



3. Daten für Controlling und Unterstützung durch die Schulaufsicht

Die Funktion der Schulaufsicht ergibt sich aus Art. 7 GG, demzufolge „das gesamte Schulwesen unter der Aufsicht des Staates“ steht (Heckel & Avenarius, 2000, S. 232). Die Aufgaben der Schulaufsicht haben sich allerdings im Zuge der Übertragung dienstrechtlicher Aufgaben auf die Schulleitung in den letzten zwei Jahrzehnten deutlich verändert. Neben der Rechts- und Fachaufsicht, die nach wie vor wichtige Aufgaben der Schulaufsicht sind, werden inzwischen in allen Ländern auch Qualitätssicherung und -entwicklung als zentrale Aufgaben der Schulaufsicht betrachtet. Konkret reicht dies von der Unterstützung der Schulprogrammarbeit und internen Evaluation über die Veranlassung zentraler Lernstandserhebungen bis hin zur Koordination von Evaluationen (Kroupa et al., 2019; KMK, 2020a). Lange (2003a) beschreibt diese Veränderung im Aufgabenprofil als Abkehr von einem „tradierte[n] Modell bürokratisch-hierarchisch organisierter Schulaufsicht“ (S. 139). Dies geht mit einer Betrachtung der einzelnen Schule als „pädagogische Handlungseinheit“ (Fend, 1986) einher (vgl. Kapitel 2). Das neue, ergebnisorientierte Steuerungsmodell zielt darauf, das „Selbstorganisationspotenzial der operativen Ebene durch Dezentralisation, Delegation und die Einbeziehung der Leistungsadressaten in Entscheidungsprozesse [zu] nutzen“ sowie „die Handlungsfreiheit der beteiligten Organisationseinheiten [zu] erweitern“ (Lange, 2003a, S. 143). Dies gilt einerseits für das Verhältnis von oberster zu operativer Schulaufsicht und andererseits für das Verhältnis von operativer Schulaufsicht zur Einzelschule. Rechenschaftslegung bzw. Ergebnisklärung sind die Voraussetzungen dafür, dass die „operativen Einheiten im Rahmen ihrer eigenen Professionalität handeln“ können (ebd., S. 143). Evaluation ist deshalb von entscheidender Bedeutung. Sie erfolgt nicht in Form der „Kontrolle mit dem Ziel der Sanktion“, sondern als „Controlling“ (ebd., S. 143) mit dem Ziel der Identifikation von Handlungsbedarf und der Unterstützung von Entwicklung. Die Beratungsfunktion, die die Schulaufsicht bereits vor Einführung des neuen Steuerungsmodells wahrgenommen hat, wird deshalb aufgewertet (Avenarius et al., 2003).

Schulaufsicht muss ihr Handeln am Grundsatz der erweiterten Selbstständigkeit von Schulen ausrichten und darf schulische Gestaltungsspielräume nicht dadurch unterlaufen, dass sie als „geheime Schulleitung“ agiert und schulinterne Prozesse kleinschrittig kontrolliert (Lange, 2003b, S. 490–491). Avenarius (2007) leitet aus dem Grundsatz der schulischen Eigenständigkeit eine *Mäßigungsformel* ab. Sie findet Ausdruck z. B. in § 127 Abs. 2 HSchG, wonach die „Befugnis der Schule, Unterricht und Erziehung selbstständig zu planen und durchzuführen, durch Rechts- und Verwaltungsvorschriften und Anordnungen der Schulaufsicht nicht unnötig eingeengt werden“ darf. Professionelle Gestaltungsspielräume finden allerdings ihre Grenze dort, wo der staatliche Anspruch auf ein qualitativ hochwertiges und gleichwertiges Bildungsangebot nicht gewährleistet wird. Deshalb gilt: „Je mehr es der einzelnen Schule gelingt, ihre Qualität durch interne Evaluation selbst zu verbessern, umso weniger bedarf es der Gegensteuerung durch die Schulaufsichtsbehörde. Umgekehrt gilt: Schulen, die den Qualitätsstandards nicht gerecht werden, laufen Gefahr, ihre Selbstständigkeit zu verlieren“ (Avenarius et al., 2003, S. 163). Schulaufsicht ist dann aufgefordert zu intervenieren, und es greifen idealerweise verbindliche Angebote der Schul- und Fachberatung.

Mit der Neudefinition von Entscheidungskompetenzen und -spielräumen von Schule und Schulaufsicht war seit Ende der 1990er-Jahre auch die Einführung einer externen Evaluation verbunden, die Schulen eine



Rückmeldung zur Qualität ihrer Arbeit geben soll. Die externe Evaluation basiert auf Qualitätsrahmen (Elsing & van Ackeren, 2017; Thiel & Tarkian, 2019) und erfasst vor allem Prozessmerkmale auf der Schul- und Unterrichtsebene. Die institutionelle Verankerung erfolgte entweder in den Landes- oder Qualitätsinstituten oder als Organisationseinheit der Schulaufsicht (Avenarius & Füssel, 2008, S. 81). Kritiker sehen in der externen Evaluation eine Verletzung der „Mäßigungsformel“, wenn Prozessqualität detailgenau operationalisiert wird (Avenarius, 2007, S. 874): „Erhebliche Gefahren drohen der schulischen Eigenverantwortung durch die Art und Weise, wie in einigen Ländern externe Evaluation betrieben wird. Ihr wird zumeist ein Referenz-, Orientierungs- bzw. Handlungsrahmen oder ein Qualitätstableau mit einem Set von Indikatoren zugrunde gelegt, die bestimmen, was eine gute Schule ausmachen soll. Diese Indikatoren beruhen teilweise auf einem Evaluationskonzept, das [...] sämtliche Schulen einem einheitlichen Qualitätsmuster unterwirft. Ein solches Vorgehen verstößt gegen die den Schulen gesetzlich eingeräumte Selbständigkeit und Eigenverantwortung“ (Avenarius & Füssel, 2008, S. 82).

Grundsätzlich gilt, dass sich Rolle und Funktion der Schulaufsicht in den letzten zwei Jahrzehnten stark verändert haben. Herausforderungen der Rollendiffusion und Schnittstellenprobleme im Verhältnis zu anderen Akteur:innen wurden von Klein und Bremm (2020) in juristischen und empirischen Analysen dargestellt.

In der nachfolgenden Situationsanalyse wird für die Länder dargestellt, welche Daten für Schulaufsichten zur Verfügung gestellt, wie diese Daten aufbereitet und für Dialoge sowie Zielvereinbarungen mit Schulen genutzt und wie Schulaufsichten für die Nutzung von Daten qualifiziert werden (3.1.1). Anschließend wird das Verhältnis von externer Evaluation und Schulaufsicht beleuchtet und dargestellt, ob und wie externe Evaluation in den einzelnen Ländern implementiert ist (3.1.2). Abschließend wird ein Blick auf die Unterstützungsprozesse im Anschluss an ein Datenfeedback durch die Schulaufsicht oder die externe Evaluation geworfen (3.1.3). Hier ist zwischen Schul- und Fachberatung zu unterscheiden.

3.1 Situationsanalyse

3.1.1 Erfassung, Aufbereitung und Nutzung von Daten für schulaufsichtliches Handeln

Die Arbeit der Schulaufsichten nutzt unterschiedliche Daten etwa zu Schülerzahlen, der Zusammensetzung der Schülerschaft, dem Unterrichtsausfall oder den Ergebnissen. Die Länder unterscheiden sich darin, welche Daten genau in welcher Form den Schulaufsichten rückgemeldet werden. VERA-Ergebnisse z. B. werden gemäß der SWK-Länderabfrage Schule in elf Ländern auch der Schulaufsicht zur Verfügung gestellt.

Die Aggregationsebene der rückgemeldeten Daten unterscheidet sich dabei (vgl. Abbildung 6). In acht Fällen werden Daten der Schulaufsicht auf Jahrgangs- und in neun Fällen auf Landesebene zurückgemeldet sowie in sechs Fällen zusätzlich adjustierte Vergleiche. Für fünf Länder ist zudem eine Rückmeldung auf Klassenebene angegeben. Weiterhin zeigt sich eine Varianz im genutzten Format der Datenrückmeldung. In vielen Fällen erfolgt die Datenaufbereitung in Berichten und/oder Datenblättern. In vier Ländern wird zusätzlich ein Gespräch zwischen den Schulen und der Schulaufsicht geführt.

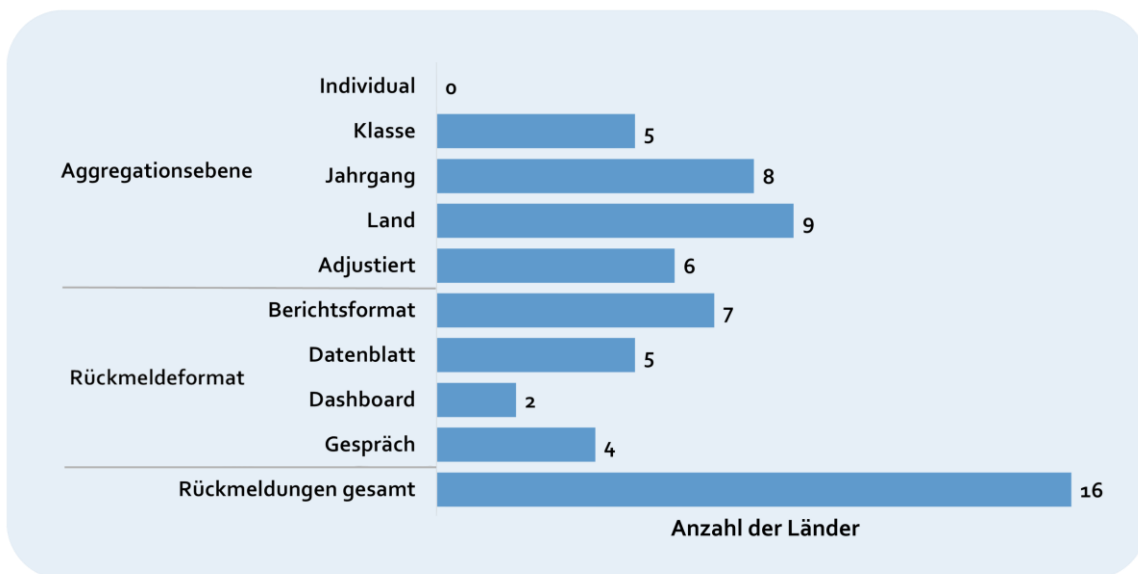


Abbildung 6: VERA-Rückmeldungen an die Schulaufsicht
Quelle: SWK-Länderabfrage Schule

Neben den Schulleistungsdaten zeigt sich die Varianz in der Erfassung und Rückmeldung von Daten an die Schulaufsicht auch bei weiteren Indikatoren. So werden in acht Ländern die Abwesenheiten administrativ/statistisch erfasst, aber nur in vier Ländern an die Schulaufsichten rückgemeldet. Das Gewalt- und Disziplinarverhalten wird in sechs Ländern teils administrativ/statistisch und teils durch Befragungen/Beobachtungen erhoben und auch den Schulaufsichten zur Verfügung gestellt. Daten zum Schulklima werden in drei Ländern erfasst und in einem der Schulaufsicht rückgemeldet. Motivation wird in zwei Ländern erfasst, ohne dass eine Aufbereitung für die Schulaufsicht vorgesehen ist. Einzelne Länder geben zudem Rückmeldungen zu den Faktoren Wohlbefinden, Selbstregulation und wahrgenommene Unterrichtspraxis an.

Aufbereitung der Daten in Datenblättern und Dashboards

In einigen Ländern erhält die Schulaufsicht die aufbereiteten Daten in Form von Schuldatenblättern. Baden-Württemberg führte diese beispielsweise im Schuljahr 2023/2024 ein und informiert mit unterschiedlichen Datenquellen u. a. die Schulaufsichten. Schulleistungsdaten sowie weitere Daten bezüglich der Rahmenbedingungen werden in einfachen Kennwerten wie Prozentanteilen und Spannweiten sowie mit Vergleichswerten zur Verfügung gestellt.²⁸ Ähnlich sieht es in Hamburg aus – neben absoluten Häufigkeiten und Prozentanteilen wird beispielsweise für die Ergebnisse der Schulinspektion eine vierstufige Skala verwendet (Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ), 2024a) (vgl. Kapitel 3.1.2).

Nach der SWK-Länderabfrage Schule arbeiten bisher erst wenige Länder mit Datenaufbereitungen in Dashboards, die auch die Schulaufsichten adressieren. Deutlich mehr Länder befinden sich in der

²⁸ <https://ibbw-bw.de/Lde/Startseite/Systemanalysen/Schuldatenblatt> (aufgerufen am 20.02.2026).



Entwicklung interaktiver sowie mehrperspektivischer Dashboards und tauschen sich dazu auch im länderübergreifenden Netzwerk Dashboards aus.

Datengestützte Dialoge und Zielvereinbarungen zwischen Schulaufsicht und Schule

Im Idealfall werden auf der Grundlage von Daten *Zielvereinbarungen* zwischen der Schulaufsichtsbehörde sowie den Schulen bzw. Schulleitungen in einem partizipativen Prozess vereinbart. Zielvereinbarungen legen Maßstäbe für die Bewertung der zu erbringenden Leistungen fest (Avenarius & Füssel, 2010). Für alle Länder sind Regelungen über den Abschluss von Zielvereinbarungen zwischen Schulaufsichtsbehörden und den Schulen bekannt; in Niedersachsen werden sie derzeit im Rahmen des Startchancen-Programms eingeführt. Unter ihnen gibt es begriffliche Abweichungen in Baden-Württemberg und Hamburg, hier wird von *Ziel- und Leistungsvereinbarungen* gesprochen, in Bayern von *Ziel- und Handlungsvereinbarungen* sowie in Berlin von dem *Schulvertrag*. In den meisten Ländern sind entsprechende Regelungen zu den Zielvereinbarungen in den Schulgesetzen verankert. Ausnahmen gibt es z. B. in Nordrhein-Westfalen mit Regelungen in einer Qualitätsanalyse-Verordnung²⁹ oder in Brandenburg mit einer Verwaltungsvorschrift³⁰ (vgl. Anhang 5). Die genaue Ausgestaltung von Zielvereinbarungen, einschließlich ihres Inhalts und ihrer Laufzeit, ist gesetzlich und untergesetzlich nicht festgelegt. In einzelnen Ländern gibt es allerdings Handreichungen³¹ zur Erstellung von Zielvereinbarungen (Kroupa et al., 2019). Die Laufzeiten der Zielvereinbarungen variieren in einem Zeitraum von einem bis fünf Jahren (vgl. Anhang 5).

Die Zielvereinbarungen werden in aller Regel auf der Grundlage der externen Evaluationsergebnisse sowie von Schulleistungsdaten geschlossen. Welche weiteren Daten den Vereinbarungen zugrunde liegen, unterscheidet sich zwischen den Ländern. So ist bekannt, dass in Einzelfällen auch Ergebnisse der internen Evaluation sowie schulstatistische Daten betrachtet werden (vgl. Anhang 5).

Qualifizierung der Schulaufsicht für Datennutzung

Während in anderen Ländern nicht nur Standards, sondern auch Karrierewege für Führungspositionen in und außerhalb der Schule auf der Grundlage klarer Anforderungsprofile und entsprechende Weiterbildungsprogramme entwickelt wurden (SWK, 2023), sieht das deutsche Beamtenrecht zwar Funktionsstellen und Laufbahnen für Schulaufsichtsbeamt:innen vor, ein zwischen den Ländern abgestimmtes Aufgabenprofil, das als Grundlage für die Entwicklung von Qualifikationsrahmen dienen könnte, liegt für die Schulaufsicht indes bislang noch nicht vor. Die KMK (2025) hat allerdings kürzlich das *Zielbild zur Rolle und Arbeit der Schulaufsicht* verabschiedet, das perspektivisch auch die Entwicklung eines Qualifikationsrahmens orientieren könnte.

In einigen Ländern existieren Verordnungen oder Qualifikationsangebote, die Voraussetzungen für die Bewerbung um eine Position in der Schulaufsicht definieren. Die Qualifikationsmaßnahmen unterscheiden sich allerdings deutlich. So ist beispielsweise in Sachsen-Anhalt eine vom Ministerium begleitete

²⁹ https://recht.nrw.de/lmi/owa/br_text_anzeigen?v_id=1000000000000000545 (aufgerufen am 20.02.2026).

³⁰ https://bravors.brandenburg.de/verwaltungsvorschriften/vv_sv (aufgerufen am 20.06.2026).

³¹ In Bayern gibt es beispielsweise den *Leitfaden für die Erstellung von Zielvereinbarungen* (Lankes & Huber, 2014).



Einführungszeit zu absolvieren³² und in Nordrhein-Westfalen ein vier Module umfassendes Qualifikationsprogramm, das Themen wie Qualitätssicherung, rechtliche Grundlagen oder Veränderungsmanagement umfasst.³³ Gemäß der SWK-Länderabfrage Schule sind in 13 Ländern Fortbildungen oder spezifische Qualifikationsangebote für die datengestützte Schulaufsicht vorgesehen. Die Teilnahme an Fortbildungsinhalten zur datengestützten Schulaufsicht ist in drei Ländern verpflichtend im Rahmen der Schulräteausbildung bzw. der Ausbildung für den Schulaufsichtsdienst. Darüber hinaus gibt es freiwillige Angebote für Schulaufsichten.

3.1.2 Externe Evaluation und schulaufsichtliches Handeln

Aktueller Stand der externen Evaluation in den Ländern

Bremen führte als erstes Land 1994 im Rahmen einer Pilotphase eine externe Evaluation bzw. Schulinspektion ein. Andere Länder folgten. Während in einigen Ländern wie Berlin oder Hamburg die externe Evaluation in den ersten Jahren des 21. Jahrhunderts in den Regelbetrieb überführt wurde, setzen sie andere Länder bereits wieder aus oder stellten sie auf ein anlassbezogenes Verfahren um. Lediglich im Jahr 2009 existierte in allen Ländern eine regelhaft-verpflichtende externe Evaluation (Tarkian, Lankes & Thiel, 2019). Die institutionelle Verankerung erfolgte entweder in den Landes- oder Qualitätsinstituten oder als Organisationseinheit direkt in der obersten Schulbehörde (Kroupa et al., 2019). Die Abgrenzung der Aufgaben von Schulaufsicht und externer Evaluation ist häufig nicht eindeutig.

Derzeit wird das Verfahren in elf Ländern mit zum Teil unterschiedlichen Bezeichnungen angeboten. Während am häufigsten von der *externen Evaluation* gesprochen wird (u. a. Hessen, Mecklenburg-Vorpommern), heißt es in anderen Fällen *Schulvisitation* (Brandenburg), *Schulinspektion* (Hamburg), *Qualitätsanalyse* (Nordrhein-Westfalen), *Schulfeedback* (Schleswig-Holstein) oder *Qualitätsentwicklung* (Thüringen). Ausgesetzt ist das Verfahren derzeit in fünf Ländern (Bayern, Berlin, Brandenburg, Bremen, Rheinland-Pfalz, Sachsen). In Bayern wurde das Verfahren zum Schuljahr 2025/26 zur Neugestaltung ausgesetzt, weitere Pläne zur Wiederaufnahme sind bisher nicht bekannt. In Brandenburg ist das Verfahren im Schuljahr 2025/26 temporär ausgesetzt, für das Schuljahr 2026/27 ist eine Wiederaufnahme vorgesehen (vgl. Anhang 4).

In einigen Ländern wurde die regelmäßige Inspektion inzwischen auf eine anlassbezogene externe Evaluation umgestellt. Bei der anlassbezogenen Evaluation können Schulen entweder selbst eine externe Evaluation anfordern, oder das Verfahren wird von der Schulaufsicht angeordnet. Schleswig-Holstein hat als erstes Land 2016 eine anlassbezogene Evaluation eingeführt, die eng mit der Schulaufsicht auf der einen Seite und externen Unterstützungssystemen auf der anderen Seite verknüpft ist. Außerdem wurde ein Schulentwicklungsfonds eingerichtet, um gezielte Maßnahmen auf der Grundlage der Evaluation anzuschließen (Schleswig-Holsteinischer Landtag, 2016). Anlassbezogen findet die externe Evaluation aktuell in Baden-Württemberg (derzeit in Erprobung), Mecklenburg-Vorpommern und Sachsen-Anhalt statt. In regelmäßigen Zyklen wird die externe Evaluation in Hamburg, Hessen³⁴, Nordrhein-Westfalen, Saarland und

³² <https://www.landesrecht.sachsen-anhalt.de/bsst/document/VVST-VVST000010792> (aufgerufen am 20.02.2026).

³³ <https://fah.nrw.de/seminarangebote/angebote-auf-abruf/neu-in-der-schulaufsicht> (aufgerufen am 20.02.2026).

³⁴ Die turnusmäßige externe Evaluation findet in Hessen für die selbstständigen allgemeinbildenden und selbstständigen beruflichen Schulen statt ([Externe Evaluation selbstständiger Schulen in Hessen | Lehrkräfteakademie.hessen.de](https://www.hessen.de/Lehrkräfteakademie), aufgerufen am 20.02.2026).



Thüringen durchgeführt. In Schleswig-Holstein ist das Verfahren derzeit für die Schulen im Startchancen-Programm vorgesehen (vgl. Anhang 4).

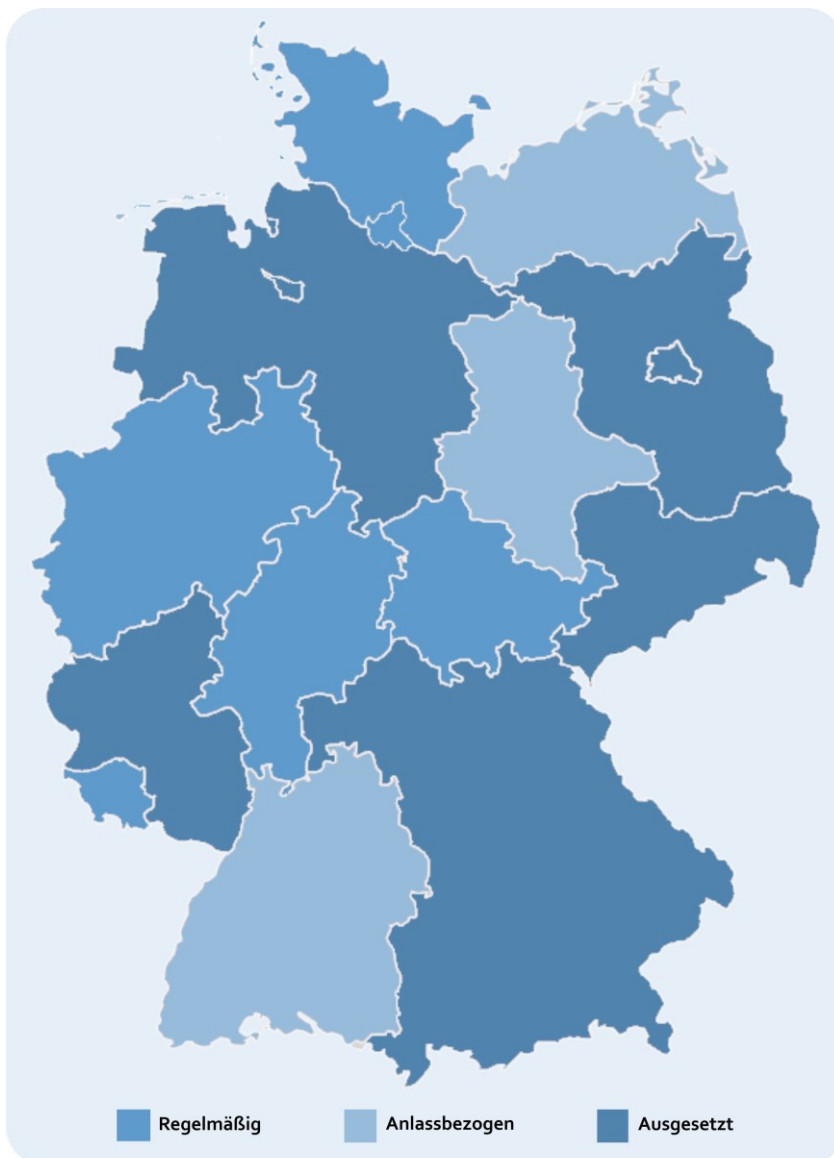


Abbildung 7: Externe Evaluation, Schulinspektion³⁵
Quelle: eigene Recherchen (vgl. Anhang 4)

Die Schulinspektionen sind in den Ländern unterschiedlich ausgestattet. Dazu liegen allerdings nur wenige verlässliche Daten vor. Einem Policy-Papier der Böll-Stiftung zufolge hat etwa Rheinland-Pfalz die Kosten für die 45 Mitarbeitenden der 2015 aufgelösten externen Evaluation auf 3,5 Mio. Euro im Jahr beziffert. In Hessen waren vor der Einstellung im Jahr 2015 45 Lehrkräfte für die externe Evaluation abgeordnet. Ein Teil der Mitarbeiter:innen ist inzwischen in der Schulberatung tätig. In Baden-Württemberg waren vor der Aussetzung des Verfahrens im Jahr 2017 80 Lehrkräfte für die externe Evaluation tätig (H.-J. Kuhn, 2019).

³⁵ Externe Evaluationen an allgemeinbildenden Schulen sind ausgesetzt, an berufsbildenden Schulen finden sie regelmäßig statt.



Externe Evaluation und Schulaufsicht

Die Schulaufsicht wird in unterschiedlicher Form in die Evaluationsverfahren einbezogen. Kroupa et al. (2019) haben auf der Basis von Rechtsdokumenten eine Analyse aller Länder, auch für jene, in denen das Verfahren ausgesetzt wurde, durchgeführt. Für die Steuerung der externen Evaluation ist in allen Ländern die oberste Schulaufsichtsbehörde (Ministerium bzw. Senatsbehörde) zuständig. Die Umsetzung obliegt in den meisten Fällen den Landesinstituten. In einzelnen weiteren Fällen liegt die Verantwortung in selbstständigen Organisationseinheiten innerhalb der obersten oder oberen Schulaufsichtsbehörde

Schnittstellen zwischen externer Evaluation und Schulaufsicht sind häufig nur unzureichend definiert (Tarkian, Lankes & Thiel, 2019). Eine Ausnahme bildet Hamburg. Daten zur Prozessqualität aus der externen Evaluation werden hier auch für Vereinbarungen zwischen Schulaufsicht, Schulinspektion und Landesinstitut genutzt.

Akzeptanz der externen Evaluation in der Praxis und Anschlussprozesse

Es liegen einige Studien zur Implementation der externen Evaluation in unterschiedlichen Bundesländern vor, beauftragt von den Ländern. Diesen Studien zufolge genießt die externe Evaluation eher Akzeptanz bei Schulaufsichten (Gärtner, 2021; Preuße et al., 2019) und Schulleitungen (Böhm-Kasper et al., 2016) als bei Lehrkräften. Ein erheblicher Teil der Schulen bescheinigt der externen Evaluation allerdings mangelnde Relevanz (Besa & Gesang, 2020) sowie wenig Problemlösepotenzial (Pietsch, 2011) und interpretiert sie vor allem als Kontrolle (Gärtner, 2021; Preuße et al., 2019). Damit korrespondiert, dass häufig eine fehlende Unterstützung im Anschluss an die Datenrückmeldung beklagt wird (Drinck et al., 2013; Gärtner, 2021; Müller, 2015; Pietsch, 2011; Preuße et al., 2019). Eine systematische Verknüpfung von externer Evaluation und Unterstützung bzw. Beratung scheint diesen Befunden zufolge in der Praxis häufig nicht zu existieren. Belastbare Studien zu Effekten auf Unterrichtsqualität und Schülerleistungen liegen für die deutschen Inspektionen bislang nicht vor.

3.1.3 Schul- sowie Fachberatungen und schulaufsichtliches Handeln

Schulentwicklungsberatung und datengestützte Entwicklung

Alle Länder haben einer aktuellen Studie zufolge Angebote der Schulentwicklungsberatung eingerichtet (Dedering et al., 2022). Diese sind in zehn Ländern direkt bei den Landesinstituten verankert. In einzelnen Fällen wurden sie sowohl in den Ministerien als auch in den zugehörigen Landesinstituten eingerichtet. In elf Ländern sind Beratungsangebote systematisch mit der Schulaufsicht vernetzt. Bezüglich der Professionalisierung der Berater:innen zeigt sich, dass in elf Ländern das erste und zweite Staatsexamen notwendig sind. Eine ältere Studie kam für Nordrhein-Westfalen allerdings zu dem Ergebnis, dass in der Praxis lediglich 53,4 Prozent der Berater:innen über einen schulnahen Hintergrund verfügten (Dedering et al., 2013, S. 95). Eine besondere Qualifizierungsmaßnahme ist in acht Ländern verpflichtend (Dedering et al., 2022).

Vernetzungen zwischen der Schulentwicklungsberatung und der externen sowie internen Evaluation existieren in elf Ländern, in sieben Ländern ist Schulentwicklungsberatung mit interner Evaluation verknüpft. Die Berater:innen wirken zum Teil bei der externen Evaluation mit und übernehmen danach die Beratung der



Schulen (Dedering et al., 2022). In einzelnen Ländern bietet die Schulentwicklungsberatung extra Angebote, um die Schulen bei ihren internen Evaluationen zu begleiten.³⁶ Dabei ermitteln sie den Zweck der Evaluation, erstellen gemeinsam Ziele und unterstützen bei der Durchführung und Auswertung der Evaluation.

Fachberatungen und datengestützte Entwicklung

Neben Schulentwicklungsberatungen existieren Fachberatungen, die die Schulen bei vorliegenden Problemen durch eine unterrichtszentrierte Beratung unterstützen sollen. Fachberatung lässt sich zwar dem Namen nach von der organisationszentrierten Schulentwicklungsberatung abgrenzen, eine klare Trennung der beiden Beratungsarten ist allerdings nicht immer gegeben (Dedering et al., 2022). Fachberatung wird dem Namen nach ebenso wie Schulentwicklungsberatung in allen Ländern angeboten. In den meisten Fällen richtet sie sich an allgemeinbildende sowie an berufsbildende Schulen. Die Fachberatung ist wie die Schulentwicklungsberatung in Ministerien, Landesinstituten und direkt bei den Schulbehörden angesiedelt. Die Fachberater:innen führen Lehrkräftefortbildungen u. a. zu fachlichen, didaktischen und organisatorischen Fragestellungen durch und können vereinzelt auch individuell fachdidaktisch coachen. Für die meisten Länder ist bekannt, dass die Fachberatungen für einzelne Schulfächer angeboten werden, darüber hinaus gibt es auch Angebote für überfachliche Bedarfe. Jene konzentrieren sich auf unterschiedliche Bereiche, wie beispielsweise die berufliche Orientierung, Begabtenförderung oder Gewaltprävention. Bei einer Bedarfsfeststellung durch die Schulaufsicht werden Fachberatungen angeboten. Darüber hinaus haben Schulen bzw. Schulleitungen die Möglichkeit, die Beratung selbst anzufordern (vgl. Anhang 6). Weiterhin geben die Länder an, dass Fachberatungen im Anschluss an Ziel- und Leistungsvereinbarungen angeboten werden (vgl. Kapitel 3.1.1, Anhang 5).

In Nordrhein-Westfalen wurden im Rahmen des *Masterplans Grundschule* über 100 Stellen für Fachberater:innen geschaffen, um die Unterrichtsentwicklung in den Fächern Deutsch und Mathematik in der Grundschule optimieren zu können (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021). Erfahrungen in großen Programmen zur Schulentwicklung wie *Schule macht stark – SchuMaS* zeigen allerdings, dass trotz Betonung der zentralen Bedeutung fachbezogener Unterrichtsentwicklung nicht genügend Fachberatungen zur Verfügung stehen, um alle Schulen mit fachbezogenen Angeboten zu versorgen. Dies entspricht den Befunden einer neueren österreichischen Studie zur Schulentwicklungsberatung, die einen nicht gedeckten Bedarf an Unterstützung insbesondere im Bereich der Unterrichtsentwicklung feststellt (Hofbauer & Bernhard, 2023).

3.2 Konzepte und Maßnahmen

Schulaufsichten haben einen nicht zu unterschätzenden Einfluss auf die Entwicklung von Schulen. Marzano und Waters (2009) haben auf der Grundlage einer Metaanalyse folgende Ansatzpunkte wirksamen schulaufsichtlichen Handelns identifiziert: Entwicklung von Zielen gemeinsam mit Schulleitungen; Etablierung

³⁶ Zum Beispiel in Hessen: [Qualitätssicherung und Qualitätsentwicklung an Schulen | Lehrkräfteakademie.hessen.de](https://www.lehrkraefteakademie.hessen.de) (aufgerufen am 20.02.2026).



nicht verhandelbarer Ziele bezüglich Schülerleistungen und Unterrichtsqualität; Herstellung von Alignment zwischen Zielen unterschiedlicher Akteur:innen; Monitoring der Erreichung nicht verhandelbarer Ziele sowie gezielte Allokation von Ressourcen zur Unterstützung der Zielerreichung. Diese empirische Identifikation zentraler Ansatzpunkte wirksamer Schulaufsicht entspricht den in der Einleitung dieses Kapitels herausgearbeiteten Funktionen des Controllings und der Unterstützung. Daten sind für die Erfüllung dieser Funktionen unverzichtbar.

Die nachfolgende Darstellung der internationalen Befundlage zu wirksamen Konzepten und Maßnahmen der datengestützten Schulaufsicht orientiert sich an diesen Funktionen. Zunächst werden Befunde zur daten- bzw. testbasierten Rechenschaftslegung mit Fokus auf Low-Stakes-Systeme dargestellt. Dies umfasst auch die Fragen, welche Daten für Zwecke der Rechenschaftslegung genutzt werden und nach welchen konzeptionellen (Bezugsnomen) und technischen Kriterien (Dashboards) sie aufbereitet werden (3.2.1). Anschließend werden Fragen des Datenfeedbacks, der Vereinbarung von Zielen sowie der Unterstützung der Datennutzung thematisiert (3.2.2). Abschließend wird auf die Qualifikation von Schulaufsichten für datengestütztes Controlling und Unterstützung eingegangen (3.2.3).

3.2.1 Konzepte und Effekte datengestützter Rechenschaftslegung

Rechenschaftslegung in High- und Low-Stakes-Systemen

Figlio & Loeb (2011) haben in einem Übersichtsartikel für das *Handbook of the Economics of Education* den Forschungsstand zur Einführung standardbasierter Testsysteme für Rechenschaftslegung systematisch aufgearbeitet und kommen zu dem Schluss, dass der Großteil der Studien Effekte auf Schülerleistungen belegt. Einige Studien weisen darauf hin, dass testbasierte Rechenschaftssysteme dann am wirksamsten sind, wenn sie mit Autonomie- bzw. Gestaltungsspielräumen für Schulen kombiniert werden (Torres, 2021; Wößmann, 2007b; Wößmann et al., 2009). Dies wird dadurch begründet, dass Schulen am besten auf die unterschiedlichen Bedarfe vor Ort eingehen könnten und Daten es ihnen ermöglichen, entsprechende *Site-based Decisions* zu treffen (Wohlstetter et al., 2008).

Neben den positiven Effekten testbasierter Rechenschaftslegung sind allerdings auch negative Effekte belegt. Levatino et al. (2024) fassen in einem aktuellen Review Befunde zu nicht-intendierten Effekten einer testbasierten Rechenschaftslegung zusammen. Dazu zählen Teaching-to-the-Test (Einschränkung des Curriculums, Veränderung des Unterrichts zulasten selbstregulierten Lernens), Fokussierung des Unterrichts auf leistungsstärkere Schüler:innen, erhöhter wahrgenommener Stress und verringerte professionelle Autonomie bei Lehrkräften. Diese Befunde gelten vor allem für High-Stakes-Systeme, die ein Verfehlen von Standards mit teilweise drastischen Sanktionen belegen. Zwar werden auch für Low-Stakes-Systeme nicht-intendierte Nebenwirkungen berichtet, allerdings ist unmittelbar plausibel, dass in diesen Systemen die Anreize für einen oberflächlichen oder gar manipulativen Umgang mit Daten ungleich niedriger sind als in High-Stakes-Systemen. Dass Low-Stakes-Systeme gleichwohl einen beträchtlichen Impact auf den Lernerfolg der Schüler:innen haben können und nicht-intendierte Effekte zumindest teilweise reduzieren, belegt eine bildungsökonomische Studie von Woo et al. (2015) am Fall von Korea. Im Unterschied zu High-Stakes-Systemen werden hier keine Sanktionen oder Zulagen für Lehrkräfte implementiert, sondern Schulentwicklung wird mit zusätzlichen Ressourcen für gezielte Maßnahmen und vor allem mit



professionellen Beratungsangeboten unterstützt. Ähnliche Befunde werden in einer Studie mit Daten aus dem Low-Stakes-System in British Columbia (Kanada) berichtet (Filiarov & Sweetman, 2023). Zudem scheinen Daten hier stärker zur Förderung schwacher Schüler:innen genutzt zu werden, als dies in High-Stakes-Systemen der Fall ist (Filiarov & Sweetman, 2023). Neuere bildungsökonomische und politikwissenschaftliche Studien legen nahe, dass Low-Stakes-Systeme eher langfristiges organisationales und professionelles Lernen auslösen, während High-Stakes-Systeme Anreize für strategische Reaktionen setzen (Andersen & Nielsen, 2020; Jakobsen et al., 2018).

Damit auch in Low-Stakes-Systemen Verbindlichkeit sichergestellt wird, müssen Schulaufsichten die Rolle von *Accountability Agents* übernehmen. Diese sind dafür verantwortlich, klare Erwartungen und Qualitätsstandards zu kommunizieren und für eine stabile institutionelle Umwelt zu sorgen, „which ensures that schools behave trustworthily“ (Ehren & Bachmann, 2020, S. 113; vgl. Marzano & Waters, 2009). Standardbasierte Testdaten spielen hier eine wichtige Rolle.

Multiple Erfolgsindikatoren für Rechenschaftslegung

Mit dem ESSA im Jahr 2015 wurde in den USA eine Nachjustierung bei der Definition von Indikatoren für Schulerfolg vorgenommen. Anstelle einer ausschließlichen Fokussierung auf standardbasierte Tests in den Kernfächern wurden nun *Multiple Measures* eingeführt, die Schulerfolg breiter operationalisieren. Ein Beispiel ist das Kalifornische *School Quality Improvement System*, das neben Leistungen und Leistungsentwicklung in den Kernfächern Abschlussquoten, sozial-emotionales Lernen (Growth-Mindset, Selbstwirksamkeit, Selbstregulation, soziales Bewusstsein), Absentismus, Drop-out-Quoten, Schulkultur und -klima umfasst (Bae, 2018) (vgl. Kapitel 1, 2).

Ein ähnliches multiples Indikatorensystem wird der datengestützten Steuerung und Entwicklung in Alberta (Kanada) zugrunde gelegt (Klopsch & Sliwka, 2020). Neben exzellenten Leistungen werden Chancengerechtigkeit, Inklusion sowie Wellbeing von Schüler:innen und Lehrkräften als Qualitätsindikatoren operationalisiert.

Konzeptuelle und technische Aufbereitung von Daten für Rechenschaftslegung

Die Interpretation und Nutzung der Testdaten sind u. a. davon abhängig, in welchen Vergleichshorizonten (Bezugsnormen) sie präsentiert werden. Schul- und jahrgangsbezogene Mittelwerte sowie Standardabweichungen werden mit Bezug auf Kompetenzstufen und – wo längsschnittliche Daten vorliegen – auch längsschnittlich oder als Tendenzquerschnitt präsentiert. Was soziale Vergleiche betrifft, erfolgen Vergleiche mit Landesmittelwerten oder adjustierte Vergleiche. Für Vergleichsarbeiten geben Maier et al. (2011) erste Hinweise darauf, dass ein adjustierter Vergleich die Akzeptanz von Datenrückmeldungen erhöht.

Nachdem Daten zunächst in Form schriftlicher Berichte rückgemeldet wurden, erfolgte seit einigen Jahren vor allem in den USA und Kanada die flächendeckende Einführung von Datendashboards als technologische Infrastruktur, die ein fokussiertes Monitoring erlauben (Bae, 2018). Während Dashboards für Lehrkräfte und Schüler:innen im Sinne von Classroom Analytics (*Teacher-facing* und *Student-facing Dashboards*; Curran et al., 2024; Knoop-van Campen et al., 2023) bereits breit untersucht wurden, liegen für systembezogene Dashboards (Systems-Level Dashboards), die Daten für Schulleitungen und Schulaufsichten



aufbereiten und auf eine organisationsbezogene Datenanalyse abzielen, noch wenige Studien vor. Gemäß einem systematischen Review (Curran et al., 2024) umfassen die Dashboards neben Leistungsdaten in der Regel auch nicht-kognitive Outcomes sowie Kontextdaten. Allerdings kommen mehrere Studien zu der kritischen Einschätzung, dass zu viele Daten zur Überforderung der Akteur:innen und zu einem Misalignment führen können (Curran et al., 2024). Copeland (2016) empfiehlt aus diesem Grund eine Beschränkung auf maximal zehn bis zwölf zentrale Indikatoren, die zentrale Qualitätsziele des Systems abbilden. Grundsätzlich wird das Potenzial von Dashboards für eine kontinuierliche datengestützte Entwicklung und Steuerung von Schule positiv beurteilt. Allerdings wird auch auf nicht-intendierte Nebenwirkungen wie ein mögliches Labeling von Schulen hingewiesen. Die wenigen vorliegenden Studien zeichnen laut Curran et al. (2024) zudem ein eher skeptisches Bild, was die tatsächliche Nutzung betrifft. Eine systematische Qualifikation für Datennutzung scheint – wie es alle Studien zu *Data-based Decision Making* belegen (Mandinach & Schildkamp, 2021) – auch für den effektiven Einsatz von Dashboards eine grundlegende Voraussetzung zu sein.

Prozessdaten zu Schule und Unterricht aus Schulinspektion bzw. externer Evaluation für Rechenschaftslegung

Prozessdaten zur Unterrichts- und Schulqualität werden insbesondere in europäischen Ländern von Schulinspektionen erfasst, aufbereitet und rückgemeldet. Schulinspektionen sind dort Teil des Rechenschafts-systems, ihre Funktion ist häufig nicht eindeutig von der der Schulaufsicht abgegrenzt (für die deutschen Länder vgl. Kapitel 3.1). Die meisten internationalen Studien zu Schulinspektionen kommen zu dem Schluss, dass die Inspektorate keine oder nur geringe Effekte auf Schulentwicklung (Ehren & Shackleton, 2016a; Gärtner et al., 2014) sowie Schülerleistungen (Choi, 2019; Hofer et al., 2020; Luginbuhl et al., 2009) haben. Teilweise werden negative Effekte von Inspektionen berichtet (Rosenthal, 2004). Bedenken bestehen außerdem hinsichtlich der Validität und Reliabilität der Urteile von Inspektor:innen (Bokhove et al., 2023).

Zahlreiche Länder haben in den letzten Jahren die externe Evaluation bzw. Schulinspektion auf anlassbezogene oder risikobasierte Inspektionen umgestellt (A. C. Timmermans et al., 2015) mit dem Ziel, die Agilität zu erhöhen, Ressourcen gezielter einzusetzen und die Belastung für Schulen zu reduzieren (Standing International Conference of Inspectorates, 2013; Verkroost, 2024). Grundlage für die Entscheidung über eine vertiefte Inspektion ist beispielsweise in den Niederlanden eine jährliche Risikoanalyse, für die u.a. adjustierte Schulleistungsdaten, Selbstevaluationsberichte und Beschwerden von Eltern genutzt werden. Basierend auf Schulleistungsdaten erfolgt zunächst eine Kategorisierung in Schulen ohne, mit erhöhtem oder mit hohem Risiko. Von Schulen mit erhöhtem Risiko werden weitere Daten angefordert. Auf dieser Grundlage und ggf. nach einem weiteren Interview wird über eine Inspektion entschieden. Schulen mit einem hohen Risiko erhalten in jedem Fall eine Inspektion vor Ort (Ehren & Shackleton, 2016a). Bislang liegen wenige belastbare Studien zu Effekten risikobasierter Inspektionen vor. Eine niederländische Studie, die das risikobasierte Inspektionsmodell in Sekundarschulen untersucht, gibt erste Hinweise darauf, dass anlassbezogene Inspektionen (Ehren & Shackleton, 2016b) bei schwachen Schulen zwar mit einer Verbesserung von Sprachkompetenzen einhergehen, allerdings sinken die Zufriedenheit sowie die Anmeldezahlen der Schüler:innen, was mit einem Labelingeffekt erklärt wird.



Wenn Schulinspektion anlassbezogen ausgerichtet wird, kommt der Selbstevaluation von Schulen ein größerer Stellenwert zu. Zwei Aspekte sind hier zentral: erstens, dass Selbstevaluation bzw. externe Evaluation sich auf dieselben evidenzbasierten zentralen Standards für Unterrichts- und Schulqualität bezieht, sowie zweitens, dass Schulen valide Indikatoren sowie Instrumente für die interne Evaluation ihrer Prozesse nutzen und dass sie die personellen und organisationalen Kapazitäten haben, diese Daten richtig zu interpretieren (OECD, 2013; Timperley, 2014; vgl. Kapitel 2).

3.2.2 Datenfeedback, Zielvereinbarungen und Unterstützung datengestützter Entwicklung

Schulaufsichten haben als Accountability Agents nicht nur die Aufgabe, für Rechenschaftslegung der Schulen Sorge zu tragen, sondern auch Schulentwicklung zu unterstützen. Die Art des Datenfeedbacks, die Ableitung von Zielen und Maßnahmen sowie die Unterstützung der Schulen mit zusätzlichen Ressourcen und Beratungsangeboten sind dafür entscheidend.

Gestaltungsmerkmale eines wirksamen Datenfeedbacks

Aus der Schulforschung liegen zahlreiche Befunde dazu vor, dass Daten häufig nicht für Entwicklung genutzt werden (Andersen & Nielsen, 2020). Die Chance einer effektiven Datennutzung steigt nicht nur deutlich, wenn diese Daten relevant sind sowie klar und konzise präsentiert werden (Verger et al., 2024), sondern auch, wenn sie nicht unter der Androhung von Zwang, sondern in einem kollegialen und kollaborativen Setting kommuniziert werden (Yan, 2019). Dies entspricht Befunden aus der Arbeits- und Organisationspsychologie, die darauf hinweisen, dass Objektivität, Akkuratheit, Spezifität, Regelmäßigkeit sowie ein aussagekräftiger Vergleichshorizont Voraussetzungen für eine effektive Feedbackverarbeitung sind (Johnson et al., 2024) und dass die Einbettung von Feedback in eine feedbackfreundliche Organisationskultur eine Bedingung für eine effektive Informationsverarbeitung darstellt (Baker et al., 2013). Drei Merkmale kennzeichnen eine solche Feedbackkultur im Schulsystem: Sie ist auf relevante und zeitkritische Informationen zu Lernen und Entwicklung ausgerichtet, durch Vertrauen und Fehlerfreundlichkeit gekennzeichnet und schließlich im Sinne einer dialogischen Kommunikation verankert (Baker et al., 2013).

Zielvereinbarungen zwischen Schulaufsicht und Schulen als Steuerungsinstrumente

Die Ableitung von Maßnahmen aus Daten macht eine Formulierung und Priorisierung von Zielen erforderlich (Schildkamp, 2019). Ziele sind ein entscheidendes Werkzeug der Steuerung von Organisationen, deren Entwicklung in hohem Maße von der Qualität der Zielformulierung abhängt (Hoffmann & Pfister, 2019). Ziele orientieren nicht nur Aktivitäten und Maßnahmen, indem sie diese auf erwartete Ergebnisse ausrichten, sie sind auch für eine Beurteilung des Erfolgs unverzichtbar (Locke & Latham, 2002). Um eine strategische Steuerungswirkung zu entfalten, müssen Ziele fokussiert und spezifisch sein, als erreichbar wahrgenommen, von allen Mitgliedern einer Organisation akzeptiert und bezüglich ihrer Realisierung regelmäßig überprüft werden (Hoffmann & Pfister, 2019).

Was die strategische Schulentwicklung betrifft, müssen sowohl Ziele der einzelnen Schule als auch übergeordnete Ziele der Schulpolitik Beachtung finden. Dass zwischen schulpolitischen und schulischen Zielen ein Alignment stattfindet, ist maßgeblich für eine nachhaltige Schulentwicklung. Bislang liegen aus dem



Schulbereich nur wenige Studien zur Implementation von Zielvereinbarungen zwischen Schulaufsicht und Schule in Deutschland vor; sie berichten widersprüchliche Befunde zu Akzeptanz und Effekten von Zielvereinbarungen (Keitel, 2024; Kleine, 2012; Ulber, 2011).

Positive Erfahrungsberichte liegen allerdings aus Kanada vor, wo in mehreren Regionen ein systematischer Prozess der Vereinbarung von Zielen zwischen Schulaufsicht und Schulen implementiert wurde (Klopsch & Sliwka, 2020). Auf der Grundlage übergeordneter Ziele für die Bildungsregion verabreden Schulaufsichten und Schulleitungen konkrete strategische Ziele für die Einzelschule. Strategische Entwicklungsziele werden datengestützt für drei Jahre formuliert und in einer Zielvereinbarung fixiert. Im Schulentwicklungsplan werden diese Ziele jährlich präzisiert und durch Maßnahmen, Ressourcen und Indikatoren der Zielerreichung unterlegt. Schulleitungen und Schulaufsichten treffen sich außerdem in Schulnetzwerken, um die gesamte Bildungsregion koordiniert weiterzuentwickeln und sich über Hindernisse und erfolgreiche Maßnahmen auszutauschen (Klopsch & Sliwka, 2020). Die über die Systemebenen abgestimmte Zielsetzung erfolgt mit Bezug auf ein übersichtliches Qualitätsmodell, das auf Kernprozesse beschränkt ist, die wiederum – das gilt auch für alle Managementprozesse – eindeutig auf Lernen und Entwicklung der Schüler:innen ausgerichtet sind.

Auch in Österreich wurde 2012 ein Steuerungsansatz implementiert, der auf einer datengestützten Vereinbarung von Zielen zwischen Schulaufsicht und Schulen beruht (Ammann et al., 2020). Bilanz- und Zielvereinbarungsgespräche werden, neben den Schulentwicklungsplänen (vgl. Kapitel 2), als zentrales Instrument der systemweiten Schulentwicklung betrachtet. Rahmenzielvorgaben des Ministeriums bilden entsprechend die Grundlage für schulische Schwerpunktsetzungen, die in den Schulentwicklungsplänen dokumentiert und in Zielvereinbarungen mit der Aufsicht verbindlich formuliert werden. Das Bilanz- und Zielvereinbarungsgespräch soll im Sinne der *dialogischen Führung* gestaltet werden. Vor dem Hintergrund einer Befragung von Schulleitungen und Schulaufsichten identifizieren Amman et al. (2020) die Etablierung einer nachhaltigen und systematischen Schulentwicklung sowie einer Vertrauenskultur als Stärken des Ansatzes. Schulentwicklungspläne sowie datengestützte Bilanzgespräche und Zielvereinbarungen werden als ineinandergreifende Komponenten eines kohärenten Steuerungsansatzes betrachtet. Der österreichische Ansatz stärkt gleichermaßen schulische Eigenständigkeit und Qualitätssicherung durch die Schulaufsicht.

Beratungs- und Unterstützungssysteme für eine wirksame Entwicklung

Verger et al. (2024) machen deutlich, dass eine datengestützte Rechenschaftslegung häufig nur dann erfolgreich ist, wenn Schulen im Anschluss an Zielvereinbarungen mit notwendigen Ressourcen und themenbezogenen (z. B. fachbezogenen) Unterstützungsangeboten versorgt werden, die sie benötigen, um Veränderungen einzuleiten. Eine ganze Reihe von Studien zur Datenrückmeldung zeigen, dass eine Verknüpfung mit thematisch fokussierten Unterstützungsangeboten geboten ist (de Hoyos et al., 2017; Drinck et al., 2013; Gärtner, 2021; Slavin et al., 2013). Deutliche Effekte zeigen sich, wenn diese Unterstützungsmaßnahmen evidenzbasiert und standardisiert erfolgen (Slavin et al., 2013). Unstrittig ist, dass Kernprozesse des Unterrichts und der Förderung von Schüler:innen nur durch lern- und entwicklungsbezogene, d. h. insbesondere auch fachbezogene, Beratungs- und Professionalisierungsangebote wirksam verändert werden können (Kraft et al., 2018; van der Linden et al., 2021). Im Widerspruch dazu war der Aufbau schulischer Beratungssysteme im deutschsprachigen Raum lange Zeit nahezu ausschließlich an Ansätzen der



Organisationsentwicklung und des Prozessmanagements orientiert (Altrichter et al., 2022). Belastbare Befunde zu Effekten organisationsbezogener Schulberatung liegen nicht vor. Die wenigen Studien stützen sich auf Befragungen der Akteur:innen zu förderlichen Bedingungen und Wirkungen ihrer Arbeit (Dedering et al., 2022). Eine neuere österreichische Studie zur Schulberatung kommt zu dem Ergebnis, dass Schulberater:innen in der Regel ihren Fokus nicht auf Unterricht richten und sich dafür auch nicht qualifiziert sehen. Vor dem Hintergrund der überwältigenden Befundlage, dass die Qualität des Unterrichts für den Lernerfolg ausschlaggebend ist, formulieren die Autor:innen der Studie die Forderung nach einer Stärkung der Unterrichtsberatung im Rahmen datengestützter Beratung (Hofbauer & Bernhard, 2023). Dem entspricht die Schlussfolgerung aus einer älteren Studie in Florida, dass die Analyse von Daten und Prozesse der Unterrichtsentwicklung ineinandergreifen sollten (Marsh et al., 2010). Dazu sind auch fachbezogene Unterstützungsmaßnahmen entscheidend.

Fachübergreifende und fachdidaktische Unterrichtscoachings sind die effektivste Methode zur Weiterentwicklung professioneller Kompetenzen. Sie unterstützen nicht nur die Veränderung von Unterrichtspraxis, sondern haben mittlere bis hohe positive Effekte auf Schülerleistungen (Allen et al., 2011; Kraft et al., 2018; L. West & Staub, 2003). Voraussetzungen eines effektiven Coachings sind nicht nur fachliches bzw. lernpsychologisches Wissen der Coachenden, sondern auch spezifische Coachingkompetenzen wie Feedback, Modeling oder Scaffolding (Kennedy, 2016; Kochmanski & Cobb, 2023). Coachende müssen entsprechend qualifiziert werden, und dem Coaching muss ein fundiertes fachdidaktisches oder lernpsychologisches Modell zugrunde liegen (Solarski, 2021). Der Zugang zum Unterricht der gecoachten Personen ist für ein effektives Coaching unverzichtbar. Besonders effektiv sind Ansätze, die Co-Planning und Co-Teaching im Rahmen eines Einzelcoachings vorsehen (Kochmanski & Cobb, 2023) und die über einen längeren Zeitraum dauern. Einzelcoaching in situ ist allerdings ein sehr zeit- und kostenintensiver Ansatz. Um die Skalierbarkeit zu erhöhen, wurden einerseits Gruppencoachings (Kreis & Staub, 2008) oder Ansätze zur Unterstützung professioneller Lerngemeinschaften in Schulen wie *Quality Teaching Rounds* entwickelt, in denen Lehrkräfte zunächst Input zu evidenzbasierten Unterrichtsstrategien erhalten und dann selbstständig in professionellen Lerngemeinschaften arbeiten (Gore et al., 2023). Andererseits werden Einzelcoachings mit ähnlichem Erfolg wie in situ auch online durchgeführt (Matsumura et al., 2019; Pianta et al., 2008). Ein typisches fachbezogenes Online-Coaching beispielsweise zur Weiterentwicklung fachlicher Klassendiskussionen umfasst einen Online-Workshop bzw. Reader, der in die fachlichen Kernkonzepte einführt, sowie Videobeispiele evidenzbasierter Strategien sowie Anregungen zur Implementation und Reflexion dieser Strategien im eigenen Unterricht. Der eigentliche Coachingzirkel umfasst eine Online-Planungssession, in der ein Unterrichtsentwurf besprochen wird, einen Kommentar des Coaches bzw. der Coachin zu ausgewählten Stellen eines hochgeladenen Unterrichtsvideos sowie eine anschließende dreißigminütige Online-Coaching-Session (Matsumura et al., 2019).

Die Befundlage lässt den Schluss zu, dass Online-Coachings und Gruppencoachings bzw. begleitete professionelle Lerngemeinschaften skalierbare und gleichwohl effektive Unterstützungsmaßnahmen für fachbezogene und überfachliche Ziele darstellen. Jene Maßnahmen können im Anschluss an Zielvereinbarungen mit der Schulaufsicht bzw. an eine fokussierte externe Evaluation implementiert werden.



3.2.3 Qualifizierung der Schulaufsicht für datengestütztes Controlling und Unterstützung

Aufgabenprofil und Kompetenzen für eine wirksame Schulaufsicht

Eine effektive datengestützte Entwicklung von Schulen ist nicht nur auf entsprechende Kompetenzen von Lehrkräften und Schulleitungen angewiesen, sondern auch auf die Expertise von Schulaufsichten (U. Maier et al., 2011; Mandinach et al., 2011). Data Literacy ist eine wichtige Voraussetzung der Datennutzung, die bei unterschiedlichen Akteur:innen im Schulsystem systematisch entwickelt werden muss (Mandinach & Schildkamp, 2021). Was das Training von Schulaufsichten im Umgang mit Daten betrifft, konnte eine Interventionsstudie von Slavin et al. (2013) in den USA sogar Effekte auf Schülerleistungen belegen.

Kompetenzerwartungen an Schulaufsichten gehen über Data Literacy allerdings deutlich hinaus. Auf der Grundlage einer systematischen Analyse von Aufgaben und Erwartungen an die Schulaufsicht beschreiben Björk et al. (2014) für die USA ein Kompetenzprofil, das folgende Aspekte umfasst: Kenntnisse der Unterrichtsforschung/Pädagogik/Lernpsychologie; Kenntnisse über rechtliche, verwaltungstechnische und finanzielle Fragen; Kenntnisse über politische Entscheidungsfindung; Kenntnisse der Datenanalyse/grundlegende forschungsmethodische Kenntnisse; Kommunikations- und Feedbackkompetenzen; Kenntnisse über Human Relations und Change Management. In diesem Kompetenzprofil spiegelt sich die doppelte Aufgabe der Schulaufsicht: Rechenschaftspflicht zu gewährleisten und Entwicklung zu unterstützen (vgl. auch Tripses et al., 2015). Standards für Schulaufsichten, wie sie z. B. in den USA vom National Policy Board for Educational Administration (2015) oder in Kanada vom Government of Alberta (2023) verabschiedet wurden, entsprechen weitgehend dem von Björk et al. (2014) ermittelten Aufgabenprofil. Diese Standards stellen die Grundlage für die Entwicklung und Zertifizierung spezifischer Leadership-Qualifizierungsprogramme dar.

3.3 Schlussfolgerungen und Empfehlungen

Die Rolle der Schulaufsicht hat sich mit der Erweiterung der Gestaltungsspielräume von Schulen substantiell verändert. Schulaufsichtliches Handeln wird nicht mehr als Kontrolle schulischer Kernprozesse begriffen, sondern einerseits als nachgelagerte Intervention, die dann zum Zuge kommt, wenn Schulen nicht in der Lage sind, Qualitätsstandards in eigener Verantwortung zu sichern. Schulaufsicht übernimmt in diesem Sinne Aufgaben des Controllings bzw. der Rechenschaftslegung. Andererseits soll Schulaufsicht alle Schulen in ihrer Schulentwicklung gezielt unterstützen. Im Idealfall stellt sie eine Schnittstelle zu Beratungs- und Unterstützungssystemen dar und kann Ressourcen für schulische Entwicklungsmaßnahmen gezielt zuweisen. Aussagekräftige Daten sind für die Wahrnehmung der Controlling- und Unterstützungsfunktion zentral.

Ein klares Aufgabenprofil der Schulaufsicht in Abgrenzung zu anderen Akteur:innen, insbesondere zu Schulleitungen und Schulinspektionen, ist eine Voraussetzung für eine effektive Erfüllung schulaufsichtlicher Funktionen. In der Praxis ist allerdings eine Rollendiffusion nach beiden Seiten festzustellen. Ein klar konturiertes Aufgabenprofil für die Schulaufsicht liegt in den meisten Ländern Deutschlands bislang nicht vor. Die KMK hat allerdings erst kürzlich ein Leitbild für die Schulaufsicht beschlossen, das als Grundlage für die Entwicklung eines Aufgabenprofils dienen kann.



Bezogen auf die Rolle der Schulinspektion besteht nicht nur aufgrund der ernüchternden Befunde zur Validität und Effektivität ein Anpassungsbedarf, sondern auch hinsichtlich der mangelnden Verschränkung mit überfachlichen und vor allem fachdidaktischen Unterstützungssystemen. Zudem muss eine bessere Abstimmung mit der schulischen Selbstevaluation (vgl. Kapitel 2) erfolgen. Wenn Prozessmerkmale des Unterrichts und der Schule mit validen Instrumenten von den Schulen selbst erfasst oder durch zentrale Erhebungen zur Verfügung gestellt werden, ist die Erfassung derselben Merkmale durch die Schulinspektion verzichtbar. Eine Umstellung auf ein anlassbezogenes Verfahren externer Evaluation, das mit einer stärkeren Fokussierung und Tiefenschärfe Ansatzpunkte für schulische Weiterentwicklung identifiziert, würde der notwendigen Abgrenzung zur Schulaufsicht Rechnung tragen. Wenn Schulen externe Evaluation auch selbst anfordern können, kann einem möglichen Labelingeffekt einer risikobasierten Inspektion entgegen gewirkt werden.

Von zentraler Bedeutung ist ein Alignment der Ebene der Aufsicht mit den Ebenen der Schule und Schulpolitik. Zu diesem Zweck ist die Einigung auf ein evidenzbasiertes fokussiertes Rahmenmodell schulischer Qualität notwendig, das sich auf Kernprozesse der Qualität beschränkt, überfachliches sowie fachliches Lernen und Entwicklung der Schüler:innen in den Mittelpunkt stellt, relevante Rahmenbedingungen schulischer Arbeit einbezieht und multiple Ergebnisse schulischer Bildung ausweist.

Was die Unterstützungssysteme betrifft, muss nicht nur ein Schnittstellenproblem zur Schulaufsicht festgestellt werden. Es stehen in den meisten Ländern auch zu wenig evidenzbasierte fachdidaktische Unterstützungskapazitäten für Unterrichtsentwicklung (z. B. durch Fachberatungen und Fortbildende) zur Verfügung.

Aus den vorausgehend dargestellten Befunden lassen sich Hinweise auf Ansätze zur Weiterentwicklung und Professionalisierung der Schulaufsicht in Deutschland ableiten. Im Einzelnen empfiehlt die SWK:

Empfehlung 9) Daten für die Schulaufsicht aufbereiten und bereitstellen, die die Ergebnisse schulischer Arbeit sowie ihrer Rahmenbedingungen mit Bezug auf Kernindikatoren der Schulqualität abbilden; dabei

1. auf Daten zu Ausgangslagen und Ergebnissen (u. a. zentrale Lernstandserhebungen (wie StarS, BKT, VERA), zentrale Prüfungen, Abschlüsse, Übergänge, Wohlbefinden, schulische Zufriedenheit, soziale Eingebundenheit, Wiederholungen, Schulabsentismus) sowie Rahmenbedingungen (personelle, räumliche und sachliche Ausstattung; besondere Anforderungen, Unterrichtsausfall) fokussieren, die eine Kontextualisierung der Ergebnisdaten ermöglichen,
2. auf Schul- und ggf. Jahrgangsstufenebene aggregierte Daten (Mittelwerte, Streuungsmaße) rückmelden mit Bezug auf kriteriale Normen, auf soziale (Landesmittelwerte) und adjustierte Vergleiche³⁷ sowie auf Entwicklungen der Einzelschule. Im optimalen Fall werden diese Daten auf Dashboards bereitgestellt.

³⁷ Sofern vorhanden, sollten für die adjustierten Vergleiche die Lernauslagen der Schüler:innen bzw. die Sozialindizes der Länder genutzt werden.



Empfehlung 10) Verbindliche Verfahren zur Unterstützung einer datengestützten Schulentwicklung durch die Schulaufsicht implementieren; dabei

1. jährliche datengestützte Dialoge und Zielvereinbarungen mit Schulen vorsehen,
2. Schnittstellen von Aufsicht und überfachlichen und fachbezogenen Unterstützungssystemen (externe Evaluation, Schulberatung, Fachberatung, Schulnetzwerke) optimieren,
3. flächendeckende externe Evaluation zur anlassbezogenen externen Evaluation umbauen, die im optimalen Fall modular aufgebaut ist, gemeinsam von Schulaufsicht und Schule verabredet wird und eng verknüpft ist mit Beratungssystemen zur Schul- und Unterrichtsentwicklung,
4. Schulentwicklungsfonds schaffen zum Zweck der Zuweisung von Ressourcen für gezielte Maßnahmen durch die Schulaufsicht im Anschluss an Zielvereinbarungsgespräche,
5. einen maßgeblichen Teil der Ressourcen der allgemeinen Schulberatung umschichten zugunsten einer Fachberatung (bzw. Einzel- und Teamcoaching), die Fachkonferenzen und Jahrgangsteams bzw. Fachteams in Schulnetzwerken in fachbezogenen Fragen der Unterrichtsentwicklung und Lernförderung berät.

Empfehlung 11) Schulaufsicht für datengestütztes Controlling und Unterstützung qualifizieren; dabei

1. ein ländergemeinsames Aufgabenprofil für Schulaufsichten entwickeln, mit dem die Funktionen des Controllings und der Unterstützung operationalisiert und mittelfristig in Standards übersetzt werden,
2. verbindliche Qualifizierungen für Schulaufsichten vorsehen, die folgende Inhalte umfassen:
 - forschungsbasierte Indikatoren und Modelle für Schul- und Unterrichtsqualität,
 - Grundlagen der Qualitätsentwicklung und der Steuerung von Schulen,
 - Grundlagen der Datenerfassung, -beurteilung und -interpretation, Datenethik und -schutz,
 - Grundprinzipien der Kommunikation, Beratung sowie des Datenfeedbacks,
 - konzeptionelle und technische Grundlagen digitaler Feedbacksysteme/Dashboards,
 - Anforderungen an Zielvereinbarungen/Leistungsvereinbarungen,
 - verwaltungsrechtliche und -ökonomische Grundlagen der Schulentwicklung.



4. Daten für Erziehungs- und Bildungspartnerschaften zwischen Eltern, Schüler:innen, Lehrkräften und pädagogischen Fachkräften

Eltern und Schule tragen gemeinsam Verantwortung für die Entwicklung von Kindern und Jugendlichen. Das Konzept der Erziehungs- und Bildungspartnerschaft steht für eine gleichberechtigte und kooperative Zusammenarbeit von Schule und Elternhaus, die eine geteilte Verantwortung sowie eine gemeinsame Entscheidungsfindung in den Mittelpunkt stellt (Killus & Paseka, 2022; Wild, 2021). Eine Erziehungs- und Bildungspartnerschaft geht über eine reine Information der Eltern bzw. Erziehungs-/Sorgeberechtigten (im Folgenden kurz: Eltern) durch die Schule hinaus und grenzt sich vom Verständnis der Elternarbeit ab. In erweiterten Konzepten wird darüber hinaus auch die Zusammenarbeit mit pädagogischen Fachkräften und der Jugendhilfe sowie weiteren Partnern:innen im Sinne multiprofessioneller Kooperation innerhalb der Schule sowie im Sozialraum betont; ebenso wird eine aktive Rolle der Schüler:innen selbst hervorgehoben (Killus & Paseka, 2022; Stange et al., 2012). Die grundsätzliche Bedeutung von Erziehungs- und Bildungspartnerschaften sowie multiprofessioneller Kooperation wurde bereits im Gutachten der SWK zur Grundschule hervorgehoben (SWK, 2022a).

Auf *individueller Ebene* der Unterstützung einzelner Schüler:innen können Daten im Rahmen von Erziehungs- und Bildungspartnerschaften den Dialog über Stärken, Entwicklungsbedarfe und Maßnahmen zur Förderung ihrer Entwicklung unterstützen und so zum Lernerfolg der Schüler:innen beitragen. Auch für die Förderung sozial-emotionaler Merkmale (vgl. Kapitel 1.1.4) werden elterliche Beteiligung und klare Zielsetzungen hervorgehoben (Jones et al., 2021; Sánchez Puerta et al., 2016). Eine kontinuierliche Beteiligung von Eltern wird insbesondere auch in sozioökonomisch belasteten Kontexten als relevant erachtet (Dearing et al., 2006). Ein regelmäßiges Feedback der Schule steht in Zusammenhang mit elterlicher Zufriedenheit und Beteiligung (Prabhakar et al., 2023). Daten können dazu beitragen, dass Schüler:innen die eigene Lernausgangslage besser erfassen sowie individuelle Stärken und Entwicklungsbedarfe erkennen können und Impulse zur Selbstregulation erhalten (vgl. Panadero et al., 2018; Winne & Hadwin, 2008). Auf *institutioneller Ebene* der einzelnen Schule können Daten aus systematischen Elternbefragungen eine wichtige Quelle für die Schulentwicklung sein (z. B. zu Schulklima, Kommunikation, Passung von Unterstützungsangeboten und zur Wahrnehmung der Erziehungs- und Bildungspartnerschaft) (vgl. Kapitel 2).

Im Kontext datengestützter Schulentwicklung ist die Qualität der Erziehungs- und Bildungspartnerschaft eng mit der Qualität multiprofessioneller Kooperation innerhalb der Schule sowie im Sozialraum verknüpft. Lehrkräfte und pädagogische Fachkräfte der Schulsozialarbeit, Schulpsychologie des Ganztags sowie der Kinder- und Jugendhilfe können z. B. Informationen zu Leistungen, Wohlbefinden und Teilhabe bündeln, gemeinsam – auch mit den Eltern und Schüler:innen – reflektieren und Förderstrategien abstimmen.

Als unverzichtbar werden ethische und rechtliche Leitplanken zu Zweckbindung, Datenminimierung, Transparenz sowie sicheren Kommunikationskanälen hervorgehoben (Education Endowment Foundation



(EEF), 2018; Europäische Union, 2016). Zugleich ist leitend, dass Daten als Grundlage für entwicklungsorientierte Förderung dienen und nicht als Instrument der Etikettierung oder Stigmatisierung von Schüler:innen (Relevant Assessment and Pedagogies for Inclusive Digital Education (RAPIDE), 2023; Scholes, 2016).

4.1 Situationsanalyse

4.1.1 Erziehungs- und Bildungspartnerschaften zwischen Schule, Eltern und außerschulischen Partnern

Im Folgenden wird skizziert, wie der Auftrag zur Zusammenarbeit mit Eltern rechtlich in Deutschland verankert ist und wie sich die Praxis unter Berücksichtigung vorliegender Befunde darstellt.

Rechtlicher Rahmen und bildungspolitischer Auftrag der Zusammenarbeit

Eltern haben das natürliche Recht und die Pflicht zur Erziehung (Art. 6 Abs. 2 GG); der Staat verantwortet die schulische Bildung (Art. 7 Abs. 1 GG). Die Kultusministerkonferenz (KMK, 2018) betont in ihren Empfehlungen zur Zusammenarbeit mit Eltern die Bedeutung einer intensiven Kooperation für erfolgreiche Bildungsprozesse. Auch in inklusiven Kontexten wird diese Zusammenarbeit als konstitutiv für schulische Qualitätsentwicklung beschrieben (Arndt & Werning, 2013; 2016; SWK, 2022a; Yotyodying & Wild, 2016). Die KMK verweist zudem auf die Bedeutung der angemessenen Einbindung der Schüler:innen in die Erziehungs- und Bildungspartnerschaft von Schule und Elternhaus (KMK, 2018, S. 2).

Die vorgesehene Zusammenarbeit mit Eltern variiert zwischen den Ländern. Eine Gesamtauswertung aller Länder sprengt den Rahmen; daher werden exemplarisch einzelne Regelungen dargestellt. Grundlegend wird in den Schulgesetzen festgehalten, dass das Recht der Eltern, über die Erziehung und Bildung ihrer Kinder mitzubestimmen, geachtet werden soll (z. B. § 1 Abs. 3 SchulG BW, § 1 Abs. 2 Bay EUG). In einigen Ländern wird konkret die Forderung nach einer aktiven Einbeziehung der Eltern in die Gestaltung der Schule und des Unterrichts formuliert (§ 6 BremSchulG). Ein Recht auf Hospitation (mit Absprache/Einvernehmen) ist z. B. ausdrücklich in Berlin vorgesehen (§ 47 Abs. 2 SchulG Berlin); Nordrhein-Westfalen erlaubt Unterrichtsbesuche und – mit Gremienzustimmung – aktive Mitarbeit in geeigneten Bereichen (§ 44 Abs. 3 Schulgesetz NRW). Andere Länder (z. B. Schleswig-Holstein) sehen die Schulkonferenz in der Verantwortung, solche Mitarbeit schulweit zu regeln. Regelmäßig ist zur Einbindung der Eltern die Bildung von Elternvertretungen, Klassenpflegschaften oder Elternbeiräten in den Schulgesetzen vorgesehen.

Ebenso müssen Eltern Ergebnisse und Schlussfolgerungen sonderpädagogischer Diagnostik mitgeteilt und erläutert werden, da sie Grundlage für Förderentscheidungen und ggf. die Wahl eines Schulplatzes sind. Beispielsweise gilt in Rheinland-Pfalz, dass das sonderpädagogische Gutachten einschließlich der Fördermöglichkeiten mit den Eltern besprochen und protokolliert wird (§ 27 Abs. 10 InklSchulO Rheinland-Pfalz). In Nordrhein-Westfalen regelt die *Ausbildungsordnung sonderpädagogische Förderung (AO-SF)* das formelle Verfahren von der Antragstellung bis zur Entscheidung der Schulaufsicht über Förderbedarf und Förderschwerpunkt; die Eltern sind Verfahrensbeteiligte und erhalten den Bescheid. Die datenschutzrechtliche Rahmung sichert Informations-, Einsichts- und Aushändigungsrechte und bestimmt zugleich Grenzen.



Berlin konkretisiert in der *Schuldatenverordnung* die Auskunfts- und Einsichtsrechte nach DSGVO für schulische Datenverarbeitungen – maßgeblich auch für Testergebnisse und Gutachten (§ 7 SchuldatenV Berlin). In Nordrhein-Westfalen legt die *Verordnung über die Verarbeitung personenbezogener Daten (VO-DV I)* zulässige Datenarten, Empfänger:innen (einschließlich Erziehungsberechtigter) sowie organisatorische Sicherungen fest; sie flankiert damit die rechtmäßige Weitergabe an Eltern (§ 3 VO-DV I NRW).

Die KMK empfiehlt, Eltern niedrigschwellig, umfassend und adressatengerecht zu informieren, um gemeinsame Absprachen zum Wohl des Kindes zu ermöglichen und sie von Beginn an zur aktiven Mitarbeit im Schulalltag und in den Gremien einzuladen (KMK, 2018). Konzepte zur Elternbeteiligung sollen in Leitdokumenten ausgewiesen sowie im Rahmen schulischer Evaluationsvorhaben weiterentwickelt werden. Angestrebt wird eine angemessene Vertretung aller Eltern in den Gremien, auch z. B. von „Eltern mit Migrationshintergrund“ (KMK, 2018, S. 8), solchen in sozial benachteiligten Lebenslagen oder bei Vorliegen sonderpädagogischer Unterstützungsbedarfe. Auch bei der Bearbeitung konkreter Herausforderungen im Bildungssystem wie z. B. Berufsorientierung im MINT-Bereich (KMK, 2024a), Medienbildung (KMK, 2024b) oder Förderung leistungsschwächerer Schüler:innen (KMK, 2024c) wird die Elternbeteiligung hervorgehoben. Im *Bericht der Länder zum Stand der Umsetzung der Förderstrategie für leistungsschwächere Schülerinnen und Schüler* (KMK, 2024c) wird die Bandbreite der Elternbeteiligung in den Ländern sichtbar. Explizit benannt wird eine intensive Elternbeteiligung im Sinne einer Erziehungs- und Bildungspartnerschaft vorwiegend auf Grundschulebene oder im sonderpädagogischen Förderbereich. Einzelne Länder beschreiben Aspekte einer solchen Zusammenarbeit auch für die Sekundarstufe I (z. B. Sachsen oder Thüringen; KMK, 2024c, S. 60, 66). Über die rechtlich vorgesehenen Formen hinaus können Schulen zusätzliche Beteiligungsangebote entwickeln, um die Mitwirkung breiter und inklusiver zu gestalten. Explizit benannt ist auch die Rolle außerschulischer Partner (z. B. Jugendhilfe, freie Träger, Akteure im Sozialraum), die Erziehungs- und Bildungspartnerschaften unterstützen (KMK, 2018).

Von der Elternarbeit zu (multiprofessioneller) Partnerschaft

Goodall und Montgomery (2014) beschreiben elterliche Beteiligung modellhaft als ein Kontinuum vom *Parental Involvement* in der Schule (vor allem schulzentrierte, von der Schule gesteuerte Aktivitäten und Informationsweitergabe) hin zum *Parental Engagement* mit dem Lernen des Kindes im Fokus (lernbezogene Interaktionen vor allem im häuslichen Kontext). Entscheidend ist dabei eine Verschiebung von einer einseitigen Steuerung durch die Schule hin zu einer ausgehandelten, partnerschaftlichen Verantwortung von Schule und Eltern für das Lernen der Kinder. Die Autorinnen beschreiben, dass *Engagement* mit den größten positiven Effekten auf Motivation, Selbstkonzept und Leistungen der Schüler:innen verknüpft ist; schulische *Involvement*-Formate bleiben zwar wichtig, reichen allein aber nicht aus, um das Potenzial auszuschöpfen. Epstein (2018) entwickelte das Modell der *Overlapping Spheres of Influence* (Familie – Schule – Community) und zeigt, wie enge Kooperation gemeinsame, konsistente Botschaften an Kinder über die Bedeutung von Lernen und Anstrengung erzeugt.

In der deutschsprachigen Debatte hat sich der Begriff der *Erziehungs- und Bildungspartnerschaft* etabliert. Damit wird die traditionelle, eher einseitig von der Institution ausgehende Elternarbeit programmatisch weiterentwickelt zu einer kooperativen Beziehung, in der Eltern als Expert:innen für ihr Kind anerkannt und als aktive Mitverantwortliche für Lern- und Entwicklungsprozesse adressiert werden. Im Zentrum stehen



dialogische Kommunikation, geteilte Verantwortung und Kooperation zwischen Familie und pädagogischen Fach- bzw. Lehrkräften, zunehmend auch im Zusammenspiel mit weiteren Akteuren im Sozialraum (z. B. Jugendhilfe, Ganztage, Beratungsstellen) (Wild, 2021). Eine solche Form der Elternarbeit ist anspruchsvoll und verlangt aufseiten der Lehrkräfte entsprechende Kompetenzen, die systematisch entwickelt werden müssen (Antony-Newman, 2024).

Nach Stange zeichnen sich Erziehungs- und Bildungspartnerschaften auch dadurch aus, dass Eltern, Kindertageseinrichtungen, Schulen, Jugendhilfe und weitere Akteure im Sozialraum systematisch und verbindlich zusammenarbeiten (Stange, 2012).

Herausforderungen

Der Begriff der Erziehungs- und Bildungspartnerschaft markiert ein Ideal, das Asymmetrien und Hierarchien zwischen Familien und Institutionen sowie ungleiche Ressourcenlagen von Eltern häufig nur unzureichend berücksichtigt (vgl. Betz, 2015). Eine implizite Adressierung von Eltern als Kund:innen, Zuliefernde oder Klient:innen kann dazu beitragen, Erwartungen, Kommunikationsstile und Beteiligungsmöglichkeiten zu strukturieren sowie Partnerschaftlichkeit zu unterlaufen (Killus & Paseka, 2021). Gerade in sozial benachteiligten Kontexten und in Hinblick auf Schüler:innen mit sonderpädagogischen Unterstützungsbedarfen besteht die Gefahr, dass fehlende oder brüchige Zusammenarbeit vorrangig als Defizit der Eltern interpretiert und Verantwortung für Bildungserfolg einseitig verschoben wird (Schürer & Lintorf, 2023; Wild, 2021). Empirische Studien zeigen, dass vor allem besser gebildete, deutschsprachige Eltern mit hohem Bildungs- und Systemwissen die etablierten Beteiligungsformate nutzen, während Eltern in sozial benachteiligten Lagen oder mit Migrationsgeschichte deutlich seltener erreicht werden (Betz et al., 2017; Lokhande et al., 2014; Sacher, 2012; Turney & Kao, 2009). Erziehungs- und Bildungspartnerschaften eröffnen zwar Chancen für eine stärker geteilte, lernbezogene Verantwortung und für konsistente Botschaften an Kinder, erfordern aber eine reflexive, inklusionssensible Praxis und darauf bezogene Professionalisierung, die strukturelle Asymmetrien und ungleiche Teilhabebedingungen berücksichtigt.

Zudem verweisen die verfügbaren Befunde darauf, dass Schulen zwar eine Vielzahl von – vor allem formalen – Beteiligungsgelegenheiten anbieten (Informationsabende, Beratungsgespräche, Gremienarbeit), dass die Kooperation aber häufig im Muster traditioneller, schulzentrierter Elternarbeit verbleibt und seltener an das Ideal einer gleichberechtigten Verantwortungsgemeinschaft anschließt. Analysen des SVR-Forschungsbereichs (Sachverständigenrat deutscher Stiftungen für Integration und Migration) zu Grundschulen zeigen, dass stärker lernbezogene Unterstützungsformate – etwa systematische Elternbildung oder Materialien zur Förderung des Lernens zu Hause – deutlich seltener sind und zwischen Schulen erheblich variieren (Lokhande et al., 2014). Dabei gibt es Beispiele von Schulen, die Elternarbeit partnerschaftlich gestalten (Wild, 2021). Insgesamt erscheint Erziehungs- und Bildungspartnerschaft aber noch als ambitioniertes Ziel, dessen Realisierung stark von lokalen Ressourcen, professionellen Haltungen und der Führungskultur einzelner Schulen abhängt (Bronnert, 2024).

Auch die Jugendhilfe oder Gesundheitsdienste arbeiten mit Kindern, Jugendlichen und ihren Familien – jedoch in anderen Rechtskreisen, anderen Organisationslogiken und Dokumentationspflichten. Dies



erschwert den Informationsaustausch sowie eine abgestimmte Hilfeplanung im Einzelfall; Kooperationsvereinbarungen, definierte Schnittstellen und gemeinsame Verfahren können helfen (Dittmann et al., 2018).

4.1.2 Nutzung und Bereitstellung von Daten für die Zusammenarbeit mit Eltern

Daten unterliegen keinem Selbstzweck, sondern können als Gesprächsanlässe dienen, die Orientierung geben und die Verabredung konkreter Schritte unterstützen können. In diesem Sinne können Daten dazu beitragen, in Gesprächen zwischen Lehrkraft und Eltern eine geteilte Sicht auf Lernstand und Lernprozess zu entwickeln und daraus begründete Entscheidungen abzuleiten (Graham-Clay, 2024; Mandinach & Gummer, 2016). Dazu müssen sie so aufbereitet sein, dass alle Beteiligten der Erziehungs- und Bildungspartnerschaft sie verstehen und daraus gezogene Schlussfolgerungen nachvollziehen können (vgl. auch Empfehlung 2).

Rechtlicher Rahmen für Datenerhebungen, -nutzung und -weitergabe

In den schulbezogenen Rechtsnormen sind Aspekte der Datenerhebung und damit verbundene Aspekte in den Ländern unterschiedlich genau geregelt. Zentrale Akteur:innen bei der Datenerhebung und -nutzung sind in der Regel die Schule oder Lehrkräfte. Explizite Regelungen, wie mit personenbezogenen Daten umzugehen ist, insbesondere zur Anonymisierung solcher Daten, sind dagegen selten (zu finden z. B. in Berlin: §§ 64, 65 SchulG Berlin; in Hamburg: §§ 98, 100 HmbSG).

Neben den Lehrkräften können auch Eltern als Empfänger:innen von Daten benannt sein, z. B. zur Verbesserung der Zusammenarbeit zwischen Schule und Eltern. Die Weitergabe von Ergebnissen aus Schulleistungstests und Diagnostik an Eltern ist im deutschen Schulrecht durchaus vorgesehen, aber in Zweck, Form und Tiefe je nach Land unterschiedlich konkretisiert. Explizit regelt dies beispielsweise Nordrhein-Westfalen: Der einschlägige Erlass verpflichtet die Schule, Eltern sowohl über die Einzelergebnisse ihres Kindes als auch über aggregierte Ergebnisse der Lerngruppe und der Schule zu informieren.³⁸ In Berlin regelt das Schulgesetz die Weitergabe von Daten von Vergleichsuntersuchungen (§ 65 SchulG Berlin). Hier erstellt z. B. das ISQ bei VERA 3 für jede Schülerin und jeden Schüler eine individuelle Ergebnisübersicht, die über die Lehrkraft an die Eltern ausgegeben wird.³⁹ Hamburg stellt klar, dass Evaluationsergebnisse in geeigneter Weise zu veröffentlichen sind (§ 100 HmbSG). So erhalten hier Eltern z. B. im Rahmen von KERMIT die Ergebnisse über die Schule.⁴⁰

Nutzung von Daten aus lernprozessbegleitenden und lernstandsbezogenen Erhebungen

In den Ländern werden unterschiedliche Datenerhebungen zu Leistungen und Kompetenzentwicklung der Schüler:innen durchgeführt (vgl. Kapitel 1). Einige Länder konzentrieren sich auf VERA und ggf. noch zwei, drei zusätzliche Instrumente, während andere Länder annähernd zwanzig solcher Instrumente zur Verfügung stellen, davon jedoch die meisten nicht verpflichtend (SWK-Länderabfrage Schule; vgl. Anhang 2).

³⁸ RdErl. d. Ministeriums für Schule und Bildung v. 12.07.2021: <https://bass.schule.nrw/19448.htm> (aufgerufen am 26.02.2026).

³⁹ <https://www.isq.berlin/wordpress/vera-3-individualrueckmeldungen-verfuegbar/> (aufgerufen am 20.02.2026).

⁴⁰ <https://www.kermit-hamburg.de/index.php?action=download&id=641> (aufgerufen am 20.02.2026).



Mit wenigen Ausnahmen werden Eltern in den Ländern, die VERA einsetzen, auch über die individuellen Ergebnisse informiert – bzw. können auf Anfrage darüber informiert werden. Inwieweit eine Einordnung in Klassen-, Schul- oder Landesergebnisse stattfindet, variiert stark, ebenso, ob damit Hinweise zur Unterrichtsentwicklung oder Lernförderung verbunden sind. Dies betrifft auch das Format, in dem darüber informiert wird (z. B. Datenblatt, Bericht oder Gespräch zwischen Lehrkraft und Eltern) (vgl. Abbildung 8). Dieser Befund einer insgesamt häufig unsystematischen Weiterarbeit mit VERA-Ergebnissen deckt sich mit den Ergebnissen der WeSU-Studie, in der an einigen Schulen eher spontane, wenig institutionalisierte Praktiken der Ergebnismutzung beschrieben werden (Weyrauch et al., 2025).

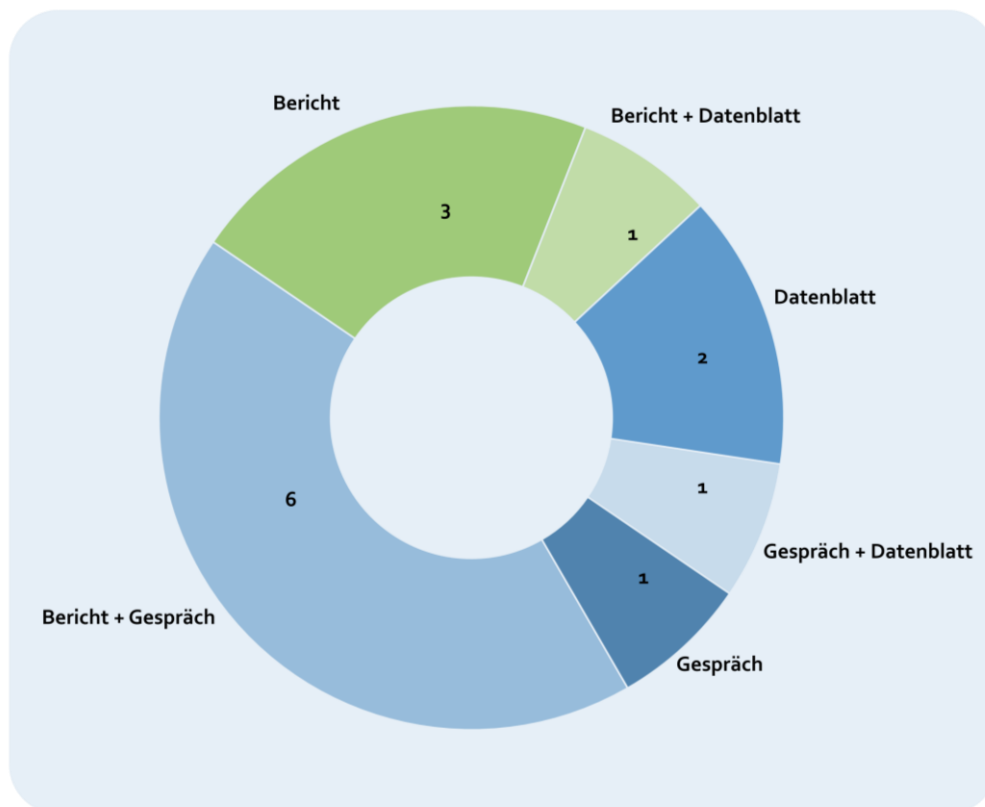


Abbildung 8: Datenrückmeldung an Eltern bei VERA 3 und 8 inkl. genutzter Formate

Anmerkung: Zwei Länder haben keine Angaben zu Rückmeldungen an Eltern gemacht.

Quelle: SWK-Länderabfrage Schule

Neben VERA/KERMIT werden in den Ländern weitere diagnostische oder formative Instrumente verpflichtend umgesetzt (teilweise nur in einzelnen Ländern, z. B. ILeA plus) bzw. zur Nutzung empfohlen (z. B. quop in fünf Ländern, Mathe sicher können in vier Ländern) (s. Anhang 2). Die Anzahl der in den Ländern verfügbaren Instrumente variiert, ebenso der Umfang, in dem Eltern Rückmeldungen erhalten. Die überwiegende Zahl der Länder bietet max. zehn verschiedene Instrumente an. In acht Ländern ist dabei die Rückmeldung von Ergebnissen an die Eltern nur vereinzelt vorgesehen (vgl. Abbildung 9). Über alle Länder und alle Instrumente gemittelt erfolgt nur bei 41 Prozent der Erhebungsinstrumente eine Rückmeldung an die Eltern (SWK-Länderabfrage Schule).



Anzahl der verfügbaren Instrumente in den Ländern	Anzahl der Instrumente mit Rückmeldung an die Eltern			Gesamtzahl der Länder (nach Anzahl der verfügbaren Instrumente)
	< 35 % (= einzelne Instrumente)	35–65 % (= etwa die Hälfte der Instrumente)	> 65 % (= überwiegende Zahl der Instrumente)	
1–5	4 Länder	1 Land	1 Land	6
6–10	2 Länder	2 Länder	3 Länder	7
> 10	2 Länder	1 Land	--	3
Gesamtzahl der Länder (nach Anzahl der Instrumente mit Rückmeldung an die Eltern)	8	4	4	

Abbildung 9: Umfang der Rückmeldungen an Eltern in den Ländern relativ zur genutzten Anzahl der verfügbaren Instrumente

Anmerkungen: Je dunkler die Farbe, auf desto mehr Länder treffen die Randbedingungen der Zelle zu. Übersicht inkl. VERA/KERMIT
Quelle: SWK-Länderabfrage Schule

Beispielhaft kann hier die *Lernausgangslage Berlin (LauBe)*⁴¹ im Schuleingang genannt werden: Lehrkräfte erhalten nach Eingabe der Ergebnisse unmittelbar eine aufgabenspezifische Rückmeldung mit Vergleichswerten in Deutsch und Mathematik. Für die Elternkommunikation ist zentral, dass es ein eigenes Heft mit Aufgaben für alle Kinder der Schulanfangsphase gibt, die Rückmeldungen dienen auch als Grundlage für Gespräche der Schule mit den Erziehungsberechtigten. Hamburg hat mehrere digitale Kurzdiagnosen aufgebaut, die sich für Elterngespräche nutzen lassen, weil sie Ergebnisse in Kompetenzrastern und mit Beispielen rückmelden. Das Instrument *SCHNABEL* erzeugt Individualrückmeldungen zu Rechtschreibkompetenzen. In der Handreichung wird festgehalten, dass diese Rückmeldung auf Wunsch den Erziehungsberechtigten zur Verfügung gestellt und idealerweise im Rahmen eines Lernentwicklungsgesprächs mit ihnen besprochen wird (IfBQ, 2024b).

In Bremen dokumentieren Lehrkräfte mit der für die Grundschule verbindlich eingeführten *Kompetenzorientierten Leistungsrückmeldung* den Lernstand der Kinder zunächst in Entwicklungsübersichten (KMK, 2024c). Auf dieser Basis werden in mindestens zwei Lerngesprächen pro Schuljahr, an denen Eltern, Lehrkraft und Kind teilnehmen, die erreichten Kompetenzen besprochen und die nächsten Lernziele in einer

⁴¹ <https://www.isq.berlin/wordpress/diagnostik/laube/> und <https://bildungsserver.berlin-brandenburg.de/lernausgangslage-laube> (aufgerufen am 05.03.2026).



Lernvereinbarung festgehalten. Ergänzend führt die Schule ein Portfolio als kindgerechtes Lerndokument, in dem ausgewählte Arbeiten und Produkte des Kindes gesammelt werden, sowie einen notenfreien Lernentwicklungsbericht, der die Leistungen des Schuljahres – teilweise in Form von Kompetenzrastern – zusammenfasst und den Eltern ausgehändigt wird. Diese Kernelemente der Leistungsrückmeldung werden durch weitere Diagnosedaten flankiert, etwa den *Primo-Sprachtest* im Vorschulalter, eine zweite Sprachstandserhebung nach der Einschulung, die verpflichtende Teilnahme an VERA 3 (Deutsch, Mathematik) sowie zusätzliche Screeningverfahren in den Jahrgängen 1 und 2.

Inwiefern die Instrumente tatsächlich für die Kommunikation mit Eltern genutzt werden, ist weder flächendeckend dokumentiert noch empirisch erforscht. Man kann wohl eher von einem Potenzial für eine breite Elternkommunikation sprechen als von einer etablierten, systematischen Praxis.

Instrumente und Daten zu überfachlichen Aspekten

Auch überfachliche Daten, insbesondere zu Wohlbefinden, sozialer Eingebundenheit und schulischer Zufriedenheit, aber auch zur Lernmotivation, sind für Erziehungs- und Bildungspartnerschaften wertvoll. So lassen sich u. a. belastende Bedingungen sichtbar machen (z. B. fehlendes Zugehörigkeitsgefühl zur Klasse oder häufige Erschöpfung) und in Erziehungs- und Bildungspartnerschaften thematisieren. In den Ländern wird die Relevanz dieser Aspekte z. B. für die Förderung leistungsschwächerer Schüler:innen adressiert (KMK, 2024c).

So stellt Baden-Württemberg beispielsweise über das Befragungsportal des Instituts für Bildungsanalysen (IBBW) standardisierte Module zum schulbezogenen Wohlbefinden bereit.⁴² Für Schüler:innen ab Klasse 3 bis Sekundarstufe II erfasst ein Fragebogen u. a. positive Einstellungen zur Schule, Zugehörigkeit, soziale Einbindung, Erfahrungen mit Mobbing sowie wahrgenommene Fairness und Unterstützung durch Lehrkräfte. Dazu gehören auch passende Module, die sich an Lehrkräfte und Eltern der Klassen 1 und 2 richten und u. a. das Wohlbefinden der Kinder aus Elternsicht erfassen. Die Ergebnisse dieser Befragungen werden auf Schulebene aggregiert und im Schuldatenblatt als ein zentrales Element datengestützter Qualitätsentwicklung zusammengeführt (vgl. Kapitel 2 und 3). Schulen verfügen damit über eine systematische, empirisch fundierte Grundlage, um mit Elternvertretungen über das erlebte Wohlbefinden der Schülerschaft zu sprechen und daraus gemeinsame Entwicklungsziele für Schule und Unterricht abzuleiten. Gleichzeitig können diese Daten als Kontextfolie in individuelle Lernentwicklungsgespräche einfließen.

Erfassung und Nutzung von Daten zu Fehlzeiten (Absentismus)

Daten zu Fehlzeiten liefern wichtige Hinweise, wo Bildungsteilnahme und Lernen gefährdet sein können. Internationale Analysen zeigen, dass wiederholtes Fernbleiben vom Unterricht deutlich mit geringeren Kompetenzen und erhöhter Drop-out-Wahrscheinlichkeit verknüpft ist (OECD, 2019a). Eine systematische Erfassung von Daten zu Fehlzeiten ist Voraussetzung für vereinbarte Schritte für die Rückkehr, Stabilisierung und Teilhabe – möglichst mit sozialpädagogischer, schulpсихологischer und ggf. Jugendhilfe-Anbindung.

⁴² https://schul.befragung-bw.de/de_DE/start/download.html#Frageboegen-zum-schulbezogenen-Wohlbefinden (aufgerufen am 05.03.2026).



Bundesweit existieren bislang allerdings keine einheitlichen amtlichen Kennzahlen zu Schulabsentismus. Derzeit liegen vor allem regionale Erhebungen vor, aus denen sich Werte von 6 bis 25 Prozent für mindestens einmalig unentschuldigtes Fehlen und 2 bis 4 Prozent für mehr als 10 Tage unentschuldigtes Fehlen ergeben (Kreitz-Sandberg et al., 2022). Die Datenlage ist fragmentiert, und es dürfte eine erhebliche Dunkelziffer geben (vgl. auch Sälzer et al., 2024).

Nachfolgend werden beispielhaft Regelungen in den Ländern zum Thema Absentismus betrachtet:

Hamburg arbeitet mit einer verbindlichen Richtlinie und Handreichung, die ein abgestuftes Verfahren von der ersten unentschuldigten Fehlzeit bis zu anhaltendem Absentismus beschreibt.⁴³ Die Richtlinie verpflichtet Schulen, Anwesenheit in jeder Stunde zu kontrollieren, Fehlzeiten im Klassenbuch bzw. digital zu dokumentieren und ab definierten Schwellenwerten (z. B. mehr als 3 unentschuldigte Tage bzw. 20 Schulstunden in 4 Wochen) standardisierte Schritte einzuleiten. Das dazugehörige Praxishandbuch *Jeder Schultag zählt* (Ricking et al., 2024), auf das Hamburg verweist, stellt Hintergrundwissen zum Schulabsentismus bereit und enthält u. a. Steckbriefe zum Monitoring von An- und Abwesenheit, Musterbriefe an Eltern, Gesprächsleitfäden und Dokumentationshilfen, mit denen Schulen die systematische Auswertung von Fehlzeiten und die Elternkommunikation organisieren können.

Schleswig-Holstein definiert in seinem Konzept *kritische Fehlzeiten* in mehreren Stufen (problematische, gravierende, massive Fehlzeiten) und verbindet diese mit klaren Verfahrensvorgaben: tägliche Anwesenheitskontrolle, systematische Dokumentation, schulinterne Fallbesprechungen, Einbindung der Schulsozialarbeit und – bei Überschreiten der Grenzwerte – Meldung an Schulaufsicht und Jugendhilfe (Ministerium für Bildung, Wissenschaft und Kultur Schleswig-Holstein, 2022). Eltern sollen systematisch in Rückmelde- und Interventionsprozesse einbezogen werden. Seit dem Schuljahr 2022/23 baut Schleswig-Holstein zudem eine landesweite, definitionsgestützte Statistik zu kritischen Fehlzeiten auf, die es erlaubt, die Verteilung von Absentismus über Kreise und Schularten hinweg zu beobachten.

In Nordrhein-Westfalen existieren demgegenüber Leitfäden oder Handreichungen auf Ebene der Städte und Bezirksregierungen (z. B. Bezirksregierung Düsseldorf, 2024; Stadt Mönchengladbach, 2023; eine Übersicht bei der Qualitäts- und Unterstützungsagentur – Landesinstitut für Schule QUA-LiS NRW, o. J.). Neben Hintergrundwissen und Verfahrensvorgaben wird die Dokumentation der Fehlzeiten, auch stundenweise, als wichtiges Mittel zur Prävention formuliert.

Bereitstellung von Instrumenten für Elternbefragungen an Schulen

In mehreren Ländern stellen die Landesinstitute inzwischen für Elternbefragungen Instrumente bereit, die für die Schulentwicklung genutzt werden können – meist in Form standardisierter Fragebögen, die online eingesetzt und ausgewertet werden.

Beispielsweise können über das Selbstevaluationsportal SEP-SCHULE des ISQ Berlin ausgearbeitete Fragebogenbausteine zu Themen wie Ganztage, Inklusion, Gesundheitsförderung oder schulisch angeleitetes

⁴³ <https://www.hamburg.de/resource/blob/813740/9c494610127a78e60240e1e1c4e5bdoi/richtlinie-handreichung-schulabsentismus-data.pdf> (aufgerufen am 26.02.2026).



Lernen zu Hause ausgewählt werden; für jede Themenfamilie liegen jeweils angepasste Bausteine für verschiedene Gruppen vor, darunter auch Eltern. Die Fragebögen werden online erhoben und standardisiert ausgewertet und sind als Instrumente der schulinternen Evaluation und Qualitätsentwicklung konzipiert.⁴⁴ Rheinland-Pfalz verfügt mit *EVA Schule* über ein Befragungsportal, das es Schulen ermöglicht, die Einschätzungen verschiedener Befragtengruppen – darunter Sorgeberechtigte – im Perspektivenvergleich zu erfassen.⁴⁵ In Nordrhein-Westfalen stellt die QUA-LiS mit *LimeSurvey/SIBA* ein datenschutzkonformes Online-Befragungssystem bereit, das für strukturierte, schulspezifisch adaptierbare Evaluationen in Schulen vorgesehen ist; als typische Befragtengruppen werden auch Eltern genannt.⁴⁶ Dies trifft auch auf Baden-Württemberg und das Befragungsportal *BEF-BW* bzw. das *Schul-Tool* zu.⁴⁷

Somit existiert inzwischen eine durchaus ausgebaute Infrastruktur für Elternbefragungen, die von den Landesinstituten in Form von Onlineportalen, Fragebogenpools und Evaluationshandreichungen bereitgestellt wird und explizit der schulinternen Qualitätsentwicklung dienen soll. Untersuchungen zum Nutzungsverhalten liegen noch nicht vor. Studien mit ähnlichen Fragestellungen aus den Niederlanden und Israel lassen vermuten, dass die Nutzung der bereitgestellten Fragebögen stark variiert und von der Evaluationskultur der einzelnen Schule abhängt (Aderet-German & Ben-Peretz, 2020; Schildkamp & Visscher, 2009).

Digitale Kanäle für die Bereitstellung von Daten: Portale und Dashboards

Digitale Portale und Dashboards können eine adressatengerechte Aufbereitung entsprechender Daten unterstützen, indem sie zentrale Informationen bündeln, verständlich visualisieren und Anschlusskommunikation anregen (vgl. Empfehlung 2). Dashboards als Tool zur Darstellung von Schülerdaten sind international in mehreren Ländern (z. B. Kanada) im Einsatz. Für Deutschland zeigt sich ein heterogenes und insgesamt noch frühes Stadium bei der Nutzung von Dashboards zur Darstellung von Schülerdaten: Nur in einer Minderheit der Länder gibt es eine etablierte Nutzung, während mehrere Länder Dashboards erst in Ansätzen erproben und ein weiterer Teil aktuell Dashboards gar nicht nutzt (eigene Erhebungen, SWK-Länderabfrage Schule). Allerdings ist der Begriff Dashboard nicht eindeutig definiert und wird sehr heterogen verwendet. Vor allem bezüglich der Interaktivität der dargestellten Daten unterscheiden sich einzelne Dashboard-Lösungen stark. Ohne eine genaue Betrachtung des Tools, das als Dashboard bezeichnet wird, sind Aussagen zu Dashboard-Lösungen daher kaum vergleichbar. Insgesamt scheinen Dashboards bislang eher punktuelle bzw. pilotierende Lösungen zu sein und werden (wo vorhanden) vor allem als Instrument der internen datengestützten Steuerung in Schule/Administration verstanden, weniger als systematisch ausgebautes Format für die Zusammenarbeit mit Eltern. Eltern als Nutzergruppe von Dashboards stehen nicht im Fokus, noch weniger die Einbindung anderer, schulexterner Nutzergruppen wie z. B. Jugendamt oder Gesundheitsbehörden.

⁴⁴ https://sep-schule.isq.berlin/de_DE/start/index.html (aufgerufen am 20.02.2026).

⁴⁵ https://schule.eva-rlp.de/de_DE/start/index.html (aufgerufen am 20.02.2026).

⁴⁶ <https://www.qua-lis.nrw.de/schulentwicklung-nrw/evaluationdiagnose/evaluation/schulinterne-evaluation/fokus-schulentwicklung-o> (aufgerufen am 20.02.2026).

⁴⁷ https://befragung-bw.de/de_DE/auswahl/portal.html (aufgerufen am 20.02.2026).



4.1.3 Lernentwicklungs- und Förderplangespräche als Struktur für datengestützte Erziehungs- und Bildungspartnerschaften

Daten für Erziehungs- und Bildungspartnerschaften können dann wirksam werden, wenn sie in dialogischen Formaten aufgegriffen, gemeinsam gedeutet und in konkrete Vereinbarungen übersetzt werden. Lernentwicklungs- und Förderplangespräche sind hierfür entsprechende Anlässe: Sie bündeln Informationen aus Unterricht, Diagnostik, Lernstandserhebungen und Beobachtungen und machen sie zum Ausgangspunkt einer gemeinsamen Verständigung zwischen Schüler:innen, Eltern, Lehrkräften und weiteren pädagogischen Fachkräften hinsichtlich konkreter Ziele und Entwicklungsstrategien. Dies gilt grundsätzlich über die gesamte Leistungsspanne hinweg – sowohl für Kinder und Jugendliche mit besonderen Unterstützungsbedarfen als auch für leistungsstarke und (hoch-)begabte Schüler:innen, deren Potenziale im Schulalltag häufig unzureichend in den Blick kommen.

Lernentwicklungsgespräche

Lernentwicklungsgespräche (LEG) als Form der Leistungsbeurteilung sind vor allem in der Primarstufe verankert (z. B. Baden-Württemberg, Bayern). Teilweise ist ein Beschluss der Schulkonferenz Voraussetzung dafür, ggf. ergänzt um die Zustimmung weiterer Gremien oder Instanzen wie z. B. der Fachaufsicht, teilweise ist es verpflichtend vorgesehen. In einer Reihe von Ländern liegen allerdings auch Regelungen vor, die bis in die Sekundarstufe I LEG mindestens in Ergänzung (z. B. Hamburg, Rheinland-Pfalz) oder sogar statt Notenzeugnissen erlauben (Berlin bis Stufe 9 in Integrierten Sekundarschulen und Gesamtschulen) (vgl. Abbildung 10 und Anhang 7). Auch hier bestehen unterschiedliche Verpflichtungsgrade. Zusätzlich zu den Lernentwicklungsgesprächen sind in den Ländern Beratungsgespräche vor Übergängen, bei Versetzungsproblemen oder disziplinarischen Problemen vorgesehen. Nehmen Eltern an einem Gespräch nicht teil, muss die Schule ggf. andere Formen der Rückmeldung nutzen (Zeugnis, schriftliche Empfehlung). Erst wenn sich eine gravierende Gesamtsituation für das Kind ergibt (z. B. Vernachlässigung, massiver Schulabsentismus), werden Jugendhilfe oder weitere Stellen relevant.

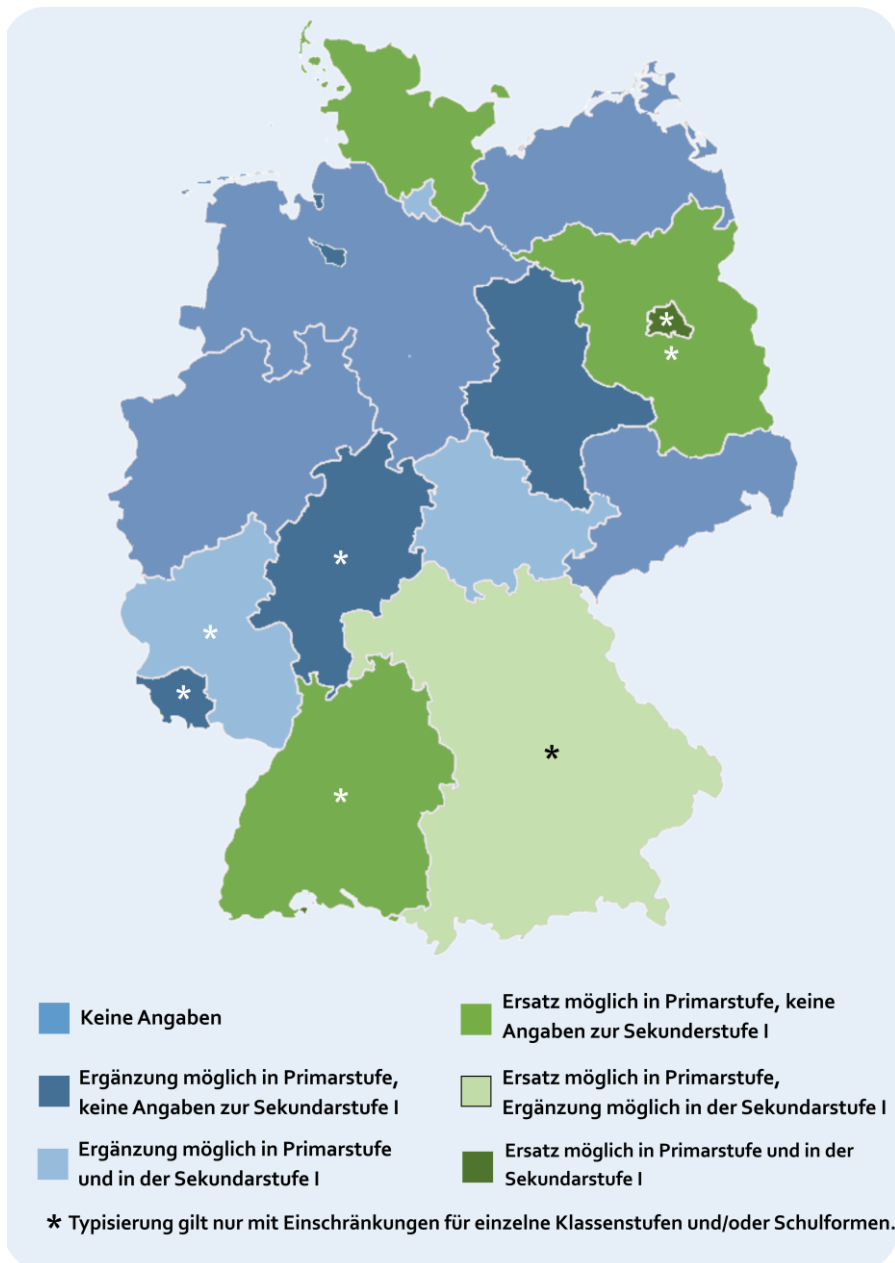


Abbildung 10: Rechtlich geregelte Möglichkeiten zur Nutzung von Lernentwicklungsgesprächen als Alternative zum Halbjahreszeugnis

Anmerkung: für weitere Details und eine Übersicht über die entsprechenden Rechtsnormen siehe Anhang 7

Quelle: eigene Recherchen

In der Grundschule verankert beispielsweise Berlin *lernprozessbegleitende Gespräche* im Schuljahr, mit dokumentierten Absprachen und Ausrichtung auf nächste Lernschritte. Bremen stärkt eine datengestützte Gesprächskultur über Portfolios und Kompetenzübersichten, wodurch kriteriale Erwartungen und individuelle Lernwege zusammengeführt werden. Rheinland-Pfalz etabliert das *Lehrer-Schüler-Eltern-Gespräch* als Standardinstrument, das nicht nur informiert, sondern konkrete Zielvereinbarungen festhält. Thüringen verbindet in der Grundschule feste Zeitpunkte über Entwicklungsphasen hinweg mit einer Beratungslogik: persönliche, fachliche und soziale Entwicklung werden gemeinsam reflektiert und in verbindlichen Absprachen fixiert.



Daneben stehen Länder, die LEG mit Zeugnisformaten verknüpfen. Baden-Württemberg, Bayern, Sachsen-Anhalt und – in einer spezifischen Erstjahrgangsregel – Schleswig-Holstein kennen Modelle, in denen ein dokumentiertes Gespräch die Halbjahresinformation bzw. das Halbjahreszeugnis ersetzen kann. Diese Strategie nutzt das Gespräch als primäre Rückmeldestruktur; die Entscheidung liegt typischerweise bei schulischen Gremien, teils mit Opt-out-Rechten für Eltern. Inhaltlich sind diese LEG weniger Gespräch über Noten, sondern strukturierte Rückblicke und Ausblicke: Wo steht das Kind, was zeigt das Portfolio, welche Förderimpulse folgen daraus – und wer macht was bis wann?

Andere Länder belassen es bei allgemeinen Gesprächs- und Beratungspflichten rund um Zeugnisse oder Elternsprechtage. In Niedersachsen sind in Klasse 4 mindestens zwei Beratungsgespräche zur Lernentwicklung und zum Übergang verbindlich. In Nordrhein-Westfalen, Hessen, Mecklenburg-Vorpommern, im Saarland, in Sachsen, Sachsen-Anhalt (Sekundarstufe) und Schleswig-Holstein werden Elterngespräche regelmäßig praktiziert, aber landesweite Pflicht-LEG (losgelöst von Zeugnissen oder besonderen Anlässen) sind nicht der Regelfall; hier prägen anlassbezogene Beratungen die Praxis.

In der Sekundarstufe klappt das Feld weiter auseinander. Hamburg schreibt mindestens ein LEG pro Jahr schulartübergreifend fest und hält an der triadischen Beteiligung von Lehrkraft, Lernenden und Eltern fest. Thüringen sichert LEG über Doppeljahrgänge (5/6 und 7/8) ab – mit dokumentierten Zielen und Anschlussfähigkeit an Fach- und Laufbahnberatung. Rheinland-Pfalz ermöglicht, dokumentierte Lehrer-Schüler-Eltern-Gespräche als Ersatz für klassische Elternsprechtage zu nutzen. In den übrigen Ländern dominieren in der Sekundarstufe strukturierte Elternsprechtage, Laufbahn- und Fördergespräche ohne ausdrücklich landesweit verpflichtendes LEG-Raster.

In der Praxis stützen sich Lehrkräfte bei der Vorbereitung und Durchführung von LEG in der Regel auf Einschätzungsbögen und Gesprächsleitfäden, in denen fachliche und überfachliche Kompetenzen dokumentiert werden; häufig werden jene durch Portfolios oder ausgewählte Arbeitsproben ergänzt. In einigen Ländern werden hierfür landesweit Materialien bereitgestellt – etwa Dokumentationsbögen und Handreichungen in Bayern oder der Leitfaden zum Lehrer-Schüler-Eltern-Gespräch in Rheinland-Pfalz –, die typische Gesprächsphasen, mögliche Formulierungen für Rückmeldungen und Zielvereinbarungen sowie Beispiele für Kompetenzraster strukturieren. In Hamburg sind LEG verankert, die sich explizit auf dokumentierte Lernbeobachtungen, Leistungsdaten und überfachliche Kompetenzen stützen sollen; flankierend liegen Musterbögen und Hinweise zur Gesprächsführung vor. Daneben zeigt sich in der Praxis eine große Bandbreite schuleigener Formulare und Abläufe, die von stark strukturierten, kompetenzrasterbasierten Verfahren bis hin zu eher frei geführten Gesprächen reicht.

Beobachtungs- und Befragungsstudien in der Grundschule zeigen, dass LEG grundsätzlich von den Gesprächsbeteiligten sowie von externen Beobachter:innen als lernunterstützend eingeschätzt werden (Dollinger, 2019), dabei aber zwischen Klassen und Kindern variieren (Ertl, Kücherer & Hartinger, 2022). Lehrkräfte und Eltern schätzen LEG in der Studie im Übrigen auch als gewinnbringender als ein Zwischenzeugnis ein. Rückmeldungen zum aktuellen Lernstand erfolgen allerdings häufiger als Rückmeldungen zur Lernentwicklung. Auch Hinweise auf Lernstrategien bzw. konkrete Übungen werden vergleichsweise selten gegeben (Dollinger, 2019).



Mit Blick auf Gesprächsführung und Partizipation zeigt sich, dass LEG häufig stark von der Lehrkraft gesteuert werden. Zwar sind Schüler:innen und Eltern formal beteiligt, in der Interaktion dominieren jedoch lehrkraftzentrierte Gesprächssequenzen; Beiträge der Schüler:innen und Eltern werden teilweise eher in bestehende Deutungsmuster eingebunden, als dass sie gleichberechtigt verhandelt würden (Bonanati, 2015). Zugleich erleben Schüler:innen die Gespräche insbesondere dann als sinnvoll und motivierend, wenn sie in einer wertschätzenden Atmosphäre aktiv einbezogen werden, konkrete Rückmeldungen zum eigenen Lernen erhalten und an der Zielformulierung beteiligt sind (Ertl, Kücherer & Hartinger, 2022; Häbig, 2018).

In der Gesamtschau lässt sich festhalten, dass LEG in mehreren Ländern in unterschiedlichen Formen als Instrument der Leistungs- und Entwicklungsrückmeldung sowie individuellen Förderung verankert sind, die tatsächliche Durchführung in der Schulpraxis jedoch ein breites Spektrum an Umsetzungsformen aufweist. Die vorliegenden Studien verweisen auf ein grundsätzlich hohes Potenzial der Gespräche für individualisierte Förderung, Motivation und Bildungsbeteiligung. Zugleich ist dies eng an die Qualität der Rückmeldung, der Gesprächsstruktur und der tatsächlichen Beteiligung von Schüler:innen und Eltern sowie der Orientierung an Zielen gebunden (Ertl, Kücherer & Hartinger, 2022; eher praxisorientierte Hinweise finden sich bei Dollinger & Hartinger, 2020; Ertl et al., 2021).

Förderplangespräche

Während LEG grundsätzlich alle Schüler:innen in den Blick nehmen, gibt es bei einem Teil der Kinder und Jugendlichen Hinweise auf besondere Unterstützungsbedarfe – sei es z. B. aufgrund diagnostizierter Beeinträchtigungen oder auch ausgeprägter Begabungen. In diesen Fällen reicht eine allgemeine, zyklische Lernstandsrückmeldung allein nicht aus; erforderlich ist eine multiprofessionell gestützte Förderplanung, die Ziele, Maßnahmen, Zuständigkeiten und Überprüfungszeiträume verbindlich festlegt. Diese bezieht sich in den meisten Bundesländern primär auf sonderpädagogische Unterstützungsbedarfe.

Förderplangespräche bilden hierfür eine zentrale kommunikative Struktur (vgl. auch SWK, 2022a). Sie vertiefen die Analyse des Unterstützungsbedarfs und übersetzen sie in konkrete, überprüfbare Förderziele. Popp, Melzer und Methner betonen, dass qualitativ gute Förderpläne immer an eine systematische Diagnostik, klare Zielformulierungen, konkrete Maßnahmenplanung und regelmäßige Evaluation gekoppelt sind (Popp et al., 2023). Neben schulischen Leistungsergebnissen sollten Daten zu Lern- und Lebensbedingungen, Beobachtungen aus Ganztags- und Förderangeboten sowie Einschätzungen der Schüler:innen selbst in die Zielentwicklung einfließen. Popp, Melzer und Methner heben hervor, dass Förderplanung als kooperativer Prozess gedacht ist, in dem Lehrkräfte, Sonderpädagog:innen, Schulsozialarbeit und ggf. weitere Fachkräfte gemeinsame Zielvereinbarungen treffen und Zuständigkeiten klären (vgl. Popp et al., 2023).

Für die digitale Förderplanung stehen unterschiedliche Tools zur Verfügung. In Sachsen wird beispielsweise mit *DigiDuF (Digitales Verfahren zur sonderpädagogischen Diagnostik und Förderung)*⁴⁸ ein landesweit eingeführtes Fachverfahren genutzt, das von allen allgemeinbildenden Schulen eingesetzt wird. DigiDuF stellt strukturierte Vorlagen für Förderpläne bereit, regelt Rechte- und Rollenvergabe und soll Transparenz und

⁴⁸ <https://www.foerderdiagnostik.bildung.sachsen.de/aufbau-des-fachverfahrens-digiduf-digitales-verfahren-zur-sonderpaedagogischen-diagnostik-und-foerderung-4208.html> (aufgerufen am 26.02.2026).



Nachvollziehbarkeit für alle am Förderprozess Beteiligten sicherstellen. Daneben nutzen Schulen in mehreren Ländern – teils mit Landeslizenz – digitale Anwendungen wie *SPLINT*⁴⁹, eine Web-App zur Förderplanung und Lernbegleitung. Im Rahmen der Bund-Länder-Initiative *Leistung macht Schule* wurden darüber hinaus Instrumente zur personalisierten Entwicklungsplanung und zur begabungs- und leistungsfördernden Schulentwicklung entwickelt (*LemaS-Toolbox*)⁵⁰. Nicht zuletzt eröffnet der Einsatz von KI neue Möglichkeiten für die individuelle Förderplanung. Studien aus dem Bereich der sonderpädagogischen Förderplanung zeigen, dass KI-gestützte Systeme Lehrkräfte etwa bei der Formulierung von Förderzielen unterstützen können (Waterfield et al., 2026).

4.2 Konzepte und Maßnahmen

Bislang scheint es nicht so, dass es auf Länderebene umfassende Strategien für (datengestützte) Erziehungs- und Bildungspartnerschaften gibt bzw. diese systematisch bei der Entwicklung von Instrumenten, Konzepten und Materialien mitgedacht werden. Lernentwicklungs- und Förderplangespräche sind in mehreren Ländern rechtlich verankert, weisen in der Praxis jedoch eine große Spannweite an Verfahren auf und sind noch nicht durchgängig mit einer datengestützten Strategie verbunden. Hinzu kommen Schnittstellenprobleme zwischen Schule und Jugendhilfe, die multiprofessionelle, dateninformierte Kooperation erschweren. Nachfolgend werden internationale Strategien sowie Maßnahmen im Umgang mit zentralen Herausforderungen (u. a. Verständlichkeit von Daten, Professionalisierungsanforderungen, Schnittstellengestaltung zwischen Rechtskreisen) skizziert.

4.2.1 Übergreifende Strategien für datengestützte Partnerschaften: internationale Beispiele

Internationale Beispiele (z. B. aus Alberta/Kanada, Finnland und New South Wales/Australien) zeigen, dass Eltern bei datengestützter Schulentwicklung integral mitgedacht werden. Dabei ist leitend, dass elterliches Erfahrungs- und Kontextwissen in gemeinsame Deutungs-, Entscheidungs- und Zielbildungsprozesse einbezogen wird und dass die Agenda und die intendierten Outcomes der Zusammenarbeit nicht ausschließlich von der Schule gesetzt, sondern als gegenseitig bestimmt und für beide Seiten bedeutsam verstanden werden (Pushor, 2012). In Alberta/Kanada ist Elternbeteiligung in eine Gesamtstrategie datengestützter Schulentwicklung eingebettet.⁵¹ Mehrere Aspekte greifen ineinander: die rechtlich verankerte Mitwirkung von Eltern auf Schulebene, ein jährlich eingesetztes Datenerhebungs- und Rückmeldesystem (inkl. Elternperspektiven), ein öffentliches datengestütztes Berichtswesen und eine verbindliche, datengestützte Gesprächspraxis in den Schulen. Der *Education Act*⁵² verpflichtet jede Schule, ein School Council einzurichten. Diese Räte erhalten beispielsweise Ergebnisberichte (einschließlich standardisierter Testergebnisse) und werden in Planungs- und Reportingprozesse einbezogen. Gremien werden entsprechend qualifiziert (mit

⁴⁹ <https://splint.schule/> (aufgerufen am 26.02.2026).

⁵⁰ <https://lemas-forschung.de/glossar/toolbox/> (aufgerufen am 26.02.2026).

⁵¹ <https://www.alberta.ca/accountability-education-system> (aufgerufen am 20.02.2026).

⁵² <https://www.albertaschoolcouncils.ca/education-in-alberta/the-education-act> (aufgerufen am 20.02.2026).



Leitfäden und Workshops). Der *Education Assurance Survey*⁵³ (ein jährliches, standardisiertes Befragungsinstrument für Eltern, Lehrkräfte und Schüler:innen) fragt u. a. nach der Zufriedenheit mit der Bildungsqualität, inwieweit die Lernumgebungen als einladend, fürsorglich, respektvoll und sicher empfunden werden sowie nach *Parent Involvement*. Ergebnisse werden ausschließlich aggregiert berichtet und öffentlich einsehbar mit Zielen und Maßnahmen verknüpft.

In vielen Schulbezirken in Alberta sind ab Kindergarten bis zum Ende der Junior High (Klassenstufe 9) *Student-led* bzw. *Three-way Conferences* (Lehrkraft–Schüler:in–Eltern) als zentrales Format der Lernstands- und Entwicklungsrückmeldung etabliert, häufig zweimal im Schuljahr.⁵⁴ In der Sekundarstufe kommen – je nach Schule – zusätzliche, oft fachspezifische Sprechzeiten und Beratungstermine hinzu. Die Vorbereitung dieser Gespräche stützt sich typischerweise auf mehrere, unterrichtsnahe Evidenzen: Leistungsdaten aus Tests und kurzen Kompetenzchecks, ausgewählte Verlaufsdaten zu Schlüsselbereichen (z. B. Lesen, Mathematik), Anwesenheitsinformationen sowie – in den Testjahrgängen – Ergebnisse der *Provincial Achievement Tests* und der *Diploma Exams*,⁵⁵ zu denen Eltern spezifische Informationsmaterialien mit Erläuterungen des Formats, Beispielaufgaben und Beschreibungen der erwarteten Lernziele erhalten. Betont wird auch, dass die Ergebnismeldungen dazu genutzt werden sollen, gemeinsam konkrete nächste Schritte für weitere Fördermaßnahmen in Schule und Elternhaus zu vereinbaren. Die Provinz regelt den Umgang mit personenbezogenen Daten über den *Freedom of Information and Protection of Privacy Act*. Parallel dazu gibt es abgestufte Zugriffsrechte auf Dashboards, um u. a. elternrelevante Informationen aktuell und zugleich datensparsam bereitstellen zu können (Klopsch & Sliwka, 2020).

Sliwka und Klopsch haben das kanadische Modell für Deutschland adaptiert. Die *Kernroutine zur Sicherung basaler Kompetenzen* verbindet datengestützte Diagnostik, kooperative Förderplanung und strukturierten Austausch mit Eltern (Sliwka et al., 2023). Auf eine digital gestützte Lernausgangserhebung zu Schuljahresbeginn folgen die Auswertung im multiprofessionellen Team, ein verbindliches Elterngespräch mit visualisiertem Infoblatt und Protokoll sowie gemeinsame Zielvereinbarungen als Grundlage für die weitere Förderung; den Abschluss bildet ein formatives Assessment mit individueller und kriterialer Bezugsnorm. Daten werden dabei so aufbereitet und genutzt, dass sie dialogische Erziehungs- und Bildungspartnerschaften anregen können.

In Finnland legt das *National Core Curriculum for Primary and lower Secondary (Basic) Education*⁵⁶ u. a. fest, dass Schüler:innen und Erziehungsberechtigte regelmäßig über Fortschritte informiert werden. Die Stadt Tampere gibt beispielsweise vor, dass Lehrkräfte einmal pro Schuljahr Eltern und Schüler:innen zu einem Beurteilungs- bzw. Entwicklungsgespräch einladen. Besprochen werden Lernstand, Stärken bzw. Entwicklungsfelder und die aktuelle Schulsituation; Dolmetschen ist auf Wunsch möglich. Neben Fachleistungen werden Arbeits- und Lernverhalten, altersangemessene Themen sowie gemeinsam vereinbarte Ziele

⁵³ <https://www.alberta.ca/alberta-education-and-childcare-assurance-survey> (aufgerufen am 20.02.2026).

⁵⁴ <https://www.alberta.ca/success-in-school-for-children-and-youth-in-care> (aufgerufen am 20.02.2026).

⁵⁵ <https://www.alberta.ca/k-12-provincial-assessment> (aufgerufen am 20.02.2026).

⁵⁶ <https://www.oph.fi/en/education-and-qualifications/national-core-curriculum-primary-and-lower-secondary-basic-education> (aufgerufen am 20.02.2026).



besprochen.⁵⁷ Die zweijährliche *School-Health-Promotion-Studie*⁵⁸ erhebt flächendeckend Daten zu psychischer Gesundheit, sozialer Eingebundenheit, Lernbelastung und Zufriedenheit; Kommunen nutzen die Ergebnisse für Schulentwicklung, Prävention und die Information von Eltern- und Schüलगremien. In vielen Kommunen läuft die direkte Abstimmung mit Eltern über ein *School-Home-Portal*; Eltern können Lernfortschritte einsehen, Entschuldigungen übermitteln und Benachrichtigungen zu Fehlzeiten erhalten.⁵⁹

In New South Wales (NSW)/Australien wird empfohlen, verpflichtende Gespräche mit Eltern u. a. als *Three-way Learning Conferences* zu führen.⁶⁰ Außerdem können – ggf. unter Einbindung professioneller Unterstützung – verbindliche Entwicklungspläne vereinbart werden. Zusätzlich nutzt NSW systemische Klima- und Beteiligungsdaten. Die *Tell-Them-From-Me*-Befragung erhebt die Schüler:innen- und auch Elternperspektiven zu Schulklima, Zugehörigkeit und Lernengagement.⁶¹ Schulen binden diese aggregierten Indikatoren in ihre Schulpläne ein. Die gewonnenen Erkenntnisse werden u. a. zur Adressierung von Absentismus genutzt.⁶²

Neben Finnland gibt es auch in Dänemark, England, Estland und in einzelnen Kantonen der Schweiz Beispiele für die Umsetzung datengestützter Erziehungs- und Bildungspartnerschaften (Hertweck, 2024). Während in Dänemark und England der Fokus auf Leistungsvergleichen und Bildungsstatistiken liegt, die Eltern schulscharf zur Verfügung gestellt werden, sind in Estland darüber hinaus auch Aspekte wie Schülerzufriedenheit integriert.

4.2.2 Umgang mit Herausforderungen datengestützter Partnerschaft

Besonders Eltern in sozial benachteiligten Lagen und/oder mit Migrationsgeschichte werden von Schulen seltener und vielfach nur anlassbezogen erreicht; Kontakte konzentrieren sich häufig auf Krisen- und Bewertungsanlässe statt auf kontinuierlich geteilte Information und eine gemeinsame Verantwortung für Lernen und Entwicklung (Chamakalayil et al., 2022; Otterpohl & Wild, 2017). Es stellt sich die Frage, wie Barrieren der Zusammenarbeit zwischen Schule und Eltern bearbeitet werden können.

Willkommenskultur und Beziehungsqualität datengestützt entwickeln

Der von einer wissenschaftlichen Expertenkommission entwickelte Kompass *Qualitätsmerkmale schulischer Elternarbeit* betont als erstes Qualitätsmerkmal eine einladende, respektvolle Schulkultur, in der Eltern sich als Teil der Schulgemeinschaft wahrnehmen (Vodafone Stiftung Deutschland, 2013). In diesem Zusammenhang können Schulen auch Daten zu Schulklima, Erfahrungen und Beteiligungsmustern erheben (z. B.

⁵⁷ [https://www.tampere.fi/sites/default/files/2022-09/Perusopetuksen%20oopas_englanti_RGB%20\(1\).pdf](https://www.tampere.fi/sites/default/files/2022-09/Perusopetuksen%20oopas_englanti_RGB%20(1).pdf) (aufgerufen am 20.02.2026).

⁵⁸ <https://thl.fi/en/statistics-and-data/statistics-by-topic/social-services-children-adolescents-and-families/well-being-of-children-and-youth-people-school-health-promotion-study> (aufgerufen am 20.02.2026).

⁵⁹ Zum Beispiel über die in Finnland von vielen Schulen genutzte Wilma-App: <https://www.wilma.fi/>; <https://www.vantaa.fi/en/wilma-communication-between-home-and-school> (aufgerufen am 20.02.2026).

⁶⁰ <https://education.nsw.gov.au/teaching-and-learning/curriculum/reporting-to-parents/reporting-advice#Communicating1> (aufgerufen am 20.02.2026).

⁶¹ <https://education.nsw.gov.au/about-us/education-data-and-research/cese/publications/research-reports/supporting-school-improvement-using-ttfm-surveys> (aufgerufen am 20.02.2026).

⁶² <https://education.nsw.gov.au/about-us/education-data-and-research/cese/publications/research-reports/understanding-attendance> (aufgerufen am 20.02.2026).



Elternbefragungen, Teilnahmestatistiken nach Gruppen), um systematisch zu identifizieren, welche Eltern mit welchen Formaten nicht erreicht werden und welche Barrieren sie benennen. Niederschwellige, familienfreundliche Strukturen – z. B. flexible Zeitfenster, Kinderbetreuung, vertraute Sozialräume, mehrsprachige Formate – binden insbesondere Eltern in benachteiligten Lagen besser ein (Lokhande et al., 2014). Eine besondere Rolle können dabei sozialraumorientierte Familienzentren spielen: Sie bündeln Bildungs-, Beratungs- und Begegnungsangebote und können insbesondere auch Familien in belasteten Lebenslagen erreichen; zugleich können sie dazu beitragen, Schule, Jugendhilfe und weitere Akteure in kommunalen Präventionsketten zu vernetzen (Born et al., 2019).

Verständliche und bezugsnormsensible Kommunikation über Daten

Relevant sind vor allem individuelle Daten mit Bezug zur konkreten Lern- und Lebenssituation des Kindes; sie machen Lernstände und nächste Schritte anschlussfähig und unterstützen gemeinsame Förderentscheidungen (Black & Wiliam, 2009; Brookhart, 2011; Panadero et al., 2018).

Zentral ist dabei die Bezugsnormorientierung: Soziale Vergleiche richten die Aufmerksamkeit auf das Selbstbild und die eigene Person und können Motivation und Selbstkonzept beeinträchtigen, während individuelle und kriteriale Bezüge den Fokus auf Aufgaben bzw. Lernprozesse richten und die Akzeptanz von Zielen fördern (Dickhäuser & Rheinberg, 2003; Jungjohann & Gebhardt, 2022; Rheinberg, 1980). Experimentelle Studien zeigen, dass aufgaben- und prozessorientiertes Feedback, das konkrete nächste Schritte markiert, lernwirksamer ist als Notenrückmeldungen (Butler, 1987, 1988; Harks et al., 2014; Hattie & Timperley, 2007). Standards- bzw. kompetenzbasierte Rückmeldungen mit klaren Kriterien und Beispielaufgaben erhöhen Verständlichkeit und Zielklarheit auch für Eltern (Brookhart, 2011; Guskey & Jung, 2013). Eine wachstumsorientierte Haltung, die vorhandene Lernpotenziale und nächste machbare Schritte betont, knüpft an diese Logik an und stärkt Selbstwirksamkeit (Dweck, 2006; Yeager & Dweck, 2012).

Verschiedene Studien zeigen, dass Qualität und Regelmäßigkeit der Kommunikation zentral für wirksame Bildungs- und Erziehungspartnerschaften sind. Sie belegen positive Zusammenhänge zwischen elterlicher Beteiligung und Schulleistungen sowie sozial-emotionaler Anpassung – insbesondere bei informierter, autonomiestützender Unterstützung zu Hause (Barger et al., 2019; Erdem & Kaya, 2020; Hill & Tyson, 2009; Jeynes, 2014). Kurze, regelmäßige und personenbezogene Nachrichten von Lehrkräften an Eltern erhöhen das elterliche Engagement und senken Leistungsrisiken bei den Schüler:innen (Doss et al., 2019; Kraft & Rogers, 2015). Leitfäden empfehlen dementsprechend klare, jargonfreie Sprache, kurze und zeitnahe Rückmeldungen, einfache Visualisierungen und konkrete Vorschläge für nächste Schritte im Alltag (EEF, 2018). Visuell gestützte Darstellungen (Ampeln, einfache Verlaufsdiagramme) und Materialien verbessern das Verständnis von Leistungs- und Rückmeldedaten (Graham-Clay, 2024; Yotyodying & Wild, 2019, *Effective Family-School Communication*). Mehrsprachigkeit ist Voraussetzung für gleichberechtigte Teilhabe. Studien deuten darauf hin, dass technologisch gestützte Übersetzung Sprachbarrieren in Elterngesprächen wirksam senken kann (Lawal, 2021).

In diesem Zusammenhang können Dashboards für Erziehungs- und Bildungspartnerschaften hilfreich sein, wenn sie für Eltern niederschwellig zugänglich sind, zentrale Informationen verständlich visualisieren und konkrete Anschlusskommunikation anregen (Data Quality Campaign, 2008). Gleichzeitig profitieren



insbesondere Eltern mit geringeren ökonomischen und bildungsbezogenen Ressourcen von komplexen Dashboards oft weniger und können sich von Datenfülle und Fachsprache überfordert fühlen (Curran et al., 2024). Für die Nutzung sind daher eine adressatengerechte Portalarchitektur (z. B. eigene Einstiege für Eltern) sowie die Flankierung durch ergänzende Kanäle wie Kurzbriefe in einfacher Sprache und mehrsprachig oder persönliche und digitale Sprechstunden zentral (Data Quality Campaign, 2008; Hernandez, 2012).

Damit Leistungsrückmeldungen aus Klassenarbeiten für Eltern relevante Einsichten liefern, braucht es zudem transparente Kriterien und aufgabennahe Beispiele. *Rubrics* (Kriterien- bzw. Kompetenzraster) sind hierfür ein bewährtes Instrument: In übersichtlicher Tabellenform wird beschrieben, wie unterschiedliche Niveaustufen einer Leistung konkret aussehen. Rubrics erhöhen die Verständlichkeit, unterstützen faire, kriteriale Bewertung und erleichtern gezieltes *Feed-forward* im Sinne konstruktiver Vorschläge (Brookhart, 2011; Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). Typische Fehlermuster können erläutert werden. Solche Rückmeldungen wirken vor allem dann, wenn sie spezifisch, zeitnah und auf die nächsten machbaren Schritte gerichtet sind (Hattie & Timperley, 2007; Sadler, 1989; Shute, 2008).

Datengestützte Stärkung häuslicher Lernunterstützung

Empirische Befunde legen nahe, dass vor allem die Qualität und weniger die Quantität elterlicher Beteiligung bzw. Mitwirkung relevant ist (vgl. auch SWK, 2022a, 91 ff.). Positive Effekte zeigen vor allem autonomiestützende, wertschätzende und lernprozessorientierte Formen wie etwa die Unterstützung der akademischen Sozialisation (Gespräche über Lernen und Zukunft), realistische Leistungserwartungen, häusliches Lesen und unterstützende Lernbegleitung (Boonk et al., 2018; Hill & Tyson, 2009; Jeynes, 2005; X. Ma et al., 2016; Yotyodying & Wild, 2016). Eine Metaanalyse zeigt, dass elterliche Mitwirkung in kleinen, aber robusten positiven Zusammenhängen mit schulischer, sozialer und emotionaler Anpassung steht (Barger et al., 2019). Direkte Hausaufgabenhilfe ist nicht per se förderlich; kontrollierende Muster können nachteilig sein, während strukturierende und autonomiestützende Unterstützung mit besseren Ergebnissen einhergeht (Boonk et al., 2018; Dumont et al., 2014; Moroni et al., 2015).

Internationale Leitfäden empfehlen kurze Elternbriefe, Checklisten und gemeinsame Zielvereinbarungen, die sich auf konkrete, durch Daten belegte Lernziele beziehen (EEF, 2018; OECD, 2012). Ein Beispiel für elternorientierte Aufbereitung von Diagnosedaten ist die Plattform *LONDI*, die evidenzbasierte Diagnostik und Förderhinweise zu Lese-, Rechtschreib- und Rechenschwierigkeiten bündelt und wissenschaftlich fundierte, niedrighschwellige Informationen explizit auch für Eltern bereitstellt.

Professionalisierung für datengestützte Kommunikation mit Eltern

Mapps und Kuttners *Dual Capacity-Building Framework for Family–School Partnerships* (2013) zeigt, dass familienbezogene Initiativen dann wirksam sind, wenn sie zugleich Kompetenzen aufseiten der Pädagog:innen (z. B. für dialogische, ressourcenorientierte Gespräche, Dateninterpretation, interkulturelle Sensibilität) und aufseiten der Eltern (z. B. für den Umgang mit Feedback, Verständnis von Leistungsanforderungen, Nutzung von Daten für Entscheidungen) aufbauen (vgl. auch die im SWK-Gutachten zur Grundschule beschriebenen Trainings, 2022a, S. 99).



Wild (2021) zeigt in ihrer Expertise, dass es in Deutschland bislang nur punktuell systematische Professionalisierungsangebote zum Thema Eltern gibt, und plädiert für verbindliche Module in der Lehramtsausbildung, in denen diagnostische und beraterische Kompetenzen verbunden werden. Das *BiProfessional*-Projekt zur Diagnose- und Beratungskompetenz etwa arbeitet mit simulierten Elterngesprächen auf Basis realer Diagnosedaten und zeigt, dass Studierende dadurch sicherer im Umgang mit sensiblen Befunden und in der Kommunikation gegenüber Eltern werden. Vertrauen aufseiten der Eltern entsteht, wenn Lehrkräfte gut begründet und konsistent handeln, bereit sind zuzuhören und auch schwierige Themen transparent ansprechen. Gerade in Gesprächen, die auf Daten zum Lernen, Verhalten oder zur Entwicklung der Kinder Bezug nehmen, ist dieses Vertrauensverhältnis von großer Bedeutung (Bormann et al., 2022; Fabel-Lamla, 2022). Ein besonderer Fokus sollte angesichts einer zunehmend heterogenen Schülerschaft auf der Bereitschaft liegen, sich auf unterschiedliche familiäre Lebenslagen, kulturelle Deutungsmuster und alltagsweltliche Erfahrungen einzulassen und Eltern – auch bei bestehenden kulturellen Differenzen – in die Schule einzuladen (Grüter, 2025; zu Habitussensibilität: Weitekämper & Rutter, 2023).

Sozialräumliche Daten können die Elternarbeit substanziell stärken, wenn Schulen sie nicht als Etikett für Problemquartiere, sondern als Ansatz für eine möglichst passgenaue, wertschätzende Ansprache verstehen. Ihr besonderer Mehrwert liegt darin, Lern- und Lebenslagen nicht nur fallbezogen, sondern im Umfeld zu lesen: Wo wohnen die Schüler:innen, wie sind Wege, Sprachen, Betreuungsangebote, Vereinsstrukturen, Gesundheits- und Sozialdienste verteilt, welche Treffpunkte prägen den Alltag? Diese kontextsensible Sicht hilft, Elternkommunikation realistisch zu planen (Sprach-, Zeit- und Ortsbarrieren), niedrigschwellige Formate zu entwickeln und Unterstützungsnetze sichtbar zu machen. Ziel ist es nicht, Defizite zu markieren, sondern vorhandene Ressourcen sichtbar zu machen und gemeinsam mit Schüler:innen und Eltern zu arbeiten (Forell, 2023; Forell & Schuchardt, o. D.). Gleichzeitig erfordert dies eine kritische Reflexion eigener normativer Vorstellungen darüber, wie gute Eltern agieren sollten, um einem defizitorientierten Blick entgegenzuwirken (Betz et al., 2017).

International gibt es zudem Elternworkshops, in denen z. B. das Verstehen von Lernstandsrückmeldungen unterstützt wird; das *APTT-Modell* (*Academic Parent–Teacher Teams*) verknüpft hierzu Lernzieldaten, gemeinsame Zielvereinbarungen und konkrete häusliche Aktivitäten (Paredes, 2011).

Datengestützte multiprofessionelle Zusammenarbeit von Schule und außerschulischen Unterstützungssystemen

Während der schulische Blick in der Regel leistungsbezogen und kompetenzorientiert ausgerichtet ist, fokussiert die Jugendhilfe stärker auf Lebenslagen, psychosoziale Belastungen und familiäre Ressourcen (Mack, 2014; Olk & Speck, 2012). Gerade in herausfordernden Lagen treffen jedoch Lern- und Entwicklungsfragen häufig mit sozial-emotionalen, gesundheitlichen oder familialen Belastungen zusammen. Damit Unterstützung nicht parallel, widersprüchlich oder lückenhaft verläuft, braucht es Kooperationsformate, in denen unterschiedliche Professionen ihre Perspektiven fallbezogen zusammenführen. Zugleich ist Zusammenarbeit zwischen Schule, Jugendhilfe und ggf. medizinisch-therapeutischen Partner:innen strukturell erschwert, weil die Systeme traditionell mit getrennten Diagnostiken, Berichten und Datensystemen arbeiten. Datengestützte Kooperation steht daher vor einer doppelten Herausforderung: (a) fachlich eine gemeinsame Einschätzung aus heterogenen Datenquellen herzustellen und (b) organisatorisch-rechtlich



Rollen, Zugriffe, Einwilligungen und Kommunikationswege so zu klären, dass Datenschutz, Schweigepflichten und Vertrauen gewahrt bleiben.

Ein exemplarisches Modell hierfür ist das Hamburger Projekt *DreifürEins*⁶³, das zeigt, wie sich Erziehungs- und Bildungspartnerschaften in herausfordernden Lagen gestalten lassen, wenn Schule, Kinder- und Jugendhilfe sowie ggf. medizinisch-therapeutische Partner:innen zusammenarbeiten und Eltern dabei nicht nur informieren, sondern aktiv beteiligen (u. a. durch standardisierte Elternbefragungen sowie im Rahmen dokumentierter Ziel- und Krisenvereinbarungen) (Speck et al., 2023). In diesem Rahmen werden Informationen zur klinischen Symptomatik und Bildungsteilhabe sowie weitere fallbezogene Hintergrundinformationen in multiprofessionellen Fallkonferenzen zu einer gemeinsamen Einschätzung gebündelt. Nach informierter Einwilligung der Eltern und klar geregelter Schweigepflichtentbindung werden in diesen Runden gemeinsame Ziele, Zuständigkeiten und Maßnahmen festgelegt und in Kurzzyklen überprüft bzw. angepasst. Für Eltern kann so eine anschlussfähige Gesamtsicht auf Stärken, Belastungen und Entwicklung ihres Kindes entstehen; zugleich werden Wege, Rollen und Kontaktpunkte im Hilfesystem transparent (Speck et al., 2023).⁶⁴ Angesichts hoher datenschutzrechtlicher Anforderungen – insbesondere in der Kinder- und Jugendhilfe, wo die Verarbeitung und Übermittlung personenbezogener Sozialdaten an enge Voraussetzungen gebunden ist – setzt das Verfahren auf Zweckbindung, Datenminimierung, definierte Zugriffsrechte und transparent erklärte Datenflüsse (Speck et al., 2020; Speck et al., 2023).

Ohne solche verbindlichen Kooperationsvereinbarungen, standardisierten Austauschformate (Fallkonferenzen, Hilfeplanverfahren) und regelmäßigen Reflexionszyklen bleiben Kooperationen zwischen den Systemen häufig oberflächlich (Dittmann et al., 2018).

4.2.3 Gestaltung von Lernentwicklungsgesprächen

Datengestützte Bildungs- und Erziehungspartnerschaften erfordern sowohl strukturelle Rahmenbedingungen als auch professionelle Routinen und Haltungen. Ein entsprechendes, bereits verbreitetes Format stellen Lernentwicklungsgespräche dar (vgl. auch Kapitel 4.1.3). LEG sind in Deutschland vor allem in der Grundschule beforscht. Die Studien zeigen recht konsistent: Nicht das Format an sich, sondern die Qualität der Umsetzung entscheidet über den pädagogischen Mehrwert (Partizipation der Kinder, lernunterstützendes Feedback, klare Zielarbeit und Follow-up) (Dollinger, 2019; Ertl, Kücherer & Hartinger, 2022). Wenn Kinder die Qualität ihrer Gespräche hoch einschätzen (z. B. klare, hilfreiche Rückmeldung, ernst genommene Perspektive, konkrete nächste Schritte), entwickeln sich Merkmale wie Anstrengungsbereitschaft und schulisches Selbstkonzept günstiger (Ertl, Kücherer & Hartinger, 2022).

Eine Analyse von in LEG vereinbarten Zielen prüfte die Qualität anhand der S.M.A.R.T.-Kriterien (Ertl, Weißhorn et al., 2022): Viele Ziele erfüllen mehrere Kriterien; Messbarkeit und konkreter Umfang bleiben aber oft unkonkret. Insgesamt gibt es Hinweise auf die Notwendigkeit besserer Zielarbeit und verbindlicher Nachverfolgung (Ertl, Weißhorn et al., 2022). LEG sind oft lehrkraft gelenkt; echte Aushandlung entsteht

⁶³ <https://innovationsfonds.g-ba.de/projekte/drei-fuer-eins.360> (aufgerufen am 05.03.2026).

⁶⁴ <https://www.tk.de/presse/themen/medizinische-versorgung/dreifueereins-pressegesprach-rebbz-2175650> (aufgerufen am 05.03.2026).



nicht automatisch, sondern erfordert gezielte Gesprächsführung. Partizipation ist ein herstellbares – nicht gegebenes – Qualitätsmerkmal (Bonanati, 2015). Für die Sekundarstufe ist die Evidenz dünner, aber es gibt Hinweise aus Gymnasien. LEG werden von Schüler:innen insgesamt eher positiv erlebt, gleichzeitig besteht Entwicklungsbedarf bei der Formulierung von Lernzielen sowie bei der Ausbalancierung von Beteiligung, Selbststeuerung und lehrkraftseitiger Steuerung (Häbig, 2018).

LEG entfalten ihren Mehrwert insbesondere dann, wenn Selbsteinschätzungen der Schüler:innen im Gespräch systematisch aufgegriffen und mit Leistungs- und Prozessdaten verschränkt werden. Darauf aufbauend sollten wenige, konkret formulierte und überprüfbare Ziele vereinbart werden, einschließlich klarer Terminierung und eines gemeinsamen Verständnisses darüber, woran Fortschritte erkennbar werden. Rückmeldungen sollten dabei so formuliert werden, dass sie nächste Schritte formativ unterstützen. Vereinbarungen werden in LEG typischerweise dokumentiert, zudem erscheint die Weiterarbeit mit den vereinbarten Zielen im nachfolgenden Unterricht für das weitere Lernen relevant (Dollinger, 2019). Ziele sollten in definierten Abständen nachgehalten werden, um Zielarbeit und Verantwortungsbeteiligung über das Gespräch hinaus sicherzustellen. Dabei sind die jeweiligen Möglichkeiten der Eltern für fachbezogene Kommunikation zu berücksichtigen.

4.3 Schlussfolgerungen und Empfehlungen

Schule und Elternhaus tragen gemeinsam Verantwortung für Bildung und Erziehung; Maßstab sind Respekt, Wertschätzung und Mitwirkung (KMK, 2018). Die Forschung belegt den positiven Zusammenhang zwischen qualitativ hochwertigem elterlichen Engagement und der schulischen Entwicklung von Kindern und Jugendlichen.

Datengestützte Erziehungs- und Bildungspartnerschaft benötigt relevante und verständliche Informationen sowie verlässliche Gelegenheiten, diese gemeinsam zu reflektieren, Ziele zu klären und nächste Schritte zu vereinbaren. Lern- und entwicklungsförderliches Feedback sollte zeitnah, spezifisch und knapp sein und nächste Schritte inklusive Reflexionstermin benennen; es knüpft an Vorwissen an, wird anhand von Arbeitsproben und Selbsteinschätzungen ko-interpretiert und orientiert sich an kriterialen und individuellen Bezugsnormen statt an sozialem Vergleich; Daten stützen ein professionelles Urteil und eine kooperative Deutung, sie ersetzen beides nicht.

Im deutschen Kontext sind Lernentwicklungsgespräche teils rechtlich verankert. Zugleich ist die fachlich naheliegende Kooperation mit der Kinder- und Jugendhilfe sowie weiteren externen Partnern (z. B. bei Absentismus) rechtlich anspruchsvoll; ohne klare Vereinbarungen zur Datennutzung bleibt das Potenzial ungenutzt. Internationale Beispiele, etwa Alberta/Kanada, zeigen strukturierte bilaterale und triadische Gespräche (Schule, Eltern, Schüler:innen) mit gemeinsamen Zielen und Maßnahmen sowie surveygestützte Schulentwicklungszyklen. Eltern werden an der Deutung beteiligt, und Schüler:innen übernehmen mit entsprechender Vorbereitung Teile der Gesprächsführung. Transparente Rückmeldungen stärken zudem die aktive Rolle von Schüler:innen. Daten können das Verständnis der eigenen Ausgangslage, Stärken und Entwicklungsbedarfe und die Anpassung von Lernstrategien unterstützen. Daraus können Impulse zur Selbstregulation entstehen.



Um gewünschte Wirkungen zu erreichen, braucht es Professionalisierung und Strukturen. Lehrkräfte benötigen Daten- und Beratungskompetenzen, Schulleitungen sichern Prozesse, Zeitfenster und multiprofessionelle Zusammenarbeit. Die Forschung verweist darauf, dass Elternkooperation anspruchsvoll ist und eine systematische Vorbereitung darauf oft fehlt, und sollte dementsprechend effektive Formate (weiter-)entwickeln und evaluieren (vgl. auch Empfehlung 3). Für Eltern sind barrierearme, mehrsprachige und verständliche Informationsformate zentral, die zeigen, was ihr Kind kann, woran gearbeitet wird und wie zu Hause unterstützt werden kann. Elterliche Selbstwirksamkeit kann durch kurze Workshops, Materialpakete und klare Leitfäden gestärkt werden, insbesondere dann, wenn Angebote wertschätzend und alltagsnah gestaltet sind.

Vor diesem Hintergrund empfiehlt die SWK, Perspektiven von Eltern und Schüler:innen bei der Weiterentwicklung von (digitalen) Instrumenten der Datennutzung (z. B. Dashboards) systematisch zu berücksichtigen (vgl. Empfehlung 2). Die Empfehlungslinien zielen auf drei verbundene Hebel: 1. verständliche, lern- und entwicklungsrelevante Daten sowie praxistaugliche Instrumente bereitstellen; 2. regelmäßige, entwicklungsorientierte Kommunikation verankern; 3. Rahmenbedingungen und Professionalisierung im multiprofessionellen Kontext sichern.

In diesem Sinne werden Daten für pädagogische Zwecke genutzt und professionelle Deutung und Dialog ins Zentrum gestellt. Eine vertrauliche Datennutzung wird sichergestellt, und aus Information wird so gezielte Unterstützung werden. Damit dies gelingt, braucht es Erziehungs- und Bildungspartnerschaften als wertschätzenden Dialog.

Vor diesem Hintergrund empfiehlt die SWK für die Umsetzung datengestützter Erziehungs- und Bildungspartnerschaften zwischen Eltern, Schüler:innen, Lehrkräften und pädagogischen Fachkräften:

Empfehlung 12) Valide und aussagekräftige Daten und Informationen bereitstellen und zur Unterstützung lern- und entwicklungsförderlicher Erziehungs- und Bildungspartnerschaften nutzen; dabei

1. ausgewählte adressatengerechte Daten für Bildungs- und Erziehungspartnerschaften zwischen Lehrkräften/pädagogischem Personal, Eltern/Erziehungsberechtigten und Schüler:innen berücksichtigen; hierzu eignen sich (unter Berücksichtigung der in Kapitel 1 empfohlenen Instrumente):
 - Daten aus den zentralen Lernstandserhebungen,
 - Daten aus landesweit bereitzustellenden lernprozessbegleitenden Diagnoseinstrumenten und Kurztests,
 - Ergebnisse schulischer Klassenarbeiten (entlang von Kompetenzrastern) und anderer Lernprodukte,
 - fachliche und überfachliche Selbsteinschätzungen der Schüler:innen,
 - Daten zu Fehlzeiten von Schüler:innen (z. B. im digitalen Klassenbuch oder Dashboard) mit einheitlicher Frühwarnlogik für alle Schulen einschließlich einer gestuften Elternkommunikation,
2. Datenerhebung und -nutzung ergänzen um niedrigschwellige Elternbefragungen (z. B. zu Partizipation, Schulklima, Unterstützungsbedarfen) und Schülerfeedbacks zu Schule und Unterricht (vgl. Kapitel 1.2.5),



3. barrierefreie, datenschutzkonforme und visuell aufbereitete Auswertungen (z. B. Trendgrafiken, Ampel) vorsehen; Rollen- und Rechtemanagement für alle Nutzergruppen implementieren und analoge Alternativen für Familien ohne stabilen Onlinezugang anbieten.

Empfehlung 13) Regelmäßige, entwicklungsorientierte, vertrauensvolle und datengestützte Informations- und Kommunikationsstrukturen zwischen Schule und Eltern/Erziehungsberechtigten unter Einbindung der Schüler:innen etablieren; dabei

1. datengestützte Zusammenarbeit im Schulentwicklungsplan verankern – unterstützt durch zentral bereitgestellte Gesprächsleitfäden und schlanke Dokumentationshilfen,
2. regelmäßige niedrigschwellige Informationsformate (digital/analog) in Schulen zu Lernzielen und Entwicklungsfortschritten stärken – mit Hinweisen zu Unterstützungsmöglichkeiten der häuslichen Lernbegleitung (z. B. Anregungen zu kleinen Lernroutinen, alltagsintegrierten Gesprächsanlässen, emotionaler Unterstützung; Online-/Take-Home-Materialien) in familienfreundlichen Strukturen (z. B. in Familienzentren), die alle Eltern erreichen,
3. fokussierte, datengestützte Lernentwicklungsgespräche (LEG) mit aktiver Beteiligung der Schüler:innen als gleichwertige Alternative zu Halbjahreszeugnissen etablieren:
 - Schaffung der dafür nötigen Rechtsgrundlagen in der Grundschule,
 - Prüfung der Ausweitung der LEG auf die Sekundarstufe und der Option, in den unteren Jahrgängen der Sekundarstufe Halbjahreszeugnisse durch LEG zu ersetzen,
 - bei Bedarf Integration von Förderplangesprächen in die LEG,
 - in den LEG Formulierung konkreter, realistischer Zielvereinbarungen für alle Kompetenzniveaus,
 - entsprechende Bereitstellung schlanker Dokumentationsformate für die LEG,
 - Durchführung der LEG in einer die jeweiligen Expertisen wertschätzenden Gesprächsatmosphäre.

Empfehlung 14) Rahmenbedingungen für multiprofessionelle Zusammenarbeit und zur Professionalisierung für datengestützte Erziehungs- und Bildungspartnerschaften (weiter-) entwickeln; dabei

1. multiprofessionelle Zusammenarbeit zur gemeinsam verantworteten Diagnostik verankern (Kooperationszeiten im Stundenplan, Ziel-/Zuständigkeitsdefinition, Fallprotokolle, Review-Zyklen) anhand landesseitig bereitgestellter Templates zur Lernentwicklungs- und Förderplanung (SWK, 2022a),
2. Zusammenarbeit von Schule mit außerschulischen Partnern rechtlich absichern (z. B. Jugendhilfe, Gesundheitsdienste) auf Basis abgestimmter datenethischer und datenschutzkonformer Standards (Einwilligung/Schweigepflichtentbindung, Zweckbindung, Datensparsamkeit, rollenbasierte Zugriffsrechte, Löschrufen),
3. Lehrkräfte sowie Schulleitungen für die datengestützte Gestaltung von Erziehungs- und Bildungspartnerschaften phasenübergreifend professionalisieren, insbesondere in Bezug auf lernförderliches Feedback und kompetente dialogische Gesprächsführung; einschließlich der Prozesssteuerung durch die Schulleitung (vgl. Kapitelübergreifende Empfehlungslinien).



5. Daten für Systemmonitoring und -steuerung durch die Politik

Bis zum Beginn der 1990er-Jahre fehlten in Deutschland systematische und regelmäßige Beobachtungen bzw. Überprüfungen von Erträgen institutionalisierter Bildungsprozesse. In anderen Industrienationen wie den USA, Großbritannien, Kanada und den Niederlanden waren entsprechende Programme längst etabliert. In den USA beispielsweise führte im Auftrag der Bundesregierung seit 1969 zunächst der Educational Testing Service und später das National Center for Education Statistics (NCES) das *National Assessment of Educational Progress (NAEP)* durch, wodurch sich Trends in den Leistungen verschiedener Fächer nachzeichnen ließen. In der Bundesrepublik Deutschland spielten zu diesem Zeitpunkt Untersuchungen, die auf einer breiten empirischen Basis die Beschreibung und Analyse der Erträge fachlichen Lernens in den Mittelpunkt rückten, keine Rolle bei der bildungspolitischen Steuerung bzw. wurden überhaupt nicht in Betracht gezogen. Dies galt auch noch für die Teilnahme des vereinigten Deutschlands an *TIMSS 1995 (Third International Mathematics and Science Study)*. Umso bemerkenswerter waren die besondere Aufmerksamkeit und die breite Diskussion, die die TIMSS-Ergebnisse für die Mittelstufe (Baumert et al., 1997) auslösten. Der *TIMSS-Schock* machte sich an den für die Öffentlichkeit unerwartet niedrigen Leistungsergebnissen bemerkbar (Klieme & Baumert, 2001; Schwippert, 2005). In der Folge wurde diese Dynamik durch die Veröffentlichung der ersten PISA-Studie (Baumert et al., 2001; OECD, 2001) in erheblichem Maße intensiviert: Der *PISA-Schock* bestätigte nicht nur die zuvor durch TIMSS aufgezeigten Defizite, sondern verschärfte sie in der öffentlichen und wissenschaftlichen Wahrnehmung, sodass vielfach von einem entscheidenden Impuls für einen paradigmatischen Wandel in der deutschen Bildungsforschung und Bildungspolitik gesprochen wird (Baumert et al., 2001). PISA erweiterte das Untersuchungsspektrum um Lesen als Schlüsselkompetenz für Bildung, untersuchte die Fähigkeit, Wissen auch außerhalb des Schul(buch)kontexts anzuwenden, erhöhte die internationale Anschlussfähigkeit durch das OECD-Bezugssystem und rückte (systematischer als TIMSS) Disparitäten in den Blickpunkt. Als Folge wurde in Deutschland die empirische Wende in der Bildungspolitik eingeleitet, begleitet von Maßnahmen zur Qualitätssicherung im Bildungswesen auf unterschiedlichen Ebenen (Lange, 2008). Outputmaße (Schulleistungen) der Schüler:innen sollten Informationen über die Leistungsfähigkeit des Bildungssystems auf der Individual-, Klassen-, Schul-, Länder- und Bundesebene liefern und Ausgangspunkte für Reformen schaffen.

Dieses Kapitel fokussiert sich auf die Länder- und Bundesebene. Diese Perspektive berücksichtigt vor allem die Beiträge nationaler und internationaler Large-Scale Assessments (LSA) sowie der Bildungsberichterstattung zum Bildungsmonitoring auf Systemebene in Deutschland. Ziel dieses Kapitels ist es, Empfehlungen für die zukünftige Gestaltung des LSA-gestützten Systemmonitorings in Deutschland zu erarbeiten. Dabei ist zu prüfen, ob sich beispielsweise die LSA dahin gehend bewährt haben, dass sie für Bildungsreformen, etwa bundesweit, in den 16 Ländern oder auch auf regionaler Ebene genutzt wurden, und wie diese Nutzung zu bewerten ist. Auch ist zu erörtern, wie die bisherigen Ansätze des Systemmonitorings weiterentwickelt werden sollten, um die Leistungsfähigkeit des Bildungssystems in Deutschland zu steigern.



5.1 Rückblick und Situationsanalyse

Grundlagen für die folgenden Ausführungen waren zum einen die KMK-Beschlüsse zur Gesamtstrategie zum Bildungsmonitoring (KMK, 2006, 2015) sowie weitere Dokumente von Bund und Ländern zum Bildungsmonitoring in Deutschland. Zum anderen fanden Hearings mit den leitenden Wissenschaftler:innen der unterschiedlichen LSA sowie mit ausgewählten Ländervertreter:innen statt. Die Wissenschaftler:innen wurden gebeten, die Bedeutung ihrer jeweiligen Schulleistungsstudie für das Bildungsmonitoring zu beschreiben und in den Kontext der bestehenden Gesamtstrategie einzubetten. Ländervertreter:innen als Adressat:innen der LSA und des Nationalen Bildungsberichts wurden hinsichtlich der Rezeption der Studien und der Konsequenzen für die Steuerung des Bildungssystems befragt.

5.1.1 Die Gesamtstrategie der KMK zum Bildungsmonitoring

Ausgangspunkt der Gesamtstrategie der KMK zum Bildungsmonitoring waren die sogenannten *Konstanzer Beschlüsse* (KMK, 1997), die nicht zuletzt aus dem enttäuschenden Abschneiden der Schüler:innen in Deutschland in TIMSS (Baumert et al., 1997) resultierten (vgl. Abbildung 11). Darin wurde vereinbart, dass zukünftig regelmäßig nationale und internationale Vergleichsuntersuchungen zum Leistungsstand von Schüler:innen ausgewählter Jahrgänge und Fächer (Bildungssprache, Mathematik, Naturwissenschaften, Fremdsprachen) stattfinden. Der Schwerpunkt lag zunächst auf der Sekundarstufe I. Die Konstanzer Beschlüsse waren von der Überzeugung getragen, dass Outputindikatoren des Bildungssystems in Form von Schülerleistungen wichtige Steuerungsinformationen liefern, um die Leistungsfähigkeit des Bildungssystems zu steigern (sogenannte Outputorientierung).

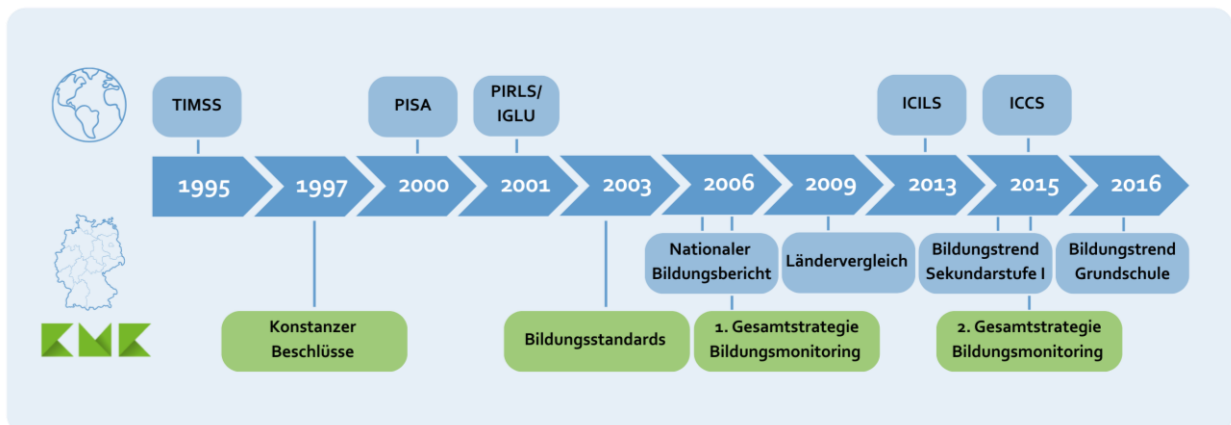


Abbildung 11: Zentrale Beschlüsse der KMK zum Bildungsmonitoring und erste Durchführung der Studien

Quelle: eigene Recherchen

Zusätzlich zu den fachbezogenen Schulleistungen sollten – so die Konstanzer Beschlüsse der KMK – auch Schlüsselqualifikationen wie Kooperations- und Kommunikationsfähigkeiten sowie Problemlösekompetenzen überprüft werden. Eine der ersten Konsequenzen aus den Beschlüssen war die Teilnahme an PISA im Jahr 2000 (Baumert et al., 2001), um die Leistungsfähigkeit des bundesdeutschen Bildungssystems gegen Ende der Sekundarstufe I im internationalen Vergleich feststellen zu können. Ergänzend zum



internationalen Vergleich wurde beschlossen, in allen 16 Ländern hinreichend große Stichproben von 15-Jährigen zu ziehen, um nationale Leistungsvergleiche zwischen den Ländern durchführen zu können, und ergänzende Tests mit Aufgaben einzusetzen, die curriculare Validität beanspruchten (Baumert et al., 2002; Prenzel, Baumert et al., 2005; Prenzel et al., 2008). Man wollte so möglichen Ergebnisverzerrungen durch die vermeintlich nicht curricular validen internationalen Items nachgehen. Mit der Fokussierung auf 15-Jährige ist PISA nach wie vor die einzige große internationale Schulleistungsstudie, die eine alters- anstelle einer jahrgangsbezogenen Stichprobe untersucht. In Deutschland wird das Programm traditionell durch eine jahrgangsbezogene Stichprobe (Schüler:innen der neunten Jahrgangsstufe) ergänzt. Somit können die Outputdaten der Schüler:innen einer Klasse mit Merkmalen der sie unterrichtenden Lehrkräfte direkt verknüpft und wichtige Schlüsse über den Zusammenhang zwischen Lehrkräfte- und Unterrichtsmerkmalen sowie Schülerleistungen gezogen werden.

Entgegen der ursprünglichen Absicht, sich auf Studien in der Sekundarstufe I zu beschränken, beteiligte sich Deutschland bereits im Jahr 2001 an der Internationalen Grundschul-Lese-Untersuchung (IGLU; international: *PIRLS – Progress in Reading Literacy Study*) in der vierten Jahrgangsstufe (Bos et al., 2003). Nachdem Deutschland in den Jahren 1994/95 an TIMSS in der siebten/achten Jahrgangsstufe und zum Ende der Sekundarstufe II teilgenommen hatte (Baumert et al., 1997), war zunächst keine weitere Teilnahme an TIMSS vorgesehen. Im Rahmen von PIRLS 2001 wurde in Deutschland allerdings eine Erweiterung vorgenommen, sodass auch Mathematikaufgaben für die vierte Jahrgangsstufe, die in TIMSS 1995 eingesetzt worden waren, bearbeitet wurden. Seit 2007 nimmt Deutschland wieder an der TIMSS-Grundschuluntersuchung (vierte Jahrgangsstufe) teil.

In den Konstanzer Beschlüssen war bereits die Erarbeitung von Bildungsstandards (zunächst) für den Mittleren Schulabschluss festgelegt. Dies folgte der damaligen Beobachtung, dass Länder, die in den internationalen Schulleistungsvergleichen besonders gut abschnitten, verbindlich geltende Standards entwickelt hatten und dies vielfach mit nationalen Assessments verknüpften. In den Bildungsstandards für Deutschland sollten grundlegende Kompetenzen in den Fächern Deutsch, Mathematik, modernen Fremdsprachen und Naturwissenschaften festgelegt und in späteren Ländervergleichen getestet werden. Der Auftrag wurde dann auf die Primarstufe und den Ersten Schulabschluss (Hauptschulabschluss) ausgedehnt. Mit der daraus abgeleiteten Erarbeitung und Verabschiedung der Bildungsstandards für den Primarbereich (Ende der vierten Jahrgangsstufe; Fächer: Mathematik und Deutsch) und für das Ende der Sekundarstufe I (Hauptschulabschluss/Erster Schulabschluss; Fächer: Deutsch, Mathematik, erste Fremdsprache; Mittlerer Schulabschluss; Fächer: Deutsch, Mathematik, erste Fremdsprache, Biologie, Chemie, Physik) in den Jahren 2003 und 2004 wuchs der Druck, eine Strategie zu entwickeln, wie zum einen das Erreichen der in den Standards beschriebenen Kompetenzziele in den 16 Ländern überprüft und zum anderen solch ein nationales Assessment-Programm mit den internationalen Schulleistungstudien und mit einer Bildungsberichterstattung für Deutschland verknüpft werden kann. Im Jahr 2004 wurde daher das Deutsche PISA-Konsortium beauftragt, Mathematikaufgaben zur Überprüfung der Bildungsstandards am Ende der Sekundarstufe I zu entwickeln. Deren Normierung erfolgte im Rahmen von PISA 2006 (Prenzel et al., 2007). Im Jahr 2004 gründeten die Länder das IQB als unabhängige wissenschaftliche Einrichtung an der Humboldt-Universität zu Berlin und entwickelten ihre erste Gesamtstrategie zum Bildungsmonitoring, die 2006 verabschiedet wurde (KMK, 2006). Diese sah konkret vor:



1. Feststellung der Leistungsfähigkeit des deutschen Bildungswesens im internationalen Vergleich durch Teilnahme an internationalen Schulleistungsstudien, vorrangig an solchen, in denen Kompetenzbereiche getestet werden, für die in Deutschland Bildungsstandards vorliegen; konkret waren das PIRLS/IGLU, TIMSS und PISA; um Redundanzen zwischen TIMSS und PISA zu vermeiden, wurde bei TIMSS die Beschränkung auf den Grundschulbereich (vierte Klasse) vorgenommen,⁶⁵
2. zentrale Überprüfung des Erreichens der Bildungsstandards durch das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) im Vergleich der 16 Länder auf der Basis länderrepräsentativer Schülerstichproben am Ende der vierten und neunten Jahrgangsstufe,
3. länderspezifische und länderübergreifende Vergleichsarbeiten – in Anknüpfung an oder Anlehnung an die Bildungsstandards –, die flächendeckend durchgeführt werden und für die Unterrichts- und Schulentwicklung genutzt werden sollen (VERA 3 und VERA 8, vgl. Kapitel 1),⁶⁶
4. gemeinsame Bildungsberichterstattung von Bund und Ländern mit dem Ziel, für sämtliche Bildungsetappen und die verschiedenen Steuerungsebenen (Input, Prozess, Output und Kontext) Indikatoren zu entwickeln und diese regelmäßig im Trend zu berichten, ergänzt um ein für jeden Bericht spezifisches Schwerpunktthema.

Die Gesamtstrategie der KMK schuf die Grundlage für eine neue Steuerungslogik, die auf Ergebnisorientierung, verbindliche Rechenschaftslegung und den systematischen Ausbau eines umfassenden Bildungsmonitorings in Partnerschaft mit der empirischen Bildungsforschung setzte. Die zeitliche Taktung der Erhebungen, die in der Gesamtstrategie verabschiedet wurden, ist in Abbildung 12 aufgeführt. Die standardbasierten Ländervergleiche, die inzwischen als IQB-Bildungstrends bezeichnet werden, begannen 2009 mit der ersten Erhebung in der neunten Jahrgangsstufe. TIMSS in der vierten Jahrgangsstufe fand in Deutschland erstmals 2007 statt. Erkennbar sind die ursprüngliche Parallelität zwischen PIRLS und dem IQB-Bildungstrend in der vierten Klasse sowie die Parallelität zwischen PISA und dem IQB-Bildungstrend in der neunten Klasse. Insbesondere durch coronabedingte Verschiebungen in mehreren Studien und den Wechsel der OECD auf einen vierjährigen Erhebungszyklus ab 2029 wurde diese Parallelität aufgegeben.

⁶⁵ Interessanterweise gilt für immerhin 20 OECD-Staaten, dass sie inzwischen wieder zu der Praxis zurückgekehrt sind, sich an TIMSS in der achten Jahrgangsstufe und an PISA zu beteiligen, nicht zuletzt deshalb, weil TIMSS durch ihren vierjährigen Zyklus dieselbe Kohorte in den Jahrgangsstufen 4 und 8 untersucht.

⁶⁶ Einige Länder haben die zentralen Lernstandserhebungen um eigene Lernausgangslagenuntersuchungen in anderen Jahrgangsstufen erweitert.



Jahr	2000	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
PIRLS		•					•					•					•					•					•				
TIMSS								•				•				•				•				•				•			
PISA	•			•			•			•		•				•			•			•			•				•		
IQB-BT; G																•						•							•		
IQB-BT; S																•			•			•		•					•		
BiBer						•		•		•		•		•		•		•		•		•		•		•		•		•	
ICLS														•					•				•					•			
ICCS										•						•						•						•			
VERA 3/8								•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•

Abbildung 12: Zeitliche Taktung der Maßnahmen zur Feststellung der Leistungsfähigkeit des Bildungswesens

Anmerkungen: PIRLS: Progress in Reading Literacy Study (Lesen in der vierten Jahrgangsstufe); TIMSS: Trends in Mathematics and Science Study (Mathematik und Naturwissenschaften in der vierten Jahrgangsstufe); PISA: Programme for International Student Assessment (Lesen, Mathematik, Naturwissenschaften plus innovative Domäne bei 15-Jährigen); IQB-BT; G: Bildungstrend in der vierten Jahrgangsstufe (Deutsch, Mathematik); IQB-BT; S: Ländervergleich in der neunten Jahrgangsstufe (abwechselnd Sprachen vs. Mathematik und Naturwissenschaften); BiBer: Nationaler Bildungsbericht; ICILS: International Computer and Information Literacy Study (digitale Grundbildung in der achten Jahrgangsstufe); ICCS: International Civic and Citizenship Education Study (politische Bildung in der achten Jahrgangsstufe). Farblich abgesetzt: Programme, die nicht Teil der Gesamtstrategie sind. Quelle: eigene Recherchen

Die Gesamtstrategie wurde in den Folgejahren weiterentwickelt und 2015 in überarbeiteter Fassung vorgestellt (KMK, 2015). Das (blau hinterlegte) Erhebungsprogramm in Abbildung 12 wurde beibehalten. Ergänzungen in der Neuauflage resultierten zum einen aus der zwischenzeitlichen Verabschiedung der länderübergreifenden Bildungsstandards für die Allgemeine Hochschulreife, zum anderen aus dem Wunsch der Länder, mehr evidenzbasiertes Erklärungs- und Handlungswissen an die Hand zu bekommen, mit dem sich Bildungs- und Erziehungsprozesse in Schulen verbessern lassen. Die dazu notwendigen Literaturreviews und Metaanalysen sollten nicht dem Bildungsmonitoring dienen, sondern Veränderungswissen liefern.

Infolge der Einführung von Bildungsstandards für die Allgemeine Hochschulreife (Fächer: Deutsch, Mathematik, fortgeführte Fremdsprache, Biologie, Chemie, Physik) wurde das IQB damit betraut, die Entwicklung ländergemeinsamer Pools von Aufgaben für die schriftliche Abiturprüfung zu koordinieren, die geeignet sind, standardbasiert Kompetenzen am Ende der gymnasialen Oberstufe zu überprüfen, und damit wichtig für die Implementation der Standards in die unterrichtliche Praxis sind. Hieraus entnehmen die Länder Aufgaben und setzen sie in den zentralen Abiturklausuren ein. Die Entwicklung der ländergemeinsamen Abiturprüfungsaufgaben ist aufwendig, reduziert jedoch den Aufwand, der mit der Entwicklung eigener Aufgaben in den einzelnen Ländern verbunden ist. Die Länder haben sich darauf verständigt, in jedem der Fächer, für die ländergemeinsame Pools zur Verfügung stehen, mindestens 50 Prozent ihrer Abiturprüfungsaufgaben zu entnehmen und die Aufgaben unverändert einzusetzen, wobei der Aufgabenpool zudem so weiterentwickelt werden soll, dass auch eine Entnahme von 100 Prozent der Aufgaben möglich ist (KMK, 2020b, S. 20). Ein Verfahren, mit dem die Erreichung der in den Standards beschriebenen Ziele mithilfe standardisierter Tests ähnlich zu den IQB-Bildungstrends überprüft wird, wurde von den Ländern hingegen nicht eingeführt.



Trotz dieser Beschränkung gilt, dass die Länder mit der Gesamtstrategie folgende Ziele im Auge haben (KMK, 2015, S. 5):

- „für Transparenz hinsichtlich zentraler Maßstäbe zur Bestimmung von Bildungsqualität zu sorgen,
- wissenschaftlich abgesicherte Ergebnisse zur Bildungsqualität für Schulen, Bildungsverwaltung, Bildungspolitik und Öffentlichkeit bereitzustellen,
- die Qualitäts- und Standardsicherung auf allen Ebenen möglichst eng mit der Qualitätsentwicklung zu verbinden,
- eine Grundlage für die länderübergreifende Zusammenarbeit im Bereich der Qualitätssicherung zu schaffen“.

LSA leisten im Rahmen der bestehenden Gesamtstrategie der KMK einen zentralen Beitrag zur Gewinnung von Beschreibungswissen und führen eine Systemdiagnose zu Input, Prozessen, Output und Kontextfaktoren im Bildungssystem durch. Darüber hinaus stellt sich die Frage, wie weiteres Wissen, insbesondere steuerungsrelevantes und schulpraktisches Erklärungs- sowie Handlungswissen, systematisch über mit LSA abgestimmte Forschungsstrategien und -programme erschlossen werden kann. Hier ist eine Reihe von Aktivitäten zu nennen, zunächst die ab 2016 vermehrt am Zentrum für Internationale Vergleichsstudien (ZIB) erarbeiteten Literaturreviews und Metaanalysen, die den internationalen Forschungsstand zu ausgewählten Themen systematisch aufbereiteten (Hillmayr et al., 2017; Hillmayr et al., 2021). Ergänzend wurden gezielte Forschungsprogramme initiiert, etwa Schwerpunktprogramme der Deutschen Forschungsgemeinschaft (DFG), Rahmenprogramme des Bundesministeriums für Bildung und Forschung (BMBF) oder das Nationale Bildungspanel (NEPS), die mit klaren Forschungszielen neben deskriptivem Wissen zur Gewinnung von Erklärungs- und Handlungswissen für Bildungspolitik und Bildungspraxis beitragen. Ein Beispiel ist das BMBF-Programm *Schule macht stark* (2021 bis 2025), in dem empirisch fundierte Maßnahmen zur Schul- und Unterrichtsentwicklung in sozial herausfordernden Lagen entwickelt und erprobt wurden. Ein anderes Beispiel ist das Bund-Länder-Programm *Leistung macht Schule* (LemaS 2018 bis 2022; LemaS-Transfer 2023 bis 2027), in dem Programme zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schüler:innen wissenschaftlich fundiert entwickelt, erprobt und implementiert wurden.

Auch das Forschungsdatenzentrum (FDZ) am IQB trägt als Teil des ZIB durch die kontinuierliche Bereitstellung von Datensätzen aus nationalen und internationalen LSA sowie durch die daraus bundesweit entstehenden Sekundäranalysen zu vertieftem Wissen für evidenzbasierte Steuerung und schulische Entwicklungsprozesse bei. Besonders eng mit den internationalen Studien verzahnt ist auch die nationale Begleitforschung, wie sie regelmäßig in PISA und anderen internationalen LSA durch gezielte nationale Ergänzungen erfolgt. So wurde nach den insgesamt schwachen Ergebnissen in Mathematik in PISA 2022 der Unterricht in ausgewählten deutschen Schulen detailliert untersucht, um Gelingensbedingungen und Entwicklungsbedarfe zu identifizieren, in enger Anlehnung an die mathematische Rahmenkonzeption von PISA sowie die Kompetenzdimensionen der IQB-Bildungsstandards. Der Transfer der im Rahmen von LSA gewonnenen Erkenntnisse in steuerungsrelevantes Handlungswissen und schulpädagogisches Praxiswissen bleibt eine Herausforderung (vgl. hierzu auch Renkl, 2022), wird jedoch zunehmend bei der Planung und Durchführung von Studien und begleitenden Programmen mitgedacht (vgl. hierzu Bromme & Prenzel, 2014).



5.1.2 Weitere internationale Studien jenseits der Gesamtstrategie

Parallel zu den in der KMK-Gesamtstrategie aufgeführten internationalen Studien (vgl. Abbildung 12) beteiligt sich Deutschland seit 2013 an der *International Computer and Information Literacy Study (ICILS)* und seit 2022 an der *International Civic and Citizenship Education Study (ICCS)*. Beide Studien adressieren Schüler:innen der achten Jahrgangsstufe, ICILS zielt auf computer- und informationsbezogene Kompetenzen sowie Computational Thinking, ICCS auf die politische Grundbildung. Während in ICILS seit Beginn versucht wird, eine bundesweit repräsentative Stichprobe zu rekrutieren und zu testen, war ICCS bislang auf zwei Länder beschränkt (Nordrhein-Westfalen und Schleswig-Holstein). Erst im Jahr 2027 wird sich Deutschland auf Beschluss von Bund und Ländern mit einer bundesweiten Stichprobe von Achtklässler:innen an ICCS beteiligen. Bund und Länder haben sich auch schon darauf verständigt, dass ICILS im dritten Zyklus 2028 in Deutschland stattfinden wird. Eine Einbindung in eine Gesamtstrategie des Bildungsmonitorings fehlt bislang für ICILS. Die Teilnahme der Schulen und der Schüler:innen an den ICCS- und ICILS-Tests ist nicht verpflichtend, was die Repräsentativität der Befunde einschränken kann.

Neben den großen Schulleistungsstudien existieren weitere internationale Vorhaben, an denen Deutschland derzeit nicht oder nur punktuell teilnimmt. Dazu zählt seit 2008 der *OECD Teaching and Learning International Survey (TALIS)*⁶⁷, der alle sechs Jahre durchgeführt wird, zuletzt 2024, mit der nächsten Erhebungsrunde entsprechend 2030. TALIS befragt Lehrkräfte und Schulleitungen zu Arbeitsbedingungen, professioneller Entwicklung, Kooperation und Lernumwelten; neuerdings werden auch Erzieher:innen und Kita-Leitungen einbezogen. Ermittelt werden (über Selbsteinschätzungen) Belastungsquellen und Ressourcen, wobei initiale Ansätze zur Verknüpfung der Angaben mit PISA-Daten bislang nicht verfolgt wurden. Die KMK hat sich in der Vergangenheit gegen eine Teilnahme an TALIS entschieden, nicht zuletzt mit dem Argument, dass Befragungen von Lehrkräften auch im Rahmen der IQB-Bildungstrends sowie der internationalen LSA durchgeführt werden könnten. Deutschland beteiligt sich daher aktuell nicht an der TALIS-Kernstudie und hat bislang nur vereinzelt (z. B. *TALIS-Videostudie 2018*, *TALIS Starting Strong* im Bereich der frühkindlichen Bildung) teilgenommen.

Der *OECD Survey on Social and Emotional Skills (SSES)*⁶⁸ erhebt bei 10- und 15-Jährigen sozial-emotionale Merkmale in Bereichen wie Aufgeschlossenheit, Kooperationsbereitschaft und soziale Einbindung vorrangig auf Basis von Selbstberichten. Diese sollen künftig durch innovative Messansätze, die eine verlässlichere Einordnung ermöglichen, ergänzt werden. Ziel ist eine international vergleichbare Datengrundlage zu förderlichen und hinderlichen Kontextfaktoren für Wohlbefinden und gesellschaftliche Teilhabe.

Das *International Vocational Education and Training Assessment (PISA-VET)*⁶⁹ der OECD erfasst arbeitsplatzrelevante Kompetenzen, die am Ende des ersten beruflichen Abschlusses erworben werden. Im Mittelpunkt stehen derzeit sechs Berufsfelder, darunter Kfz-Mechatronik, kaufmännische und pflegerische Berufe. Diese decken nicht nur zentrale Berufsfelder, sondern auch quantitativ bedeutsame Ausbildungs-

⁶⁷ <https://www.oecd.org/en/about/programmes/talis.html> (aufgerufen am 20.02.2026).

⁶⁸ <https://www.oecd.org/en/topics/social-and-emotional-skills.html> (aufgerufen am 20.02.2026).

⁶⁹ <https://www.oecd.org/en/about/projects/international-vocational-education-and-training-assessment-pisa-vet.html> (aufgerufen am 20.02.2026).



berufe ab. Deutschland ist an der Entwicklungs- und Pilotierungsphase sowohl mit wissenschaftlicher Expertise als auch in der Datenerhebung und Instrumentenentwicklung, die zentral auf die deutschen *ASCOT*- und *ASCOT+*-Initiativen⁷⁰ aufsetzen, beteiligt; über eine Teilnahme an der Hauptstudie, die frühestens 2029 stattfinden wird, ist aktuell noch nicht entschieden.

5.1.3 Welche Informationen liefern die Large-Scale Assessments?

Steuerungsmodelle für das allgemeinbildende Schulsystem folgen einer Kontext-Input-Prozess-Output-Logik (Scheerens, 2008). Während in einem Input-orientierten Ansatz Steuerung vor allem über die Inputs (z. B. Festlegung von Stundentafeln, Zuweisung von Ressourcen und Lehrkräften) erfolgt, wird einem Output-orientierten Ansatz zufolge – der sich in Deutschland infolge von TIMSS 1994/95 und PISA 2000 durchgesetzt hat – verstärkt über die Definition klarer Kompetenzerwartungen sowie deren Überprüfung gesteuert.

Die LSA haben die Aufgabe übernommen, kognitive und in jüngerer Zeit vermehrt auch nicht-kognitive Erträge in Form fachlicher oder überfachlicher Kompetenzen aufseiten der Schüler:innen (Output) zu messen. Jene dienen als Indikatoren der Leistungsfähigkeit von Bildungssystemen. Darüber hinaus beschreiben LSA Leistungsdisparitäten (nach Geschlecht, sozialer Herkunft, Migrationsgeschichte), die ebenfalls als Indikatoren der Leistungsfähigkeit von Bildungssystemen gelten. Der IQB-Bildungstrend soll den 16 Ländern in erster Linie Informationen darüber liefern, welche fachlichen Kompetenzstände Schüler:innen am Ende der vierten bzw. am Ende der neunten Jahrgangsstufe erreichen, und dabei vor allem, wie hoch die Anteile der Schüler:innen sind, die die Regel-, Mindest- und Optimalstandards verfehlen. Die internationalen Studien sollen auf der Basis national repräsentativer Stichproben über Schülerleistungen im internationalen Vergleich berichten. Ohne die dezidierte Erweiterung der Stichproben (sogenanntes *Oversampling*; siehe PISA-E 2000) lassen sich aus den internationalen Studien keine Aussagen über die einzelnen Länder in Deutschland treffen. Eine über die Landesebene hinausgehende Aussagekraft, wie sie etwa auch auf Ebene der Landkreise denkbar wäre (vgl. hierzu den Ansatz bei PISA-E 2003 in Prenzel, Baumert et al., 2005), ist aktuell anhand keiner der durchgeführten nationalen und internationalen Studien möglich.

Um politische Maßnahmen zur Weiterentwicklung des Systems ableiten zu können, benötigen Steuerungsakteur:innen plausible Annahmen über Ansatzpunkte zur Verbesserung von Outputs. Dazu ist über die Outputdaten hinausgehend nicht nur eine Erfassung relevanter Kontext- und Inputdaten erforderlich, sondern auch eine valide Erfassung von Daten zu Prozessen auf der Schul- und der Unterrichtsebene. LSA leisten die Erfassung von Kontext-, Input-, Prozess- und Outputdaten in unterschiedlicher Weise. Wiederkehrende Inhalte der Berichte zu den Studien sind z. B.:

- Kontext:
 - soziale und kulturelle Zusammensetzung der Schülerschaft auf Schul- und Länderebene,
 - Anteile der Schüler:innen mit sonderpädagogischem Förderbedarf,
 - Schulgröße,
 - Größe der Gemeinden, in denen sich die Schulen befinden,
 - außerschulische Kooperationspartner,

⁷⁰ https://www.ascot-vet.net/ascot/de/home/home_node.html (aufgerufen am 20.02.2026).



- Input:
 - Schulstrukturen in den teilnehmenden Ländern,
 - curriculare Passung der Testaufgaben,
 - Studentafeln (Lerngelegenheiten) und Lernzeiten/Lerngelegenheiten,
 - Situation der Beschulung von Schüler:innen mit sonderpädagogischem Förderbedarf,
 - Ausstattung (u.a. digitale Infrastruktur) und deren Nutzung,
- Prozess:
 - Wahrnehmung der Unterrichtsqualität (Zielführung, Kohärenz, Unterrichtsgespräch, Aufgabenqualität, Passung auf Lernvoraussetzungen) durch die Schüler:innen,
 - didaktischer Gehalt von Unterrichts- und Prüfungsaufgaben (z. B. einfache Routinen vs. komplexe Modellierungsaufgaben im Fach Mathematik),
 - professionelle Kompetenzen der Lehrkräfte,
 - Schulmanagement und Führung,
 - Fortbildungen der Kollegien an den teilnehmenden Schulen,
 - Förderangebote für leistungsschwache Schüler:innen,
- Output:
 - fachliche Kompetenzen,
 - überfachliche kognitive Kompetenzen (Problemlösefähigkeiten, kritisches Denken, Kreativität; Learning in a Digital World),
 - prozentuale Anteile der Risikogruppen,
 - soziale, migrationsbezogene und geschlechtsbezogene Disparitäten in den fachlichen Kompetenzen,
 - Einstellungen von Schüler:innen und Lehrkräften,
 - fachspezifische motivationale Maße (Selbstkonzepte, Einstellungen, Interessen),
 - überfachliche motivationale Maße und Persönlichkeitsmerkmale (schulisches und psychisches Wohlbefinden, Selbstwirksamkeitserleben, Lebenszufriedenheit, Growth-Mindset).

Durch die wiederkehrenden Erhebungen lassen sich auf allen vier Ebenen (Kontext, Input, Prozess, Output) Trends beschreiben und Implikationen ableiten, an welchen Stellen Veränderungen nötig sind, um negative Trends zu stoppen bzw. positive Trends fortzusetzen. Ein besonderes Potenzial der LSA liegt auch darin, dass Kontext-, Input-, Prozess- und Outputdaten miteinander in Beziehung gebracht werden können. In den Berichten zu den LSA werden immer wieder regressionsanalytische Auswertungen vorgestellt, die solche Zusammenhänge aufdecken. Zwar lassen diese Analysen keine kausalen Schlussfolgerungen zu. Sie unterstützen jedoch die Entwicklung plausibler Annahmen, die in weiterführenden Studien getestet werden können. Die im Rahmen der LSA durchgeführten Trend- und Regressionsanalysen unterstützen die Identifikation konkreter Ansatzpunkte für länderspezifische und länderübergreifende systemische Programme und Maßnahmen der Qualitätsentwicklung wie z. B. BISS oder QuaMath (vgl. Kapitel 5.1.10).

Auf Ebene des Outputs ergänzen überfachliche Kompetenzen wie Kreativität oder digitale Kompetenzen (vgl. Kapitel 1) die Ausgestaltung internationaler LSA. Insbesondere PISA greift hier regelmäßig wichtige innovative Strömungen auf. In PISA 2012 stand etwa Problemlösen im Fokus (Prenzel et al., 2013), in PISA 2022 Kreativität (Lewalter et al., 2023), und in PISA 2029 wird ein erweitertes Konzept digitaler



Kompetenz, einschließlich Medien- und KI-Kompetenzen, integriert. Auch Studien wie ICCS und ICILS konzentrieren sich in ihrem gesellschaftlich orientierten Bezug weniger auf fachgebundene Basiskompetenzen als auf überfachliche Kompetenzbereiche. Die erfassten fachübergreifenden Kompetenzbereiche liefern wichtige Hinweise für die Weiterentwicklung des Bildungssystems, sind jedoch bislang nicht vollständig in eine Gesamtstrategie integriert.

Außerdem werden auf der Outputseite zunehmend sozial-emotionale Merkmale erfasst, die für eine schülerzentrierte Unterrichtskultur von hoher Bedeutung sind, in der Vergangenheit jedoch nur begrenzt berücksichtigt wurden. Jüngere Studien, wie SSES, PISA oder die IQB-Bildungstrends, erfassen sozial-emotionale Merkmale inzwischen systematischer. Dies ist methodisch anspruchsvoll: Häufig beruhen die Messungen auf Selbstberichten, die wertvolle Einblicke in Selbstwahrnehmungen und Haltungen geben, können jedoch kontext- und kulturabhängig sein. Ergänzend könnten innovative, performanzbasierte Verfahren eingesetzt werden, deren Aussagekraft für das Bildungsmonitoring derzeit jedoch noch nicht abschließend beurteilt werden kann.

Indikatoren aus internationalen LSA werden ebenfalls genutzt, um supranationale Bildungsziele zu erfassen. Hierzu finden vorrangig Indikatoren aus PISA, aber auch aus anderen LSA wie ICILS, ICCS, TIMSS und PIRLS Anwendung. Als besonders relevantes Beispiel ist das vierte Nachhaltigkeitsziel der Vereinten Nationen (*Sustainable Development Goal; SDG 4*), das bis 2030 eine inklusive, chancengerechte und hochwertige Bildung für alle anstrebt, zu nennen. SDG 4 berücksichtigt beispielsweise für das Teilziel 4.1.1 *Proportion of children and young people achieving at least a minimum proficiency level in reading and mathematics* als Datengrundlage Indikatoren aus PISA, genauer den Anteil der 15-Jährigen, die mindestens Kompetenzstufe 2 erreichen. Auch der *Human Capital Index (HCI)* der Weltbank oder der *EU-Bildungsmonitor* nutzen entweder einzelne Indikatoren oder aggregierte Indikatoren mehrerer internationaler LSA, hier auch PISA, TIMSS und PIRLS. Diese enge Verknüpfung zeigt, dass internationale LSA einen wichtigen Beitrag zur Beobachtung und Steuerung von Bildungsprozessen auf globaler Ebene sowie im Kontext supranationaler Zielvereinbarungen leisten.

Zusammenfassend liefern die LSA wichtige Informationen zu Bildungserträgen, wie sie in der Gesamtstrategie der KMK (2006, 2015) intendiert wurden. Gleichzeitig haben sich die LSA weiterentwickelt und liefern darüber hinausgehende Informationen, die bislang gar nicht oder unsystematisch im System des Bildungsmonitorings in Deutschland genutzt werden.

5.1.4 Welche Informationen liefert der Nationale Bildungsbericht?

Anders als die LSA beschränkt sich der Nationale Bildungsbericht nicht auf den (allgemeinbildenden) schulischen Bereich, sondern nimmt im zweijährigen Turnus alle Bildungsetappen (von der Kindertagesstätte bis zur Erwachsenenbildung) in den Blick.⁷¹ Dem Bildungsbericht liegt dabei ein weites Indikatorenverständnis zugrunde: Bildungsbereiche werden nicht anhand einzelner Kennziffern beschrieben, sondern

⁷¹ <https://www.bildungsbericht.de/de/forschungsdesign> (aufgerufen am 20.02.2026).



durch die systematische Verknüpfung mehrerer Indikatoren, die unterschiedliche strukturelle, prozessuale und ergebnisbezogene Aspekte eines Bereichs abbilden.

Ein in Teilen vergleichbares, allerdings international ausgerichtetes Konzept verfolgt der von der OECD regelmäßig aufgesetzte Bericht *Education at a Glance*.⁷² Er basiert auf einem eher engen Indikatorenverständnis, das vielfach auf wenige, stark verdichtete statistische Kennzahlen zurückgreift und im Extremfall einzelne zentrale Maßzahlen in den Mittelpunkt stellt. Zudem erlaubt *Education at a Glance* zwar internationale Vergleiche, trifft jedoch keine differenzierten Aussagen zu den Ländern innerhalb der Bundesrepublik Deutschland. Internationale Vergleichsdaten (z. B. aus OECD-, EU- oder UNESCO-Statistiken) werden bislang noch nicht ausreichend systematisch in die nationale Bildungsberichterstattung integriert, um das deutsche Bildungssystem konsistent im europäischen und globalen Kontext zu verorten. Ob und in welchem Umfang solche internationalen Vergleichsdaten stärker herangezogen werden sollten, hängt wesentlich von der jeweiligen Zielrichtung des Systemmonitorings ab. Neben internationalen Vergleichsperspektiven ist für Fragen der datengestützten Qualitäts- und Systementwicklung der nationale Vergleich innerhalb Deutschlands mindestens ebenso bedeutsam – etwa zwischen Ländern, Regionen, Kommunen oder Jugendamtsbezirken –, da hier Steuerungs- und Entwicklungsimpulse unmittelbar ansetzen können.

Weiterhin widmet sich die nationale Bildungsberichterstattung in einem von Bericht zu Bericht variierenden Schwerpunktkapitel einer übergreifenden Fragestellung, die entweder durch besondere Aktualität gekennzeichnet ist oder sich mit den vorhandenen Indikatoren bislang noch nicht hinreichend abbilden lässt. Dabei wird in der Regel auf die Erhebung von Primärdaten verzichtet; stattdessen erfolgt eine problembezogene Aufbereitung und Zusammenführung existierender Daten der amtlichen Statistik, der nationalen und internationalen LSA, des NEPS sowie insbesondere für den Schwerpunkt relevanter nationaler Erhebungen. Ergänzend werden für jeden Bericht koordinierte Länderabfragen durchgeführt, um Informationen zu generieren, die bislang nicht länderübergreifend vergleichbar erhoben werden oder in bestehenden Datenquellen nicht bzw. nur eingeschränkt zugänglich sind. In den Indikatoren wie im Schwerpunkt werden Kontext-, Input-, Prozess- und Outputindikatoren in den Blick genommen. Die zweijährliche Berichterstattung ermöglicht es, für zahlreiche Indikatoren Entwicklungen im Zeitverlauf darzustellen; gleichzeitig stellt die Zweijahrestaktung eine Herausforderung für eine umfassende Rezeption und Verarbeitung der Befunde durch bildungspolitische Akteur:innen dar.

Die im Bildungsbericht angelegte Perspektive über den Lebensverlauf ermöglicht es, auch Erträge von Bildung (Outcomes) systematisch zu berichten. Bildungserträge und die Wirkungen von Bildung waren Thema des Schwerpunktes 2018, und der Bildungsbericht 2024 befasst sich schwerpunktmäßig mit der beruflichen Bildung (Autor:innengruppe Bildungsberichterstattung, 2018, 2024).

Viele der im Bildungsbericht aufgeführten Befunde sind auf Bundesebene aggregiert, häufig erfolgt zusätzlich eine regionale Differenzierung nach Ländergruppen West- und Ostdeutschland. Je nach Ausrichtung der Indikatoren werden Informationen teilweise auf die Ebene der 16 Länder heruntergebrochen oder in

⁷² https://www.oecd.org/de/publications/bildung-auf-einen-blick_19991509.html (aufgerufen am 20.02.2026).



ausgewählten Fällen weiter kleinräumiger ausgewiesen, z. B. auf Ebene der Gemeinden, Jugendamtsbezirke oder Arbeitsagenturbezirke.

Der Nationale Bildungsbericht versteht sich als evaluative Gesamtschau zentraler Entwicklungen im Bildungssystem über alle Bildungsbereiche hinweg. Aufgrund dieses Anspruchs und der damit verbundenen Aggregationsebene können die für die Steuerung vor Ort notwendigen regionalisierten Kennziffern in den Indikatoren nicht durchgängig bereitgestellt werden. Als Folge haben einzelne Länder und zahlreiche Kommunen eigene Verfahren der regionalen Bildungsberichterstattung etabliert. Im Jahr 2025 lassen sich sechs Länder identifizieren, die regelmäßig einen eigenen Bildungsbericht veröffentlichen. Sieben Länder haben in der Vergangenheit Bildungsberichte veröffentlicht, wobei eine Fortsetzung der Berichterstattung nicht erkennbar ist. Drei Länder haben bislang keinen eigenen Bildungsbericht veröffentlicht.

In allen Flächenländern gibt bzw. gab es neben den landesweiten Bildungsberichten kommunale Bildungsberichte. Gemäß der ausgewerteten Bibliothek der *Transferinitiative Kommunales Bildungsmanagement*⁷³ sind mittlerweile mehr als 280 kommunale Bildungsberichte von 120 Kreisen, kreisfreien Städten sowie kreisangehörigen Kommunen erschienen. Der Zenit der Publikationstätigkeit scheint jedoch überschritten zu sein. Wurden im Jahr 2018 noch über 30 kommunale Bildungsberichte veröffentlicht, so waren es im Jahr 2022 zwölf. Als Best-Practice-Beispiel für eine kommunale Bildungsberichterstattung kann die Stadt München genannt werden. Dort erscheinen alternierend ein allgemeiner Bildungsbericht (zuletzt Landeshauptstadt München, 2022) und ein Bericht zur beruflichen Bildung (zuletzt Landeshauptstadt München, 2024). Der allgemeine Bildungsbericht folgt in seiner Gliederung dem Nationalen Bildungsbericht.

Zusammenfassend stellt der Nationale Bildungsbericht zentrale Informationen über das Bildungssystem in Deutschland zur Verfügung, die in Umfang und thematischer Breite über die Inhalte der LSA hinausgehen, aber auch die Daten der LSA in der Berichterstattung nutzen. Neben der Darstellung von Trends ermöglicht er auch längsschnittliche Analysen und ist bildungsbereichsübergreifend konzipiert. Die zusätzlich erscheinenden Bildungsberichte auf Ebene der Länder und Kommunen sind nicht hinreichend mit der nationalen Berichterstattung abgestimmt. In diesem Zusammenhang stellt sich die Frage, welchen konkreten Nutzen die unterschiedlichen Berichtssysteme entfalten und wie sie in einem kohärenten Systemmonitoring besser aufeinander abgestimmt werden können.

5.1.5 Adressaten von Large-Scale Assessments und Nationalem Bildungsbericht

LSA liefern stichprobenbasierte Informationen zur Leistungsfähigkeit des Primar- und Sekundarschulsystems. Der Nationale Bildungsbericht ergänzt diese Informationen, indem er die übrigen Bildungsetappen über die gesamte Lebensspanne einbezieht. Darüber hinaus thematisiert der Bildungsbericht übergreifende Rahmenbedingungen des Bildungsgeschehens, etwa demografische und wirtschaftliche Entwicklungen, Fragen der Bildungsfinanzierung sowie weitere zentrale gesellschaftliche Entwicklungen. Auf diese Weise ermöglicht er eine integrierte Gesamtschau des Bildungssystems, in der Bildungsprozesse und -ergebnisse in ihren strukturellen, institutionellen und gesellschaftlichen Kontexten verortet werden. Die

⁷³ https://www.transferinitiative.de/bibliothek_bb_a-z.php (aufgerufen am 20.02.2026).



Auftraggeber und Adressaten der Vorhaben bzw. ihrer Berichte sind in erster Linie diejenigen, die Steuerungsaufgaben im System übernehmen, konkret: Bildungspolitik und Bildungsadministration auf der Ebene der 16 Länder und des Bundes. Diese Gruppen erwarten von den Berichten, dass sie ihnen wissenschaftlich fundierte Hinweise zur politischen Steuerung des Bildungssystems liefern. Ferner sind die Befunde der LSA und des Bildungsberichts für Gruppen (Stakeholder) relevant, für die es wichtig ist, evidenzbasiert beratend, teilweise gestaltend an der Steuerung im Bildungssystem beteiligt zu sein: Lehrkräfte aus-, fort- und weiterbildende Einrichtungen, die Qualitätsagenturen der Länder, Schulträger, Lehrkräfteverbände/Gewerkschaften, Stiftungen, Arbeitgeberverbände und nicht zuletzt die universitäre und außeruniversitäre Bildungsforschung mit ihren unterschiedlichen Disziplinen (Erziehungswissenschaft, Fachdidaktiken, Ökonomie, Psychologie und Soziologie) sowie – vermittelt über Medien – die bildungsinteressierte Öffentlichkeit. Die Bildungsforschung nutzt die Daten der nationalen und internationalen LSA, die hierfür am FDZ zur Verfügung gestellt werden, primär für Sekundäranalysen, deren Ergebnisse zur Beantwortung wissenschaftlicher Fragestellungen beisteuern sollen. Ergänzend werden auch die Daten und Indikatoren des Nationalen Bildungsberichts genutzt, die über die Homepage des Bildungsberichts kostenfrei zugänglich sind und insbesondere für bildungsbereichsübergreifende sowie kontextualisierende Analysen herangezogen werden können.

Zu betonen ist, dass LSA und der Nationale Bildungsbericht keinerlei Steuerungskraft auf Ebene der einzelnen Kommunen oder gar der Einzelschulen entfalten wollen und können. Dazu dienen Instrumente, die an anderer Stelle des Gutachtens (vgl. Kapitel 1, 2) beschrieben werden. Eine öffentliche Diskussion der Befunde der LSA kann allerdings Problembewusstsein sowie Interesse an Kompetenzmodellen und Aufgabentypen auf der Schulebene wecken. Eine Stärke der auf den Bildungsstandards basierenden Instrumente für Schulentwicklung, insbesondere VERA, besteht darin, dass sie auf dieselben Ziele ausgerichtet sind wie das nationale Bildungsmonitoring (Alignment).

5.1.6 Berichtsformate der Large-Scale Assessments und des Nationalen Bildungsberichts

Bereits mit der Berichterstattung zu TIMSS (Baumert et al., 1997) wurde die Praxis gepflegt, einen eigenen, deutschen Ergebnisbericht zu erstellen, der parallel zum internationalen Bericht publiziert wurde. Die nationalen Berichte boten die Möglichkeit, die Besonderheiten des deutschen Bildungssystems stärker zu berücksichtigen, Ergänzungen in den Ergebnisberichten vorzunehmen und vor allem eine höhere Passung zu den nationalen Adressaten der Berichte herzustellen. Bereits die ersten Bände waren sehr umfangreich und umfassten mehrere Hundert Seiten. Die Inhalte sind überwiegend beschreibend, im Duktus eher wissenschaftlich-analytisch. Neben deskriptiven Befunden (internationale Vergleiche mit den OECD-Mitgliedstaaten und anderen Gruppen von Vergleichsstaaten sowie soziale, migrationsbedingte und Geschlechterdisparitäten) werden Ergebnisse aus komplexeren multivariaten statistischen Analysen vorgestellt. Zusätzliche Kapitel informieren über die methodischen Details der Studien. Die Berichte werden üblicherweise im Rahmen von Pressekonferenzen der Öffentlichkeit vorgestellt. Ergänzend werden die Befunde im Rahmen eigens organisierter Fachtagungen den relevanten Adressatengruppen (s. o.) präsentiert. Zudem präsentieren und diskutieren die Studienverantwortlichen die Befunde auf Veranstaltungen der Stakeholder.



Für weite Teile der genannten Adressaten dürfte die Rezeption der umfangreichen Berichte eine erhebliche Herausforderung sein. Ein Austausch zwischen Wissenschaft und Politik über mögliche Implikationen und Konsequenzen aus den Studien findet zwar statt, etablierte Formate oder Strukturen hierfür gibt es jedoch nicht. Die fehlenden Transferstrukturen mögen auch erklären, warum die politischen Adressat:innen vielfach die Klage erheben, dass das Systemmonitoring kein ausreichendes Erklärungswissen sowie keine Interpretation der Ergebnisse liefere und daher keine Basis für daraus zu ziehende Konsequenzen biete.

Zusammenfassungen der Studien und im Duktus einfacher gehaltene Broschüren runden die Berichte ab und sollen zur Dissemination der Studienergebnisse beitragen. Die Rezeption der Berichte und der Kurzfassungen wurde bislang jedoch nicht untersucht. Digitale Formen der Ergebnispräsentation liefern die einzelnen Instrumente in unterschiedlicher Detailtiefe, aber interaktive Dashboards, wie sie international teilweise Standard sind, werden für das deutsche Systemmonitoring noch nicht vorgehalten.

5.1.7 Federführung für Large-Scale Assessments: Institutionalisierung vs. wissenschaftliche Konsortien

Die Federführung der in Abbildung 12 aufgeführten Vorhaben ist unterschiedlich. Das IQB als An-Institut (eingetragener Verein) an der Humboldt-Universität zu Berlin und unabhängige wissenschaftliche Einrichtung der 16 Länder verantwortet die Bildungstrends und die Aufgabenentwicklung für VERA sowie die Koordination der Entwicklung ländergemeinsamer Pools von Abiturprüfungsaufgaben. Es ist mit seinen Aufgaben im Bereich des Bildungsmonitorings auf Dauer gestellt und wird im Kern anteilig von den 16 Ländern finanziert. Das IQB wird regelmäßig evaluiert.

Das ZIB, das das Nationale Projektmanagement für PISA verantwortet, ist als eingetragener Verein ein An-Institut der Technischen Universität München (TUM). Mitglieder sind die 16 Länder und der Bund. Das ZIB verfügt über weitere Standorte am Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF) in Frankfurt a. M., am Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel sowie am IQB. Die Grundfinanzierung wird durch die Vereinsmitglieder gewährleistet, wobei die Sitzländer von TUM, IPN und DIPF höhere Kostenanteile übernehmen. Die Stiftungsprofessuren sowie Teile des Personals in PISA bzw. des FDZ am IQB sind entfristet. Neben dem nationalen Projektmanagement wird am ZIB Forschung betrieben, insbesondere national vertiefende Analysen der PISA-Daten sowie Methoden- und Prozessdatenforschung im Bereich LSA. Für seine Förderung stellt das ZIB regelmäßig Anträge. Das Arbeitsprogramm wird vom Kuratorium verabschiedet. Analog zum IQB wurde auch das ZIB wiederholt evaluiert.

Alle weiteren LSA in Abbildung 12 werden von unabhängig arbeitenden wissenschaftlichen Konsortien im Rahmen zeitlich befristeter Fördervorhaben finanziert. Hierfür werden Drittmittelanträge (über das Sekretariat der KMK oder das Bundesministerium) gestellt. Das Arbeitsprogramm wird von den leitenden Wissenschaftler:innen der Konsortien festgelegt. Für TIMSS und PIRLS, die Teile der KMK-Gesamtstrategie sind, übernehmen vergleichbar zu PISA die Länder sämtliche nationalen Kosten; der Bund trägt die internationalen Projektkosten. Die Kosten für ICILS trägt bisher vollständig der Bund; dasselbe gilt für ICCS 2027.



Die Finanzierung des Nationalen Bildungsberichts erfolgt über eine langfristige Projektförderung von Bund und Ländern (aktuell 2024 bis 2032). Das DIPF als federführende Einrichtung koordiniert die Arbeiten der institutionell besetzten Autor:innengruppe⁷⁴ und vertritt diese nach außen. Die Bildungsberichterstattung wird dabei eng von einer Steuerungsgruppe von Bund und Ländern begleitet, in der auch Wissenschaftler:innen als Gäste vertreten sind. Sie einigt sich über das jeweilige Schwerpunktthema des Bildungsberichts. Die Bildungsberichterstattung wurde in den Jahren 2015 und 2023 evaluiert. Die Evaluierung im Jahr 2023 war die Grundlage für die längerfristige Weiterfinanzierung der Autor:innengruppe Bildungsberichterstattung.

Hinsichtlich des Austausches zwischen den wissenschaftlichen Akteur:innen des Bildungsmonitorings über ihre jeweiligen Vorhaben ergeben sich unterschiedliche Konstellationen. Durch die Konstruktion des ZIB – das FDZ des IQB ist Teil des ZIB und der wissenschaftliche Vorstand des IQB gehört dem ZIB-Vorstand an – ist ein regelmäßiger Austausch hinsichtlich der Planungen in den IQB-Bildungstrends und PISA unmittelbar und institutionalisiert gewährleistet. Dies hat in der Vergangenheit wiederholt zum gemeinsamen Sampling oder auch zu einer Längsschnittkomponente in PISA 2012 geführt (Prenzel et al., 2013; Reiss et al., 2017), wobei in jüngerer Zeit keine gemeinsamen Vorhaben zwischen ZIB und IQB mehr durchgeführt wurden. Weiterhin sind wissenschaftliche Mitglieder der Konsortien anderer Studien teilweise auch Mitglieder im Vorstand des ZIB. So ist Prof. Dr. Goldhammer (DIPF; im Vorstand des ZIB) nationaler Projektmanager für ICILS 2028; er und Prof. Dr. Köller (IPN; im Vorstand des ZIB) sind zudem Mitglieder des TIMSS-Konsortiums. Dadurch ist ein Austausch zwischen PISA, ICILS und TIMSS möglich, der sich in der Regel vor allem auf Fragen der Digitalisierung im Unterricht und der damit verbundenen digitalen Kompetenzen von Schüler:innen und Lehrkräften bezieht. Insgesamt ist festzuhalten, dass übergeordnete Strategien zur Verzahnung der Studien ausbaufähig sind und mehr Potenzial bieten, als derzeit genutzt wird.

5.1.8 Kosten der Large-Scale Assessments und der nationalen Bildungsberichterstattung

In der Gesamtschau liegen die jährlichen Aufwendungen für das Bildungsmonitoring bei rund 9 Mio. Euro. Darin enthalten sind die internationalen Studien (PISA, IGLU, TIMSS, ICILS, ICCS), die IQB-Bildungstrends sowie die Bildungsberichterstattung. Zugleich zeigt sich, dass die einzelnen Studien bzw. die einzelnen IQB-Bildungstrends pro Erhebungsrunde jeweils Kosten in vergleichbarer Größenordnung verursachen. In der Gesamtschau entspricht der jährliche Mitteleinsatz etwa der institutionellen Jahreszuwendung eines mittelgroßen Instituts der Leibniz-Gemeinschaft.

5.1.9 Der Blick ins Ausland

In Ländern, die aus unterschiedlichen Perspektiven Vergleiche mit und Anregungen für Deutschland bieten können (Österreich, Schweiz, Kanada, Niederlande), werden divergierende Wege besprochen.

⁷⁴ Dazu zählen neben dem DIPF das Deutsche Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen (DIE), das Deutsche Jugendinstitut (DJI), das Deutsche Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), das Leibniz-Institut für Bildungsverläufe (LifBi), das Soziologische Forschungsinstitut an der Universität Göttingen (SOFI) und die statistischen Ämter des Bundes (Destatis) und der Länder.



In Österreich existiert das Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen (IQS) als nachgeordnete Behörde des österreichischen Bundesministeriums. Das IQS ist aus dem Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE), einer unabhängigen Einrichtung für Bildungsmonitoring, hervorgegangen. Es verantwortet aktuell sämtliche nationalen (standardbasierten) und internationalen Large-Scale Assessments. Das nationale Assessment (*individuelle Kompetenzmessung Plus; iKM^{PLUS}*), das die Klassenstufen 3 und 4 in der Primarstufe sowie 7 und 8 in der Sekundarstufe I umfasst, wird seit dem Schuljahr 2021/22 (Primarstufe) bzw. 2022/23 (Sekundarstufe) flächendeckend durchgeführt und ermöglicht somit sowohl ein Monitoring auf Systemebene als auch Rückmeldungen auf Schul-, Klassen- und Individualebene (Stauber et al., 2024). Die Testungen werden von den Lehrkräften selbst durchgeführt, ein Vorgehen, wie es auch in Deutschland in fast allen Ländern in den VERA-Erhebungen praktiziert wird. Durch die Erhebungen in aufeinanderfolgenden Klassenstufen lassen sich auch Veränderungsanalysen über ein Schuljahr durchführen. iKM^{PLUS} wird so zu einem Instrument, das, bezogen auf Deutschland, die Bildungstrends und die Vergleichsarbeiten kombiniert. Dass die Testungen von Lehrkräften durchgeführt werden, kann die Präzision der Ergebnisse allerdings einschränken. Das IQS ist über das nationale Assessment hinaus auch für das nationale Projektmanagement internationaler Studien zuständig. Österreich beteiligt sich an PISA, PIRLS, TIMSS (abwechselnd vierte und achte Jahrgangsstufe), TALIS und ICILS (erstmalig 2023). An ICCS nimmt Österreich nicht teil. Schließlich unterstützt das IQS die Erstellung des Nationalen Bildungsberichts, der vom österreichischen Bundesministerium für Bildung herausgegeben wird.

In der Schweiz sind die Aufgaben des Bildungsmonitorings nicht institutionalisiert. Die nationale *Überprüfung der Grundkompetenzen (ÜGK)* und PISA werden in wissenschaftlichen Konsortien unter Ägide des Interfaculty Centre for Educational Research (ICER) der Universität Bern durchgeführt. ÜGK findet alternierend in den Jahrgangsstufen 4, 8 und 11 statt; die stichprobenbasierten Daten lassen sich mit Daten des Schweizer Bildungsregisters verknüpfen. Den nationalen Bildungsbericht verantwortet die Schweizerische Koordinationsstelle für Bildungsforschung (SKBF), eine gemeinsame Institution der Schweizerischen Eidgenossenschaft und der Konferenz der kantonalen Erziehungsdirektorinnen und -direktoren (EDK). Die Schweiz beteiligt sich an PISA, jedoch nicht an TIMSS, PIRLS, ICILS und ICCS.⁷⁵

Im föderal organisierten Kanada liegt die Verantwortung für das Bildungswesen bei den einzelnen Provinzen und Territorien. Die Zusammenarbeit in Fragen der Bildungsberichterstattung und der LSA wird vom Council of Ministers of Education, Canada (CMEC) gemeinschaftlich koordiniert. Das zentrale nationale Assessment ist das *Pan-Canadian Assessment Program (PCAP)*. In der achten Jahrgangsstufe werden Basiskompetenzen in den Bereichen Lesen, Mathematik und Naturwissenschaften erfasst. Die Durchführung obliegt den einzelnen Provinzen. Die erhobenen Daten werden beim CMEC zusammengeführt, wo auch die Berichtslegung erfolgt (zuletzt Council of Ministers of Education, 2026). Für die Durchführung internationaler Erhebungen wie PISA, PIRLS, TIMSS und ICILS sind die Bildungsministerien der Provinzen zuständig, während das CMEC die nationale Koordination sowie die Abstimmung mit den internationalen Konsortien übernimmt. Nationale und internationale LSA werden in mehreren Provinzen gezielt für datengestützte

⁷⁵ Das Monitoring liegt in der Zuständigkeit des Bundes (<https://www.sbf.admin.ch/de/bildungsmonitoring>) und der Kantone (<https://www.cdpe.ch/de/themen/bildungsmonitoring>) (aufgerufen am 20.02.2026).



Schulentwicklungsprozesse genutzt, etwa durch schul- oder distriktbezogene Entwicklungspläne. Einen nationalen Bildungsbericht, vergleichbar mit dem deutschen Pendant, gibt es nicht. Mit dem gemeinsam von CMEC und Statistics Canada herausgegebenen *Education Indicators in Canada* liegt jedoch ein indikatorengestütztes Format vor, das regelmäßig zentrale Kennzahlen zum Bildungssystem Kanadas sowie zu den Provinzen publiziert.⁷⁶

Das niederländische Bildungswesen ist geprägt von dezentraler Organisation und weitreichender Schulautonomie. Nationale LSA werden unter Aufsicht der Inspectie van het Onderwijs, einer Behörde des Ministeriums für Bildung, Kultur und Wissenschaft, durchgeführt. Sie verantwortet sowohl die Qualitätsüberwachung als auch die inhaltliche Steuerung. Die nationalen Untersuchungen werden vom Centraal Instituut voor Toetsontwikkeling (Cito) entwickelt und umgesetzt. Sie finden in mehrjährigen Abständen statt und erfassen den Kompetenzstand in zentralen Lernbereichen, um langfristige Trends und Entwicklungsbedarfe zu identifizieren. Die niederländische Teilnahme an internationalen Studien wie PISA, TIMSS und PIRLS wird von der Netherlands Initiative for Education Research im Vergabeverfahren beauftragt und durch Konsortien umgesetzt (z. B. Expertisecentrum Nederlands, Universität Twente, KBA Nijmegen, Cito) (Anneke C. Timmermans et al., 2022). Die Inspectie van het Onderwijs nutzt die Ergebnisse der nationalen LSA für systeminterne Entwicklungen und die internationalen LSA zur Einordnung der niederländischen Leistungen im internationalen Vergleich sowie zur nationalen Qualitätsentwicklung. Mit dem jährlich ebenfalls von der Inspectie van het Onderwijs vorgelegten *De Staat van het Onderwijs* verfügt das Land über einen indikatorengestützten Bildungsbericht, der sowohl Ergebnisse nationaler Studien als auch internationaler Erhebungen zusammenführt.⁷⁷

Der Blick in die anderen Staaten zeigt, dass auch dort nationale und internationale LSA durchgeführt werden und diese Hand in Hand mit einer nationalen Berichterstattung gehen. Teilweise werden die Programme durch nachgeordnete Einrichtungen der Regierungen verantwortet, teilweise durch unabhängige wissenschaftliche Einrichtungen bzw. Konsortien. Da mit jedem Modell Vor- und Nachteile verbunden sind, die in Abhängigkeit vom Kontext in unterschiedlichem Maße relevant sind, lassen sich diese nicht allgemein beurteilen. So wäre es z. B. denkbar, VERA in Orientierung an IKM^{PLUS} weiterzuentwickeln und die Ergebnisse für das Systemmonitoring zu nutzen. Die potenziellen Einschränkungen der Belastbarkeit der Daten könnten jedoch das Vertrauen in die Ergebnisse ländervergleichender Analysen reduzieren.

5.1.10 Reformen und Programme im schulischen Bildungssystem infolge des Bildungsmonitorings

Mit der empirischen Wende in der Bildungspolitik haben die öffentlichen bildungspolitischen Debatten deutlich an Fahrt aufgenommen. Themen sind u. a. Leistungsniveaus, Bildungsgerechtigkeit und die Integrationsfunktion der Bildung. Die Befunde der großen Studien fanden breiten Eingang in die überregionale Presse, Bildungsjournalist:innen professionalisierten sich in der Rezeption und Interpretation der Ergebnisse. Auf der Seite der Bildungsforschung waren und sind Kompetenzen in der Wissenschafts-

⁷⁶ [Education Indicators in Canada: Report of the Pan-Canadian Education Indicators Program](#) (aufgerufen am 20.02.2026).

⁷⁷ [De Staat van het Onderwijs | Inspectie van het onderwijs](#) (aufgerufen am 20.02.2026).



kommunikation gefragt. Es wurde Expertise aufgebaut, um Befunde gegenüber der Politik zu kommunizieren und gemeinsam mit ihr über mögliche Reformen zu beraten.

Tatsächlich wurden dann viele Reformen im Bildungssystem geplant und umgesetzt, die Bezug auf die Ergebnisse des Bildungsmonitorings nehmen. Davon sollen einige exemplarisch im Folgenden aufgeführt werden. Auf TIMSS folgte das Bund-Länder-Programm *Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts (SINUS)*, das von 1998 bis 2013 zunächst im Sekundarbereich, später im Grundschulbereich umgesetzt wurde. Die TIMSS-Oberstufenuntersuchung gab wichtige Impulse für die G8-Reform, die später in den meisten Ländern wieder zurückgenommen wurde. Als Folge von PISA 2000 wurde SINUS ausgeweitet. Auf SINUS bauten später die sogenannten Kontextprojekte (*Biologie im Kontext, Chemie im Kontext, Physik im Kontext*) auf. Die ersten beiden PISA-Erhebungen gaben Impulse für die Einführung des Ganztagsprogramms mit Mitteln der Bundesregierung, die Einführung der flexiblen Eingangsphase im Primarbereich, die Reduktion der Klassenwiederholungen und die zunehmende Abkehr von der Drei- bzw. Viergliedrigkeit zugunsten eines Zwei-Säulen-Modells in der Sekundarstufe I mit dem Gymnasium und einer weiteren allgemeinbildenden Schulform, die vielfach den Erwerb des Ersten und Mittleren Schulabschlusses sowie des Abiturs ermöglicht. Im Bereich der Sprachförderung wurde in der Folge von PISA im Jahr 2007 das Bund-Länder-Programm *Förderung von Kindern und Jugendlichen mit Migrationshintergrund (FörMig)* aufgelegt und später im Anschlussprojekt *FörMig-Transfer* fortgesetzt. Es folgte im Jahr 2012 vor allem auch als Folge der PISA-Befunde das auf fünf Jahre angelegte Bund-Länder-Programm BiSS, das später in BiSS-Transfer überging.

Eine Folge von PIRLS und den IQB-Bildungstrends sind die 2015 in Hamburg und später auch in anderen Ländern eingeführten Lesebänder zur Förderung der Leseflüssigkeit. An den Universitäten in Deutschland wurden durch die *Qualitätsoffensive Lehrkräftebildung (QLB)* wichtige Innovationen angestoßen, die nicht zuletzt auf Befunde zu den professionellen Kompetenzen von Lehrkräften in den LSA zurückgingen. Deutlich größer als die genannten Programme ist das seit 2024 in 15 Ländern durchgeführte QuaMath-Programm zur Steigerung der Unterrichtsqualität in Mathematik, das ebenfalls durch Ergebnisse in den nationalen und internationalen LSA-Studien angeregt wurde und innerhalb von zehn Jahren 10.000 Schulen der unterschiedlichen Bildungsetappen erreichen soll. Schließlich soll hier das in der Geschichte der Bundesrepublik größte Bildungsprogramm, das Startchancen-Programm, genannt werden, in dem 20 Mrd. Euro über zehn Jahre zur Verfügung gestellt werden, um Disparitäten im Bildungssystem zu verringern und die sogenannte Risikogruppe zu halbieren.

Auch auf Ebene einzelner Länder haben die Ergebnisse der nationalen LSA auf Problembereiche aufmerksam gemacht und dadurch Impulse für Veränderungen gesetzt. So haben die relativ schwachen Englischleistungen in den ostdeutschen Ländern im IQB-Ländervergleich 2009 Fortbildungsinitiativen ausgelöst, und ungünstige Leistungsentwicklungen haben einzelne Länder veranlasst, eine Strategie der datengestützten Qualitätsentwicklung auf den Weg zu bringen.

Diese Aufzählung der Reformmaßnahmen macht deutlich, dass das Bildungsmonitoring in Deutschland erhebliche Impulse gegeben hat. Die durch das Bildungsmonitoring gestützten Programme wurden vielfach großflächig umgesetzt, wobei ihre Wirksamkeit nur selten evaluiert wurde. Beispiele für solche Evaluationen waren ergänzende Erhebungen zu SINUS in der Sekundarstufe im Rahmen von PISA 2003 (Prenzel,



Carstensen et al., 2005) sowie zu SINUS-Grundschule im Rahmen von TIMSS (Dalehefte et al., 2014) und des IQB-Ländervergleichs 2012 (Pant et al., 2013).

5.1.11 Impulse für die Bildungsforschung

Die empirische Wende der 1990er-Jahre und insbesondere die Veröffentlichung der PISA-2000-Ergebnisse markierten einen tiefgreifenden Wandel in der deutschen Bildungsforschung. Sie führten zu einer Neuausrichtung der Forschungsförderung, die maßgeblich durch Initiativen der Deutschen Forschungsgemeinschaft (DFG) und später des Bundesministeriums für Bildung und Forschung (BMBF) geprägt wurde. Die DFG etablierte Anfang der 2000er-Jahre mit den *Forschergruppen Empirische Bildungsforschung* Strukturen zur interdisziplinären Erforschung zentraler Bildungsfragen und trug damit entscheidend zum Aufbau universitärer Zentren in Bamberg, Tübingen und Duisburg-Essen bei. Diese Initiativen förderten nicht nur die Institutionalisierung empirischer Schul- und Unterrichtsforschung, sondern auch die Ausbildung einer neuen Generation empirisch arbeitender Wissenschaftler:innen. In eine ähnliche Richtung ging das ebenfalls zu Beginn der 2000er-Jahre gestartete DFG-Schwerpunktprogramm *Bildungsqualität von Schule (BIQUA)*, das Forschungsprojekte zur Effizienzsteigerung des Fachunterrichts förderte. Daran anschließend startete die DFG im Jahr 2007 das sechsjährige Schwerpunktprogramm *Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen*, das auf die Einführung der Bildungsstandards reagierte und Kompetenzstruktur- und Kompetenzentwicklungsmodelle systematisch erforschen sollte.

Die BMBF-Programmförderung setzte ab 2007 mit dem *Rahmenprogramm Empirische Bildungsforschung* an dieser Entwicklung an und zielte auf eine nachhaltige Stärkung, Internationalisierung und Interdisziplinarisierung des Forschungsfeldes. Ein Schwerpunkt lag dabei auf der Nachwuchsförderung und der strukturellen Verstärkung der empirischen Bildungsforschung. Als besonders erfolgreich können die Etablierung des Nationalen Bildungspanels (NEPS) und dessen Integration in die Leibniz-Gemeinschaft gelten, wodurch die langfristige institutionelle Verankerung empirischer Bildungsforschung in Deutschland weiter konsolidiert wurde.

Mit Zusatzanalysen der Daten aus den IQB-Bildungstrends konnten u. a. wichtige Informationen zu Effekten der Inklusion von Schüler:innen mit sonderpädagogischem Förderbedarf ins allgemeinbildende Schulsystem gewonnen werden (z. B. Kocaj et al., 2014). Bereits mit der TIMSS-Videostudie in den 1990er-Jahren wurde die empirisch ausgerichtete Unterrichtsforschung in Deutschland deutlich gestärkt, und dieser Schub in der fachbezogenen Forschung zu gutem Unterricht hält bis heute an. Zunehmend werden Interventionsstudien zu wirksamen fachlichen und überfachlichen Förder- und Implementationsmaßnahmen angestoßen, die die treffsichere Nutzung von Daten fundieren.

5.1.12 Die Rolle einer Bildungsverlaufsstatistik im Bildungsmonitoring

Die Einrichtung einer einheitlichen Bildungsverlaufsstatistik bzw. eines Bildungsverlaufsregisters als systematische, längsschnittliche Abbildung von Bildungsverläufen wird seit Längerem als notwendiger Bestandteil einer datengestützten Bildungssteuerung diskutiert und ist sowohl bildungspolitisch als auch administrativ mehrfach aufgegriffen worden (Giar et al., 2023; Hertweck et al., 2023; RatSWD, 2022). Bereits ein KMK-Beschluss aus dem Jahr 2020 zielte darauf ab, die Voraussetzungen dafür zu verbessern (KMK,



2020b). Im aktuellen Koalitionsvertrag auf Bundesebene⁷⁸ wird die Einführung erneut aufgegriffen. Mit der Einrichtung einer länderübergreifenden Bildungsverlaufsstatistik werden neue Möglichkeiten zur datengestützten Qualitätsentwicklung geschaffen. Dabei ist noch ungeklärt, welche Daten die Bildungsverlaufsstatistik genau enthalten und wer für welche Zwecke darauf zugreifen kann. Grundsätzlich können mit einer Bildungsverlaufsstatistik längsschnittliche Betrachtungen von Bildungsverläufen (vgl. Kapitel 1.2.1) über die unterschiedlichen Ebenen und Etappen des Bildungssystems hinweg ermöglicht werden. Sie erweitern bestehende, stichprobenbasierte Längsschnittstudien wie das Nationale Bildungspanel.

Um längsschnittliche Verknüpfungen umzusetzen, können Record-Linkage (Schnell, 2022) oder Schüler-IDs zum Einsatz kommen, die eindeutig jeder Schülerin bzw. jedem Schüler zugewiesen werden. Hiermit lassen sich (theoretisch) die Daten der Schulstatistik mit Daten aus Kompetenztestungen kombinieren. Datenschutzrechtlich und politisch ist dieses Vorgehen allerdings umstritten (vgl. A. Kuhn et al., 2025). Noch weiter geht eine Bildungs-ID, die die Verknüpfung der Daten von Lernenden über unterschiedliche Bildungsetappen – von der Kindertagesstätte bis in die Erwachsenenbildung – ermöglicht. Auch hier sind die datenschutzrechtlichen Bedenken erheblich. Gleichzeitig werden über eine Bildungs-ID und damit verbunden über eine Bildungsverlaufsstatistik mit vielfältigen Daten zu individuellen Bildungsverläufen Voraussetzungen geschaffen, um besser datengestützt im Bildungssystem steuern zu können; ganz abgesehen von den vielfältigen Möglichkeiten, die ein Bildungsverlaufsregister für die Bildungsforschung bietet.

Die Bedeutung einer Bildungsverlaufsstatistik für das Bildungsmonitoring auf Systemebene ist zu klären. Erfahrungen aus Ländern wie der Schweiz oder Österreich zeigen, dass registergestützte Ansätze das Monitoring deutlich vertiefen können, ohne stichprobenbasierte Studien zu ersetzen. Es ist weitgehend unbestritten, dass Daten von Schüler:innen aus einer Bildungsverlaufsstatistik an die Daten der LSA herangespielt werden könnten, um sie nicht mehr zusätzlich mithilfe von Fragebögen erheben zu müssen. Dasselbe gilt auf Ebene der individuellen Schule bereits jetzt für die gemäß dem Kerndatensatz in den Ländern gesammelten Daten (KMK, 2024d), deren Potenzial beispielsweise für die IQB-Bildungstrends erheblich sein kann.

5.1.13 Potenziale künstlicher Intelligenz für Large-Scale Assessments

Die internationalen LSA haben sich seit gut zehn Jahren aufgrund der digitalen Transformation und der großen Leistungsfähigkeit von Computern erheblich weiterentwickelt. Dies gilt beispielsweise für komplexere (digitale) Simulationsaufgaben in naturwissenschaftlichen Tests. In den vergangenen Jahren wurden zudem mit den Large Language Models (LLM) neue Möglichkeiten der Testentwicklung und -auswertung geschaffen (Jaewook Lee et al., 2024). LLM ermöglichen es zunehmend besser, Items mit den angestrebten psychometrischen Eigenschaften zu generieren (Yunting Liu et al., 2025). Es bestehen somit erhebliche Potenziale für Kosteneinsparungen in der Testentwicklung und -erprobung. Darüber hinaus sind die großen Potenziale von LLM bei der automatischen Kodierung von offenen Antworten und Aufsätzen dokumentiert worden (vgl. Köller, 2025). In PISA 2025 wurden beispielsweise im Rahmen einer nationalen Ergänzung schriftliche Leistungen von Schüler:innen in Form umfangreicherer Essays (bis 300 Wörter) erhoben, die mithilfe von KI-Verfahren automatisch ausgewertet werden, ohne dass dabei Kosten durch geschulte Rater entstehen.

⁷⁸ [Koalitionsvertrag von CDU, CSU und SPD](#) (aufgerufen am 20.02.2026).



5.2 Schlussfolgerungen und Empfehlungen

In Deutschland wurde in den letzten 25 Jahren durch internationale und nationale LSA sowie die nationale Bildungsberichterstattung ein systematisches Bildungsmonitoring etabliert. Dieses System liefert datengestützt wichtige Informationen zur Weiterentwicklung des Bildungswesens. Während die internationalen und nationalen LSA sowie der ländergemeinsame Pool von Abituraufgaben einen klaren Fokus auf die Etappen des allgemeinbildenden Schulsystems legen, nimmt die nationale Bildungsberichterstattung alle Bildungsetappen in den Blick, allerdings ohne selbst Primärdaten zu erheben. Die Adressaten sind typischerweise Stakeholder oberhalb der Schulebene (u. a. Politik und Administration, Verbände, Stiftungen), denn LSA und Bildungsberichterstattung verfolgen nicht das Ziel, Impulse für die Entwicklung von Einzelschulen zu geben. Die Adressaten erhalten wertvolle Informationen über Ergebnisse und Trends im Bildungswesen, die die Grundlage für die Entwicklung von Strategien, Programmen und Maßnahmen zur Weiterentwicklung der Qualität des Systems bilden. Wichtige Funktionen sind darüber hinaus die Herstellung von Transparenz sowie die Versachlichung öffentlicher Debatten über zentrale bildungspolitische Fragen.

Die aktuelle Gesamtstrategie der KMK zum Bildungsmonitoring beschränkt sich bei den nationalen und internationalen LSA auf die Fächer, für die Bildungsstandards vorliegen. Folgerichtig umfasst sie IQB-Bildungstrends, PISA, PIRLS und TIMSS. In all diesen Untersuchungen wird die Perspektive aber erheblich über fachliche Kompetenzen hinaus erweitert. So werden auf der Outputseite zunehmend fachspezifische motivationale und emotionale Orientierungen, auf der überfachlichen Ebene Maße des Sozialverhaltens und des subjektiven Wohlbefindens erhoben. Aktuell wird mit StarS ein weiteres Instrument vorbereitet, mit dem am Übergang vom Elementar- zum Primarbereich basale sprachliche und mathematische Kompetenzen sowie Aspekte von Selbstregulation und Motivation erfasst werden sollen.

Seit der Veröffentlichung der Bildungsstandards und der Verabschiedung der Gesamtstrategie wurden in Deutschland immer wieder Forderungen nach KMK-Standards für weitere Fächer sowie nach deren Aufnahme in das Bildungsmonitoring laut. Die SWK hat in verschiedenen Veröffentlichungen dazu Vorschläge gemacht. Für StarS wird eine Integration in die Gesamtstrategie des Bildungsmonitorings (SWK, 2022a) und für ICILS die Verstetigung (SWK, 2025) empfohlen. Die Stellungnahme der SWK zur politischen Bildung (2024) legt ein weiteres langfristiges LSA in diesem Bereich nahe, lässt jedoch offen, ob die in ICCS verwendeten Instrumente dies leisten können. Will man jedoch ein Monitoring im Bereich der politischen Bildung über 2027 hinaus, sieht die SWK aktuell keine Alternative zu ICCS. In diesem Fall wäre es anzustreben, dass die deutsche Perspektive stärker in das zukünftige Framework von ICCS einfließt oder dass ICCS stärker durch nationale Elemente ergänzt wird.

Anders als im Primarbereich und in der Sekundarstufe I existieren aktuell keine LSA in der Sekundarstufe II, weder im allgemeinbildenden noch im berufsbildenden Bereich. Bei den Aufgabenpools für die Abiturprüfung, die vom IQB koordiniert und evaluiert werden, handelt es sich primär um Implementationsinstrumente, die sich in der aktuellen Form nicht für ein Bildungsmonitoring eignen. Nicht zuletzt durch die großen Fortschritte im Bereich der künstlichen Intelligenz könnte ein Bildungsmonitoring in der Sekundarstufe II mit vertretbarem Aufwand entwickelt und implementiert werden.



Die IQB-Bildungstrends und PISA sind institutionalisiert, was die Kontinuität im Personal und in den Arbeitsprozessen sichert. Expertise kann so dauerhaft ausgebaut und erweitert werden. Die verantwortlichen Institutionen werden ebenso wie die Bildungsberichterstattung regelmäßig evaluiert, und auf Basis der Evaluierungsergebnisse entscheiden die Länder (IQB) bzw. der Bund und die Länder (PISA, Bildungsberichterstattung) über die Fortführung. Das IQB und das ZIB werden ebenso wie der Nationale Bildungsbericht bis in die 2030er-Jahre finanziert. Anders die Vorhaben PIRLS, TIMSS, ICILS und ICCS, die rein über Projektförderungen finanziert sind und über deren Fortführung auf Basis von Drittmittelanträgen die Länder und der Bund im jeweiligen Studienzyklus entscheiden. Institutionell ist dieses Vorgehen teilweise mit Wechseln verbunden, beispielsweise hat die Federführung in TIMSS vor einigen Jahren von der TU Dortmund an die Universität Hamburg gewechselt. Stärken dieser Projektförderung liegen darin, dass jede Neueinreichung eines Projektantrags auf innovative Entwicklungen reagieren kann. Auch kann es bei der Projektbeantragung zu einem wünschenswerten Wettbewerb zwischen unterschiedlichen antragstellenden Einrichtungen kommen, wer das bessere nationale Erhebungskonzept hat. Schwächen dieses Nebeneinanders zwischen institutioneller und Projektförderung werden in der Versäulung der Einzelvorhaben sichtbar. Die jeweiligen Studien werden konzipiert ohne Abstimmung zwischen den Vorhaben. Eine gemeinsame Strategie hinsichtlich Dissemination und Transfer der Ergebnisse fehlt ebenfalls. Bessere Abstimmungsprozesse scheinen daher unerlässlich.

Jenseits der laufenden Programme wurde immer wieder die Diskussion darüber angestoßen, an weiteren internationalen Programmen teilzunehmen, zuvorderst an TALIS. Die SWK erkennt den Wert von TALIS, um mehr über die Lehrkräfte in Deutschland zu erfahren. Eine Alternative zur Ergänzung um eine weitere Studie wäre, die Lehrkräftebefragungen an die bestehenden Schulleistungstudien anzubinden. Dies würde auch eine inhaltliche Schwerpunktsetzung auf aktuelle politische Herausforderungen ermöglichen.

Mit Blick auf die nationale Bildungsberichterstattung ist das Verhältnis zur Bildungsberichterstattung in den Ländern und Kommunen zu klären. Zudem teilt die SWK die Einschätzung, wonach die Einrichtung einer Bildungsverlaufsstatistik wichtige zusätzliche Informationen für Trendindikatoren liefern könnte. Eine Bildungsverlaufsstatik erlaubt auch (siehe Schweiz), Daten aus den LSA und der Bildungsverlaufsstatistik zusammenzuspielen.

Vor diesem Hintergrund empfiehlt die SWK:

Empfehlung 15) Die Teilnahme an nationalen und internationalen Schulleistungstudien weiterentwickeln; dabei

1. die IQB-Bildungstrends in der Grundschule und in der Sekundarstufe I fortführen,
2. StarS und ICILS in die Gesamtstrategie der KMK integrieren,
3. die Beteiligung an PISA fortführen,
4. die Funktion weiterer internationaler LSA in der Grundschule (PIRLS/TIMSS) und in der Sekundarstufe I (ICCS) für das Bildungsmonitoring in Deutschland prüfen (Länder und Bund in Abstimmung mit der Bildungsforschung),
5. den Anteil des Einsatzes von Abiturprüfungsaufgaben aus den ländergemeinsamen Pools sukzessive erhöhen und ein regelmäßiges, transparentes Monitoring des Einsatzes durchführen,



6. den möglichen Ertrag eines Systemmonitorings in der Sekundarstufe II (allgemeine und berufliche Bildung) prüfen (Länder und Bund),
7. LSA-Instrumente für die Evaluierung von Bildungsreformen und Programmen vermehrt nutzen,
8. eine einheitliche Bildungsverlaufsstatistik einrichten und die Daten für die LSA sowie die Bildungsberichterstattung bereitstellen,
9. Lehrkräftebefragungen stärker berücksichtigen, etwa indem Instrumente in Anlehnung an TALIS in die bestehenden nationalen und internationalen LSA integriert werden,
10. ein Set sozial-emotionaler Variablen definieren, das (über Selbstberichte) regelmäßig in LSA erhoben wird,
11. KI-Tools stärker bei der Testentwicklung sowie der Auswertung von Freitextantworten und komplexen Schreibaufgaben in den sprachlichen Fächern einsetzen unter klaren Qualitäts-, Transparenz- und Datenschutzstandards.

Empfehlung 16) Die nationale Bildungsberichterstattung fortführen und weiterentwickeln; dabei

1. den Nationalen Bildungsbericht mit den Bildungsberichten der Länder eng abstimmen, durch gemeinsame Indikatorenkonzepte, abgestimmte Veröffentlichungszyklen und kohärente Themenstrukturen,
2. die bislang projektförmig organisierte Bildungsberichterstattung strukturell verstetigen und dabei prüfen, in welchen zeitlichen Abständen sie erfolgen soll,
3. die Vergleichbarkeit und längsschnittliche Verknüpfbarkeit von Daten der amtlichen Statistiken gewährleisten,
4. die Bildungsberichterstattung zur Begleitung und Bewertung von Bildungsreformen und -programmen im Zeitverlauf nutzen (z. B. Umsetzung des Rechtsanspruchs auf Ganztagsbetreuung in der Grundschule).

Empfehlung 17) Die Ergebnisse des Bildungsmonitorings intensiver nutzen; dabei

1. zeitgemäße digitale Berichtsformate abgestimmt auf klar definierte Adressatengruppen nutzen,
2. eine engere Kooperation der Einrichtungen in Deutschland sicherstellen, die LSA durchführen,
3. den regelmäßigen Austausch zwischen Wissenschaft und Ländern über Ziele des Bildungsmonitorings und Schlussfolgerungen aus den Ergebnissen der Studien sichern,
4. Prozesse zur Verarbeitung der Ergebnisse und daraus resultierender Schlussfolgerungen in den Ministerien von Bund und Ländern etablieren,
5. Förderinitiativen durch Bund und Länder für Forschung auflegen, um Fragen gezielt zu adressieren, die sich aus den Befunden der LSA ergeben,
6. Rezeptionsanalysen der veröffentlichten Berichte zu den LSA und zum Bildungsbericht durchführen.



6. Datengestützte Qualitätsentwicklung in der Frühen Bildung

In Deutschland ist das frühkindliche Bildungs- und Betreuungssystem – anders als in anderen Ländern – in die Strukturen der Kinder- und Jugendhilfe eingebettet. Die Kinder- und Jugendhilfe wird rechtlich durch das Sozialgesetzbuch (SGB VIII) geregelt und war bis 2025 im Unterschied zu den Schulen auf Bundesebene dem Bundesfamilienministerium zugeordnet. Diese Trennung wurde in dieser Legislaturperiode durch die Schaffung eines gemeinsamen Bundesministeriums für Bildung, Familie, Senioren, Frauen und Jugend (BMBFSFJ) zwar aufgelöst, unterschiedliche Steuerungslogiken bei Bund, Ländern und Kommunen, die auch einen großen Einfluss auf die Potenziale und Limitationen einer datengestützten Qualitätsentwicklung haben, bleiben jedoch bestehen. Die Steuerungslogik des frühkindlichen Bildungs- und Betreuungssystems wird hier kurz skizziert.

Die Einbettung in die Kinder- und Jugendhilfe hat wesentliche Auswirkungen auf die Entwicklung sowie die Steuerung der Qualität und Zugänglichkeit von Bildungsangeboten für Kinder vor dem Eintritt in die Grundschule. Zunächst einmal stellt die Einbettung sicher, dass frühkindliche Bildung als Teil eines ganzheitlichen Ansatzes zur Förderung und Unterstützung von Kindern und Familien verstanden wird. Die Aufgaben bestehen aus der Trias Erziehung, Betreuung und Bildung. Die Verantwortung für die frühkindliche Bildung und Betreuung in Deutschland liegt auf mehreren Ebenen: beim Bund, bei den Ländern und den Kommunen. Der Bund hat Anregungskompetenz und legt einen rechtlichen Rahmen fest, der anschließend von den Ländern in Ausführungsgesetzen konkretisiert wird. Hierzu gehören z. B. die qualifikatorischen Voraussetzungen frühpädagogischer Fachkräfte oder die Bildungspläne (Rahmen- und Orientierungspläne) der Länder sowie Vorgaben zur Qualitätssicherung. Hierdurch entstehen z. B. unterschiedliche Bildungsprogramme der Länder oder unterschiedliche Mindestqualifikationen für die Tätigkeit als frühpädagogische Fachkraft in einer Kindertageseinrichtung. Ferner sind große Unterschiede zwischen den Ländern in sozioökonomischen Ausgangslagen (z. B. Anteil armutsgefährdeter Familien, von Familien mit Migrationshintergrund) sowie strukturellen Qualitätsaspekten (z. B. Fachkraft-Kind-Relation, Zeitressourcen für Leitungsarbeit) zu verzeichnen, die bei der Weiterentwicklung des frühkindlichen Bildungssystems berücksichtigt werden müssen. Der Bund finanziert Bundesprogramme für Anliegen übergeordneter Natur (z. B. sprachliche Bildung). Die Kommunen übernehmen die praktische Umsetzung, indem sie Kindertageseinrichtungen betreiben und lokale Bildungsangebote entwickeln. Diese Aufgabenverteilung gewährleistet die Berücksichtigung der spezifischen Bedürfnisse der Regionen und Gemeinden.

Ein zentrales Prinzip, das die Steuerung des Systems beeinflusst, ist das Subsidiaritätsprinzip. Das Prinzip ist ein grundlegendes Leitprinzip der sozialen und politischen Organisation in Deutschland, das besagt, dass Entscheidungen und Aufgaben dort getroffen und wahrgenommen werden sollten, wo sie am besten umgesetzt werden können. Dies bedeutet, dass höhere staatliche Ebenen, wie der Bund und die Länder, nur dann eingreifen sollten, wenn die unteren Ebenen, wie Kommunen oder soziale Träger, nicht in der Lage sind, die Aufgaben eigenständig zu bewältigen. In der frühkindlichen Bildung fördert dieses Prinzip die Selbstverantwortung und Eigeninitiative der Kommunen und freien Träger. Im Verhältnis zwischen freien und öffentlichen Trägern spielt das Subsidiaritätsprinzip eine entscheidende Rolle und definiert den Vorrang freier Träger gegenüber öffentlichen Trägern. Öffentliche Träger, wie die Kommunen, haben die Aufgabe, ein bedarfsgerechtes Angebot an frühkindlicher Bildung und Betreuung zu gewährleisten und die Rahmenbedingungen zu



schaffen, allerdings nur dann, wenn sich das Angebot nicht durch freie Träger sichern lässt. Freie Träger, wie gemeinnützige Organisationen, können flexibler auf lokale Bedürfnisse reagieren.

Die Bildungspläne der einzelnen Bundesländer definieren u. a. Bildungsbereiche und beschreiben Bildungsziele dahin gehend, welche Bildungsbereiche in den frühen Bildungseinrichtungen adressiert werden sollen, z. B. eine alltagsintegrierte Förderung sprachlicher Fähigkeiten, Erfahrungen mit Zahlen, Raum und Formen, das Erkunden von Natur und Naturphänomenen sowie die Förderung sozial-emotionaler Kompetenzen und Merkmale. Die Bildungspläne definieren allerdings bewusst keine konkreten Kompetenzziele, was in Teilen auch als unvereinbar mit dem frühpädagogischen Verständnis gilt. Im frühpädagogischen Verständnis sind eine Kind- und Bedürfnisorientierung leitend sowie ein ganzheitliches Verständnis von Bildung. Damit einhergehend werden (Kompetenz-)Unterschiede als Teil der Individualität des Kindes aufgefasst und eine Förderung spezifischer Domänen oftmals als zu eng gefasst verstanden. Ein ganzheitliches Bildungsverständnis liegt auch dem *Gemeinsamen Rahmen für die frühe Bildung in Kindertageseinrichtungen* der Jugend- und Familienministerkonferenz (JFMK) und der KMK zugrunde (JFMK & KMK, 2022), wie es auf der ersten gemeinsamen Sitzung beider Konferenzen im Oktober 2023 noch einmal bekräftigt wurde. Zugleich wurden die gemeinsame Verantwortung der Minister:innen beider Konferenzen für die Rahmenbedingungen und Grundlagen für einen erfolgreichen Übergang in die Grundschule sowie die gemeinsame Erörterung u. a. der Themen Erfassung von Entwicklungsständen, Datenübertragung, Übergang von der Kita in die Grundschule, basale Kompetenzen, Zusammenarbeit mit Familien und Sozialraum in den Gremien der Ministerkonferenzen beschlossen (JFMK & KMK, 2023).

Vor dem Hintergrund dieser Steuerungslogiken behandelt dieses Kapitel in der Situationsanalyse ebenfalls die Potenziale einer datengestützten Qualitätsentwicklung für die Bereiche Monitoring und Controlling, Programmentwicklung und Ressourcenzuweisung, Weiterentwicklung der Qualität der pädagogischen Angebote in den Einrichtungen sowie Erziehungs- und Bildungspartnerschaften. Der Fokus liegt hierbei auf der ergänzenden Nutzung von Struktur- und Prozessdaten zur Weiterentwicklung der pädagogischen Qualität. Dort, wo Daten zum kindlichen Entwicklungsstand (perspektivisch) verfügbar sind, werden auch die Potenziale der Nutzung dieser Daten diskutiert. Im folgenden Schritt werden die abgeleiteten Maßnahmen beschrieben.

6.1 Situationsanalyse

6.1.1 Monitoring und Controlling

Schaut man auf den Status quo des Systemmonitorings der frühkindlichen Bildung, Betreuung und Erziehung (FBBE), fällt auf, dass umfangreiche Daten zu den Strukturen im Sinne von Inputs und Kontextbedingungen vorliegen (vgl. Kapitel 5.1.3 und Scheerens, 2008), die daher im Folgenden dargestellt werden. Sie definieren die strukturellen Ausgangsbedingungen der frühen Bildung und bilden in Teilen im Sinne eines strukturell-prozessualen Verständnisses pädagogischer Qualität die Strukturqualität ab (vgl. Anders & Oppermann, 2024). In den Bereich der Strukturqualität fallen Aspekte wie Platz und Raumausstattung, Gruppengrößen, Fachkraft-Kind-Relation, der Anteil mittelbarer pädagogischer Arbeitszeit und formal-



fachliche Qualifikationen des Personals. Diese Aspekte werden als Voraussetzung für die pädagogische Prozess- und Interaktionsqualität gesehen, die sich auf Interaktionen zwischen Fachkraft und Kind, zwischen Kindern untereinander und zwischen Kindern in Auseinandersetzung mit Raum und Material beziehen. Daten zum Output im Sinne einer so verstandenen Prozessqualität liegen nur vereinzelt in überregionaler Form vor und stammen zumeist aus Evaluationsstudien. Eine nationale Untersuchung der Prozessqualität in Kindertageseinrichtungen fand zuletzt im Rahmen der *NUBBEK*-Studie vor 15 Jahren statt (Tietze et al., 2013). Daten zu den von den Kindern entwickelten Fähigkeiten, Fertigkeiten und Einstellungen liegen dagegen erst in Ansätzen bzw. im Rahmen von Forschungsprojekten vor, sodass die Entwicklungsmöglichkeiten im Kapitel zu den Maßnahmen dargestellt werden.

Kinder- und Jugendhilfestatistik

Die *Kinder- und Jugendhilfestatistik (KJH-Statistik)* bietet seit den 1950er-Jahren die Basis für die Publikation zentraler Ergebnisse zur Entwicklung der Kindertagesbetreuung. Zwar gab es zwischenzeitlich mehrere weitreichende Änderungen in den Erhebungen, weshalb die früheren Ergebnisse nicht mehr eins zu eins mit den heutigen verglichen werden können. Dennoch lassen sich die Entwicklungen der vergangenen 75 Jahre in ihren zentralen Linien verfolgen (vgl. Rauschenbach & Meiner-Teubner, 2024). Seit 2006 hat sich ein sukzessive weiterentwickeltes, aber hinsichtlich zentraler Grundkennziffern kaum verändertes Erhebungsinstrument etabliert, das jährlich zum Stichtag 1. März Informationen zu allen Kitas, den betreuten Kindern und dort tätigen Personen sowie Kindertagespflegepersonen und die von ihnen betreuten Kinder zusammenträgt. Aufgrund der gesetzlich geregelten Auskunftspflicht handelt es sich um eine Vollerhebung, bei der Angaben zu sämtlichen Einrichtungen und Personen sowie zu allen erfragten Merkmalen gemacht werden müssen. Stichprobenverzerrungen und fehlende Angaben sind im Grunde ausgeschlossen.

In der KJH-Statistik werden verschiedene strukturelle Rahmenbedingungen und Qualitätsparameter erfasst, die für die Analyse und Verbesserung der Qualität und Teilhabe in der frühen Bildung von zentraler Bedeutung sind. Zu den wichtigsten Parametern gehören zunächst die Anzahl und der Anteil der Kinder in der Kindertagesbetreuung, welche durch die Bildungsbeteiligungsquote dargestellt werden und somit aufzeigen, wie viele Kinder in Kitas betreut werden. Darüber hinaus werden Informationen zu den Öffnungszeiten und Schließtagen der Kindertageseinrichtungen erfasst; jene geben Auskunft über die verfügbaren Betreuungszeiten und stellen aus der Perspektive der Eltern einen zentralen Aspekt mit Blick auf die Bedarfsgerechtigkeit und die Vereinbarkeit von Familie und Beruf dar. In diesem Zusammenhang sind auch die vertraglich vereinbarten Betreuungsumfänge zu nennen, die die Details zu den Betreuungsverträgen zwischen Eltern und Einrichtungen enthalten. Zudem werden kindbezogene Merkmale erfasst, darunter die Altersstruktur der betreuten Kinder sowie Informationen über den Migrationshintergrund und die in der Familie vorrangig gesprochene Sprache, etwa Deutsch oder eine andere Sprache. Besonders wichtig ist auch die Erfassung von Kindern, die Eingliederungshilfe nach den Sozialgesetzbüchern VIII und IX erhalten, da dies Hinweise auf spezifische Förderbedarfe liefert. Zusätzlich werden unterschiedliche Parameter erfasst, die als strukturelle Qualitätsmerkmale aufzufassen sind, an erster Stelle die Personalsituation in den Kitas. Hierzu gehören die Anzahl und die Qualifikation des pädagogischen Personals, wobei Informationen zu Abschlüssen und Weiterbildungen erfasst werden. Auch die Altersstruktur sowie der Beschäftigungsumfang des Personals spielen eine wesentliche Rolle. Darüber hinaus werden Angaben zur Anstellung von Leitungs- und Verwaltungspersonal sowie deren zeitlichen Ressourcen für Leitungsaufgaben erfasst, um ein



umfassendes Bild der personellen Ausstattung zu erhalten. Die Arbeitsbedingungen und Stabilität des Personals sind ebenfalls Bestandteile der KJH-Statistik. Hierzu zählen u. a. die Dauer der Tätigkeit in der Einrichtung sowie die Art der Beschäftigung, etwa ob die Anstellung befristet oder unbefristet ist. Schließlich wird auch auf die Förderbedarfe der Kinder eingegangen, wobei Daten zu besonderen Unterstützungsbedarfen, beispielsweise in Bezug auf Sprachförderung oder Eingliederungshilfe, erfasst werden. Es steht dementsprechend hierdurch eine umfassende Datenbasis, insbesondere für Aspekte der Bildungsbeteiligung und Bildungsbedarfe, aber auch zu einzelnen strukturellen Qualitätsparametern, zur Verfügung.

Ein weiteres wichtiges Potenzial der KJH-Statistik liegt in ihrer Rolle für Personalplanung, -gewinnung und -bindung. Sie liefert wesentliche Datengrundlagen für die Erstellung von Personalbedarfsprognosen, indem sie den aktuellen Stand in den Kitas abbildet und bisherige Entwicklungen analysiert. Diese Daten sind entscheidend, da sie in ihrer Detailgenauigkeit nicht durch arbeitsmarktbezogene Daten ersetzt werden können, die mit Wirtschaftszweigen und Berufsklassifikationen arbeiten. Auch die oben dargestellten Informationen zu Arbeitsbedingungen und Qualitätsparametern der Personalsituation können als Einflussfaktoren im Kontext der Personalgewinnung und -bindung nutzbar gemacht werden. Die Ergebnisse werden teilweise durch die Statistischen Ämter des Bundes und der Länder veröffentlicht und fließen in zentrale Berichte wie den Bildungsbericht⁷⁹ oder das Fachkräftebarometer⁸⁰ ein.

Die Analysepotenziale der Einzeldatensätze gelten allerdings noch nicht als ausgeschöpft (vgl. Kapitel 5.1.12). Einzelne wichtige Parameter, wie der Bezug von Transferleistungen, sind nicht enthalten. Eine weitere zentrale Einschränkung betrifft die fehlende Verknüpfung mit Trägerdaten, außerdem schränken datenschutzrechtliche Bestimmungen die Nutzbarkeit ein. Dies führt häufig dazu, dass einige Ergebnisse, die für die Beschreibung der Quantität, der Strukturqualität sowie der personalbezogenen Entwicklungen grundlegend sind, insbesondere auf regionaler Ebene nicht zur Verfügung stehen. Ein besonderes Augenmerk gilt hierbei der seit vielen Jahren von Wissenschaftler:innen geforderten Anwendung der Cell-Key-Methode, deren Entwicklung durch die Statistischen Ämter jedoch noch aussteht. Wann welche amtlichen Statistiken Wissenschaftler:innen für Analysen zur Verfügung stehen, ist nach wie vor offen (Enderle & Vollmar, 2019; Setzer et al., 2024).

Erhebungen des Deutschen Jugendinstituts

Das Deutsche Jugendinstitut (DJI) ist ein sozialwissenschaftliches Forschungsinstitut, das an der Schnittstelle von Wissenschaft, Politik und Fachpraxis arbeitet. Es erhebt regelmäßig Daten zur frühen Bildung, Betreuung und Erziehung, die in die Sozial- und Bildungsberichterstattung einfließen. Ein zentrales Themenfeld, das seit mindestens 20 Jahren intensiv bearbeitet wird, ist die frühe Kindheit, die das Aufwachsen von Kindern ab der Geburt in relevanten Kontexten wie Familie und Kindertagesbetreuung umfasst. Aufgrund ihrer hohen Relevanz für die Politikberatung haben aktuelle gesellschaftliche Entwicklungen und Alltagsphänomene eine hohe Priorität. Dies zeigte sich besonders während der Coronapandemie, als die

⁷⁹ <https://www.bildungsbericht.de/de> (aufgerufen am 20.02.2026).

⁸⁰ <https://www.fachkraeftebarometer.de/> (aufgerufen am 20.02.2026).



*Corona-KiTa-Studie*⁸¹ in wöchentlichen Abständen aktuelle Daten zur Situation in Bildungseinrichtungen bereitstellte.

Zu den wichtigsten regelmäßigen Datenerhebungen am DJI gehören zunächst der Survey *Aufwachsen in Deutschland: Alltagswelten (AID:A)*⁸², der Daten über das Aufwachsen von Kindern, Jugendlichen und jungen Erwachsenen in Deutschland erhebt. Die bundesweit repräsentativen Daten sind ein wesentlicher Bestandteil der Sozialberichterstattung. Die erste Längsschnittstudie begann 2009 und wurde bis 2024 fortgeführt. Ein zentraler Aspekt ist die Berücksichtigung des größeren sozialen Kontextes, in dem die Befragten leben, sowie der gesellschaftlichen Rahmenbedingungen des Aufwachsens. Zudem wird der elterliche Bedarf an, die Nutzung von und das Wissen über Kinderbetreuung erfasst. AID:A analysiert auch Ungleichheiten in der Inanspruchnahme außerfamiliärer Betreuung und anderer sozialstaatlicher Leistungen, was einen wertvollen Beitrag zur Identifikation von Ungleichheiten in der frühkindlichen Bildung leistet.

Ein weiterer bedeutender Datensatz ist die *DJI-Kinderbetreuungsstudie (KiBS)*⁸³, die die Perspektive der Eltern auf die frühkindliche Bildung und Betreuung untersucht. KiBS liefert jährlich aktualisierte Kennzahlen für verschiedene Berichtsformate der Sozialberichterstattung und ermöglicht die Beobachtung relevanter Trends über die Jahre hinweg. Jährlich werden etwa 33.000 Eltern aus ganz Deutschland befragt, wobei die Daten für verschiedene Altersgruppen der Kinder repräsentativ sind. Zusätzlich werden im Projekt *Entwicklung von Rahmenbedingungen in der Kindertagesbetreuung (ERiK)*⁸⁴ die Qualität der Angebote sowie die Rahmenbedingungen in der Kindertagesbetreuung untersucht. Basierend auf einer bundesweiten Zufallsstichprobe werden Befragungen von pädagogischen Fachkräften, Leitungskräften und anderen relevanten Akteur:innen durchgeführt. Die Erhebungen fanden 2020, 2022 und 2024 statt, eine weitere Erhebung ist für 2026 vorgesehen. Es wird die Zusammensetzung der Kindergruppen in den Einrichtungen, insbesondere hinsichtlich Migrationshintergrund, Entwicklungsverzögerungen oder niedrigem sozioökonomischen Status erfasst. Darüber hinaus werden die Bedürfnisse und die Situation der pädagogischen Fachkräfte untersucht, um deren Arbeitsbedingungen und Herausforderungen besser zu verstehen. Ein besonderer Mehrwert liegt in der Perspektive der Fach- und Leitungskräfte auf ihre Arbeitssituation und die Qualität in den Einrichtungen, die über die Kinder- und Jugendhilfestatistik weit hinausgeht.

Die Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF) am DJI gibt das *Fachkräftebarometer Frühe Bildung (FKB)*⁸⁵ heraus. In diesen seit 2014 regelmäßig erscheinenden Datenberichten werden umfassende Informationen über Personal, Arbeitsmarkt, Erwerbssituation sowie Ausbildung und Qualifizierung in der Frühpädagogik und im Ganztags für Grundschulkindern auf Basis amtlicher Daten zusammengetragen.

Zusätzlich führt das DJI anlassbezogene Erhebungen zu spezifischen Themen durch. Die überwiegende Mehrheit der Daten des DJI ist über das Forschungsdatenzentrum (FDZ-DJI) für die wissenschaftliche Nachnutzung verfügbar, wobei der Fokus auf repräsentativen Studien mit hohem Nachnutzungspotenzial liegt.

⁸¹ www.dji.de/corona-kita (aufgerufen am 20.02.2026).

⁸² www.dji.de/aida/gesamtbeschreibung (aufgerufen am 20.02.2026).

⁸³ www.dji.de/KiBS (aufgerufen am 20.02.2026).

⁸⁴ www.dji.de/ERiK (aufgerufen am 20.02.2026).

⁸⁵ <https://www.fachkraeftebarometer.de/ueber-das-fkb> (aufgerufen am 20.02.2026).



TALIS-Starting-Strong-Studie

Ein internationaler Vergleich wird durch die OECD-Studie *Teaching and Learning International Survey (TALIS) Starting Strong* ermöglicht, die zentrale Aspekte der frühkindlichen Bildung und Betreuung untersucht (OECD, 2019b, 2025). TALIS Starting Strong konzentriert sich auf die Arbeitsbedingungen von Fachkräften in Kindertageseinrichtungen und untersucht deren berufliche Rollen sowie die angewandten pädagogischen Praktiken. Ein weiterer wichtiger Aspekt ist das professionelle Selbstverständnis der Fachkräfte. Die Studie ermöglicht den Vergleich der erfassten Daten mit denen anderer teilnehmender Länder, wodurch Best Practices identifiziert und Herausforderungen im Bereich der frühkindlichen Bildung analysiert werden können. Die erste Erhebungsrunde fand 2018 statt, mit der erneuten Durchführung 2024 wurden auch deutsche Einrichtungen und Fachkräfte befragt.

Bertelsmann-Ländermonitoring

Das *Bertelsmann-Ländermonitoring* (Bock-Famulla et al., 2024) ist eine wichtige Erhebung zur Qualität der frühkindlichen Bildung in Deutschland, die von der Bertelsmann Stiftung initiiert wurde, um die Bildungslandschaft im Vorschulbereich zu analysieren und zu verbessern. Das Monitoring beruht auf einer umfassenden Datensammlung, die sowohl quantitative als auch qualitative Aspekte der frühkindlichen Bildung erfasst. Die Daten stammen aus verschiedenen Quellen: amtliche Statistiken, Erhebungen von Bildungseinrichtungen und Umfragen unter Fachkräften sowie Eltern. Diese Daten umfassen u. a. Informationen zu strukturellen Merkmalen wie der Anzahl der Fachkräfte, dem Personal-Kind-Verhältnis, den finanziellen Rahmenbedingungen sowie den räumlichen Gegebenheiten der Kindertageseinrichtungen. Darüber hinaus werden qualitative Aspekte wie die pädagogischen Konzepte, die Qualität der Interaktionen zwischen Fachkräften und Kindern sowie die Zufriedenheit der Eltern mit den Betreuungsangeboten erfasst. Der Zweck des Bertelsmann-Ländermonitorings besteht darin, einen umfassenden Überblick über die Qualität der frühkindlichen Bildung in den einzelnen Ländern zu gewinnen. Dies ermöglicht Vergleiche zwischen den Ländern und die Identifizierung potenzieller Ungleichheiten in der Bereitstellung von Bildungsressourcen und -angeboten.

Qualitätsentwicklung auf Systemebene

Landesprogramme sind ein Weg, in Ländern bestimmte Qualitätsentwicklungsbereiche gezielt zu adressieren. Entsprechend dem aktuell identifizierten Bedarf an einer frühen sprachlichen Bildung berichtet die weit überwiegende Zahl der Länder (13) in der *SWK-Länderabfrage Frühe Bildung* (vgl. Anhang 1) von Landesprogrammen zur sprachlichen Bildung und Förderung im Bereich der Frühen Bildung. Spezifische Landesprogramme zu anderen Kompetenz- und Anforderungsbereichen wie frühe mathematische Kompetenzen (4 Länder), sozial-emotionale Fähigkeiten (4 Länder) oder dem Übergang von der Kita in die Grundschule (9 Länder) werden nach Angaben der Länder nur in einigen Ländern umgesetzt. All diese Themen sind jedoch in den Bildungsplänen der Länder als Bildungsbereiche regulär verankert. Sechs Länder geben an, sich bei der Bereitstellung zusätzlicher Programme auf Monitoring-Daten zu stützen, fünf Länder nennen die Kinder- und Jugendhilfestatistik als eine Basis der Ressourcenzuweisung.

Die Daten, die durch das DJI zur Verfügung gestellt werden, liefern ebenfalls wichtige Informationen für Steuerungszwecke. Sie geben Einblick in Elternbedarfe und die Fachkräftesituation und dokumentieren



regionale Ungleichheiten. Sie schaffen hierdurch eine Transparenz für politische Entscheidungsträger:innen, Träger, Fachkräfte, Eltern und Familien. Die Berichterstattung fließt hierbei auch aktuell bereits in die politische Entscheidungsfindung auf unterschiedlichen Handlungsebenen ein. Sie kann auch auf lokaler Ebene als Vergleichsgröße herangezogen werden. Das Monitoring hat großes Potenzial, Handlungsbedarf zu identifizieren und Handlungsempfehlungen für Praxis und Politik zu formulieren, um die Qualität der frühkindlichen Bildung im Sinne einer datengestützten Entwicklung und Steuerung der Bildungspolitik auf Bundes- und Landesebene flächendeckend zu verbessern. In ähnlicher Weise ist die Rolle des Bertelsmann Ländermonitorings zu bewerten. Es liefert umfassende und aktuelle Daten zur Qualität der frühkindlichen Bildung in den Ländern, die entscheidend für die politische Entscheidungsfindung sein können. Durch die Identifikation von Ungleichheiten zwischen und Qualitätsentwicklungsbedarfen in spezifischen Regionen ermöglicht das Monitoring eine gezielte Verteilung finanzieller Mittel und Unterstützungsangebote dorthin, wo sie am dringendsten benötigt werden. Die länderspezifischen Analysen erleichtern auch den Austausch von Best Practices und können zur Verbesserung der Bildungsangebote beitragen.

Blickt man allerdings auf konkrete bundes- und länderspezifische Entscheidungen in der Entwicklung der Qualität der Kindertageseinrichtungen, so wird deutlich, dass die Entscheidungen für Priorisierungen und die Analyse der Bedarfe nicht immer datengestützt erfolgen. So konnten im Rahmen des Gute-Kita-Gesetzes alle Länder Maßnahmen zur Weiterentwicklung der Qualität ergreifen und hatten hierfür unterschiedliche Handlungsfelder zur Auswahl. Die Möglichkeit der unterschiedlichen Schwerpunktsetzung in den Ländern war explizit gewollt, um auf die regionalen Gegebenheiten und Bedarfe sowie die regional unterschiedlichen Ausgangssituationen reagieren zu können. Die Evaluation des Gute-Kita-Gesetzes zeigte jedoch, dass einige – aber nicht alle – Länder sich an den tatsächlichen Entwicklungsbedarfen in ihrem Land orientierten (Rönnau-Böse et al., 2023).

Die Machbarkeitsstudie zur Ansprache und Umsetzung eines bundesweiten Monitorings der Prozessqualität in der Frühen Bildung (Kluczniok, Faas et al., 2024) zeigt, dass die derzeitigen Aktivitäten zur Erfassung der Prozess- und Interaktionsqualität in Deutschland unzureichend sind. Diese Erhebungen liegen überwiegend in der Eigenverantwortung der Träger, was zu einer uneinheitlichen Datenlage und fehlender Vergleichbarkeit führt. Vorhandene Monitorings konzentrieren sich hauptsächlich auf strukturbezogene Merkmale und erfassen prozessbezogene Qualitätsmerkmale nur sporadisch. Einzelne Länder führen projektbezogene Erhebungen durch, die jedoch primär für die Qualitätsentwicklung in den Einrichtungen selbst genutzt werden statt für übergeordnete Steuerungsfragen. Die Auswertung vorhandener Qualitätsdaten zeigt ein mittleres Qualitätsniveau hinsichtlich globaler Prozessmerkmale, während spezifische Bereiche wie Literacy und Numeracy erhebliche Entwicklungsbedarfe aufweisen.

6.1.2 Organisationsentwicklung und Ressourcenzuweisung

Für die Qualitätsentwicklung auf Ebene der einzelnen Einrichtungen und Träger und die gezielte Vergabe von Ressourcen wie Stellenmitteln, der fachlichen Unterstützung oder Fort- und Weiterbildungen stehen aktuell trägerspezifische Datenstrukturen sowie Daten aus internen und externen Evaluationen und den bereits etablierten oder im Aufbau begriffenen Sprachstandserhebungen zur Verfügung.



Interne und externe Evaluationen

Ein wichtiges Instrument der Qualitätsentwicklung und des Controllings pädagogischer Qualität auf Ebene der Träger und einzelner Einrichtungen sind externe und interne Evaluationen. In Deutschland variieren diesbezügliche Regelungen erheblich je nach Land. Interne Evaluationen sind in den meisten Ländern verpflichtend, in einigen Ländern werden sie nur empfohlen (Kluczniok, Faas et al., 2024; vgl. Abbildung 13). Die konkrete Umsetzung und auch die Transparenz für Eltern und Familien variieren allerdings oft zwischen Trägern, die für sich spezifische Qualitätsmanagement- und Qualitätsentwicklungssysteme geschaffen haben. Ebenso variieren die Evaluationssysteme und die Validität der eingesetzten Instrumente.

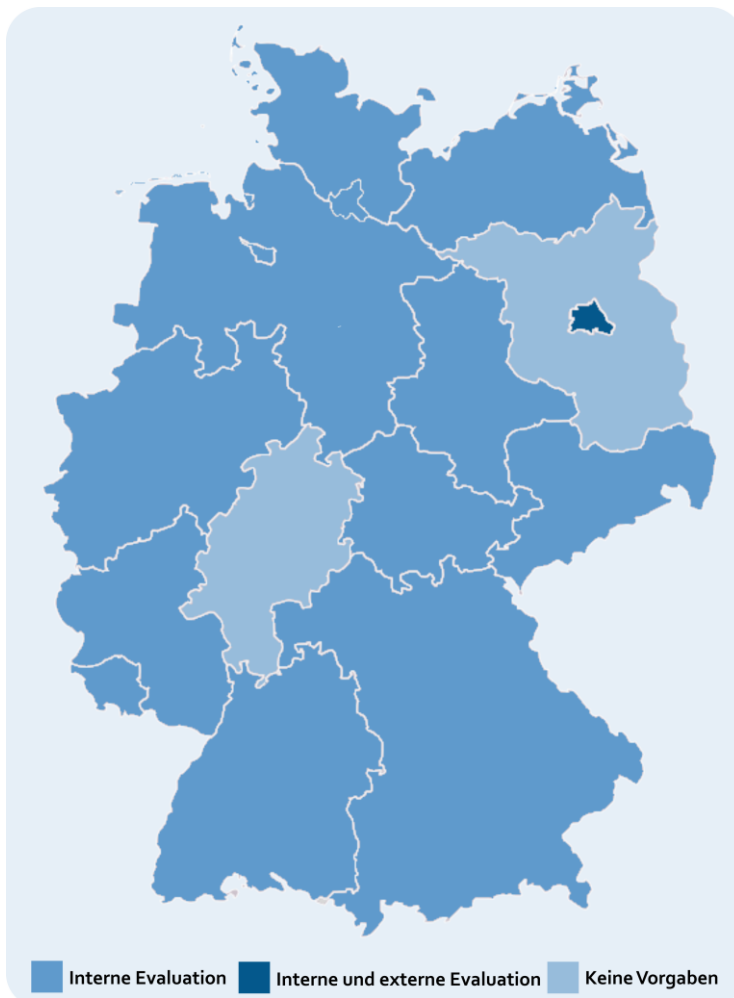


Abbildung 13: Regelungen der Länder zu internen und externen Evaluationen der pädagogischen Qualität der Kindertageseinrichtungen

Anmerkung: In Hamburg wird die Einhaltung struktureller Standards im Rahmen des Kita-Prüfverfahrens zentral extern überprüft.

Quelle: Kluczniok et al., 2024; § 6 SBEBG⁸⁶

⁸⁶ [Bürgerservice Saarland – § 6 SBEBG | Landesnorm Saarland | Qualitätssicherung und -entwicklung | § 6 – Qualitätssicherung und -entwicklung](#) (aufgerufen am 20.02.2026).



Für externe Evaluationen gibt es häufig Empfehlungen, die jedoch nicht gesetzlich verbindlich sind und oft keine Vorgaben zur Häufigkeit enthalten. Die externen Evaluationen werden von Fachberatungen oder spezialisierten Organisationen durchgeführt. Die Ergebnisse der Evaluationen konzentrieren sich in der Regel auf die Verbesserung der pädagogischen Qualität und die Unterstützung der Qualitätsentwicklung in den Kitas. In Berlin ist die externe Evaluation gesetzlich verpflichtend, interne Evaluationen sind ebenfalls Teil des Qualitätsmanagements. Die Koordination der externen Evaluation erfolgt durch das Berliner Kita-Institut für Qualitätsentwicklung (BeKi). Die Ergebnisse dieser Evaluationen sind verbindlich und fließen in die Qualitätsentwicklung sowie in die Öffentlichkeitsarbeit der Kitas ein.

Erhebungen des Sprachstands und weiterer Entwicklungsbereiche

Die meisten Länder haben spezifische Verfahren implementiert, um den Sprachstand aller Kinder oder bestimmter Gruppen zu erfassen. Die verbleibenden Länder planen in naher Zukunft Verfahren für landesweite Sprachstandserhebungen einzuführen. Dabei ist sowohl zwischen den Ländern als teilweise auch innerhalb der Länder eine Vielzahl an verschiedenen Verfahren im Einsatz, von denen ein Großteil ein bis zwei Jahre vor der Einschulung durchgeführt wird (vgl. Abbildung 14). Als besondere Herausforderung gilt hier, auch Kinder zu erreichen, die keine Kita besuchen.

Über die Sprachstandserhebungen hinaus werden nur in einzelnen Ländern noch weitere Entwicklungsdaten erhoben wie sozial-emotionale Fähigkeiten (Berlin, Bayern, Hamburg), mathematische (Berlin, Hamburg), naturwissenschaftliche (Berlin) oder motorische Kompetenzen (Berlin, Hamburg). Andere Länder bereiten weitere Erhebungen vor.⁸⁷ Die Einschulungs-/Schuleingangsuntersuchungen sowie die Vorsorgeuntersuchungen bei Kinderärzt:innen (v.a. U7 bis U9) in Verantwortung des Gesundheitssystems sind hier nicht systematisch berücksichtigt, da Gesundheitsdaten als besonders sensibel gelten und häufig kein systematischer Datenaustausch mit dem Bildungssystem bzw. dem System der FBBE vorgesehen ist. Weiterentwicklungen sind in einzelnen Ländern in der Diskussion. In Hessen wird das Sprachscreening aus den Schuleingangsuntersuchungen allerdings bereits für die Zuweisung von Vorlaufkursen zur Sprachförderung genutzt,⁸⁸ und beispielsweise in Baden-Württemberg ist dies in Vorbereitung.⁸⁹

⁸⁷ Bayern: [Beobachtungsbögen – Staatsinstitut für Frühpädagogik und Medienkompetenz \(IFP\)](#); Berlin: [BeoKiz-Verfahren – Berlin.de](#); Hamburg: [VVV 2025 Bogen A lfbQ](#) (aufgerufen am 20.02.2026).

⁸⁸ <https://schulaemter.hessen.de/schulbesuch/einschulung> (aufgerufen am 20.02.2026).

⁸⁹ <https://km.baden-wuerttemberg.de/de/schule/grundschule/sprachfit> (aufgerufen am 20.02.2026).

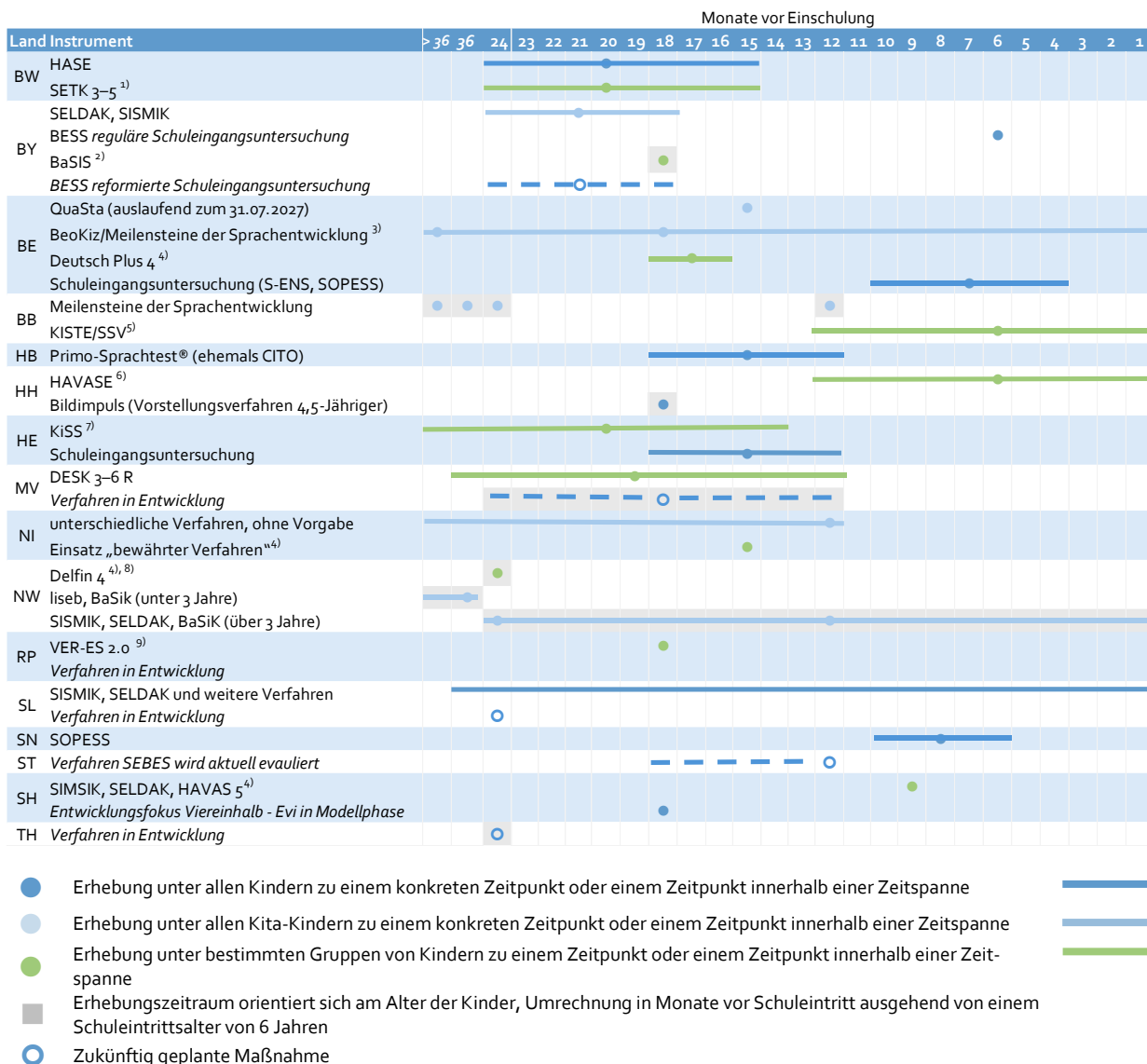


Abbildung 14: Sprachstandserhebung im Vorschulalter in den Ländern

Abkürzungen: BaSiS: Bayrisches Screening des individuellen Sprachstands | BESS: Bayrisches Einschulungssprachscreening | CITO: Centraal Instituut Toets Ontwikkeling | DESK 3-6 R: Dortmunder Entwicklungsscreening für den Kindergarten | HASE: Heidelberger Auditives Screening in der Einschulungsuntersuchung | HAVASE: Hamburger Verfahren zur Analyse der Sprachentwicklung | HAVAS-5: Hamburger Verfahren zur Analyse des Sprachstands bei 5-Jährigen | KiSS: Kindersprachscreening | KISTE: Kindersprachtest für das Vorschulalter (auslaufendes Verfahren zum 31.07.2027) | liseb – Literacy und Sprachentwicklung beobachten | QuaSta: Qualifizierte Statuserhebung Sprachentwicklung 4-jähriger Kinder in Kitas | SEBES: Schuleingangsbezogenes Entwicklungsscreening (SEBES), die Verwendung zur Sprachstandserhebung wird zurzeit evaluiert | SELDAK: Sprachentwicklung und Literacy bei deutschsprachig aufwachsenden Kindern | SELSA: Sprachentwicklung und Literacy bei Kindern im Schulalter (1.–4. Klasse) | SETK 3-5: Sprachentwicklungstest für 3-5-jährige Kinder | SISMIK: Sprachverhalten und Interesse an Sprache bei Migrantenkindern im Kindergarten | SOPESS: Sozialpädiatrisches Entwicklungsscreening für Schuleingangsuntersuchungen | SSV: Sprachscreening im Vorschulalter (ersetzt KISTE zum 01.08.2027) | VER-ES: Verfahren zur Einschätzung des Sprachförderbedarfs von Kindern im Jahr vor der Einschulung

Anmerkung: Schuleingangsuntersuchungen sind nicht systematisch berücksichtigt.

- 1) Nur für Kinder mit auffälligem Befund im HASE | 2) Für alle ca. 4,5-jährigen Kinder im vorletzten Kitajahr, ausgenommen Kinder, die nach SISMIK/SELDAK keinen Förderbedarf zeigen, und Kinder, die eine schulvorbereitende Einrichtung oder heilpädagogische Tagesstätte besuchen | 3) BeoKiz findet während der gesamten Kita-Zeit Anwendung, in diesem Rahmen finden die Meilensteine der Sprachentwicklung bei Kindern im Alter von 2,5 und 4,5 Jahren Anwendung, das entspricht in etwa 42 bzw. 18 Monaten vor Schuleintritt. | 4) Für Kinder, die keine Kita besuchen | 5) Bei auffälligem Befund im Rahmen der Beobachtung Meilensteine der Sprachentwicklung | 6) Bei festgestelltem Sprachförderbedarf oder Verdacht darauf | 7) Die Teilnahme ist freiwillig. | 8) Für Kinder, deren Sorgeberechtigte der Bildungs- und Entwicklungsdokumentation nicht zugestimmt haben | 9) Im Rahmen der Schulanmeldung von Kindern, die keine Kita besuchen.

Quelle: Autor:innengruppe Bildungsberichterstattung, 2024, Tabelle C5-4web und eigene Recherchen



Zusammenfassend werden in den Ländern eine Vielzahl verschiedener Erhebungsinstrumente eingesetzt, die teilweise auch innerhalb der Länder variieren und nicht alle validiert sind. Nicht immer ist sichergestellt, dass alle Kinder eines Jahrgangs an den Erhebungen zum Sprachstand und zu weiteren Entwicklungsmaßen teilnehmen. Damit wird das Potenzial dieser Daten für die Qualitätsentwicklung und Ressourcensteuerung eingeschränkt. Für eine Verbesserung der pädagogischen Qualität für die Kinder ist vor allem auch eine verbindliche Regelung der an die Entwicklungsstandserhebungen anschließenden Förder- und Bildungsprozesse zentral. Hierfür kommen bei Vorliegen ergänzender Förderbedarfe grundsätzlich sowohl intensivierete alltagsintegrierte Bildungsangebote als auch ergänzende Förderansätze infrage.

Qualitätsentwicklung auf Einrichtungsebene

Evaluationssysteme können nicht nur dem Monitoring und Controlling dienen, sondern auch eine wichtige Rolle in der Weiterentwicklung der pädagogischen Qualität auf Ebene der Träger und Einrichtungen spielen, da hierdurch die Bedarfe an fachlicher Unterstützung und Fort- und Weiterbildung identifiziert und aufgegriffen sowie Schwerpunkte in der pädagogischen Arbeit gesetzt werden können.

Auf Träger- und Einrichtungsebene bilden fachliche Unterstützungsangebote sowie Fort- und Weiterbildungen Kernelemente der Weiterentwicklung der pädagogischen Prozess- und Interaktionsqualität. Die Fachberatungssysteme nehmen eine zentrale Rolle ein. Die Ergebnisse aus internen und externen Evaluationssystemen sowie weiteren Daten können hier zur Steuerung der Unterstützungsangebote und zur inhaltlichen Weiterentwicklung genutzt werden. Eine besondere Rolle spielen die internen Evaluationssysteme, die in vielen Ländern – wie oben dargestellt – grundsätzlich verbindlich vorgeschrieben sind.

Unterschiedliche Studien, die die pädagogische Prozess- und Interaktionsqualität untersucht haben, haben immer wieder eine moderate globale Qualität festgestellt. Die bereichs- und domänenspezifische Qualität im Sinne einer hohen kognitiven Aktivierung weist aber oftmals einen großen Weiterentwicklungsbedarf auf, und gerade in Kitagruppen mit einem hohen Anteil von Kindern mit Zuwanderungsgeschichte ist seltener eine hohe Prozessqualität anzutreffen (vgl. Anders & Roßbach, 2019; Kluczniok, Faas et al., 2024). Weiterhin zeigt sich, dass Lern- und Bildungsangebote sich oftmals zu wenig am kindlichen Entwicklungsstand und ggf. an einem ergänzenden Förderbedarf orientieren. Ein Angebot an Lern- und Entwicklungsangeboten, das sich im Sinne einer Förderung in der Zone der nächsten Entwicklung an dem aktuellen Entwicklungsstand der Kinder orientiert, ist für positive Wirkungen auf die kindliche Entwicklung zentral. Dementsprechend entstehen auch auf organisatorischer Ebene besondere fachliche Unterstützungsbedarfe, wenn der Anteil an Kindern mit besonderem Förderbedarf besonders hoch ist. Daten aus Entwicklungsstandserhebungen können hier neben sozioökonomischen Indikatoren wie dem Anteil armutsgefährdeter Kinder und Familien einen Mehrwert bieten.

In einigen Ländern werden gemäß SWK-Länderabfrage Frühe Bildung die Daten aus Sprachstandserhebungen bereits explizit zur Steuerung und Qualitätsentwicklung genutzt. Hierdurch findet zumeist eine Verbesserung der Strukturqualität statt, die dann in die Weiterentwicklung der Prozessqualität fließen soll. In Bremen sind die Daten Teil des *KiTa-Sozialindex*, über den zusätzliche Ressourcen für Qualitätsentwicklungsprozesse (z. B. Funktionsstellen) zugewiesen werden. In Hamburg werden die Ergebnisse des Vorstellungsverfahrens für Viereinhalbjährige jährlich ausgewertet und den zuständigen Ämtern zur Verfügung



gestellt. In Brandenburg wird u. a. ein Sozialindex auf Basis der Schuleingangsuntersuchungen zur Steuerung des Programms *Kompensatorische Sprachförderung im Jahr vor der Einschulung* sowie weiterer Programme wie der Fachberatung mit dem Schwerpunkt Sprachen genutzt. Häufiger erfolgt die Zuteilung zusätzlicher Personalstellenanteile oder finanzieller Ressourcen pro Kind bzw. nach Anteil von Kindern mit nichtdeutscher Familiensprache oder aus Armutslagen⁹⁰ oder für Kitas mit einem hohen Anteil an Kindern mit besonderen Lernschwierigkeiten⁹¹. Daneben werden auf Ebene des Einzelkindes Fördermaßnahmen eingeleitet (vgl. Kapitel 5.1.3). Trotz dieser Fortschritte stehen die Länder vor Herausforderungen hinsichtlich der Umsetzung und Verknüpfung von Daten. Datenschutzrechtliche Fragen bei der Erhebung und Weitergabe von Daten sowie die Notwendigkeit einer verständlichen und adressatengerechten Darstellung der gesammelten Daten sind zentrale Themen. Zudem wird die Professionalisierung der Fachkräfte zur zielgerichteten Nutzung der Daten als entscheidend angesehen.

6.1.3 Individuelle Planung von Bildungs- und Förderangeboten sowie Übergang in die Grundschule

Eine kursorische Sichtung der Bildungs- und Orientierungspläne der Länder⁹² sowie der landesrechtlichen Regelungen⁹³ zeigt, dass auf vielfältige Weise Daten zur pädagogischen Arbeit und Entwicklungsdaten der Kinder erhoben werden. Häufig wird auf mehr oder weniger systematische oder standardisierte Beobachtungen sowie auf Dokumentationen und Portfolios mit Fotos, Lerngeschichten und Werken der Kinder als Gesprächsgrundlage verwiesen. Vereinzelt werden explizit bestimmte Beobachtungsbögen genannt, oder es wird die Erfüllung der Gütekriterien (Objektivität, Reliabilität und Validität) gefordert. Durchgehend wird deutlich, dass die Entwicklungsdokumentation keinem Selbstzweck unterliegt, sondern primär dazu genutzt werden soll, pädagogische Angebote individuell zu planen und ggf. auch zusätzliche Bildungs- und Fördermaßnahmen einzuleiten. Teilweise wird im Sinne eines zirkulären Vorgehens auch auf die Überprüfung der Wirksamkeit der jeweiligen Maßnahmen und ggf. die Anpassung verwiesen (vgl. z. B. Bayern, Baden-Württemberg, Mecklenburg-Vorpommern).

Eine Reihe von Ländern geben auch in der SWK-Länderabfrage Frühe Bildung an, die individuellen Bildungs- und Förderangebote für die Kinder an den Befunden auszurichten, sei es bei der alltagsintegrierten sprachlichen Bildung und Förderung, sei es durch ergänzende Fördermaßnahmen, die in der Kita oder in der Grundschule stattfinden bzw. perspektivisch in den nächsten ein bis zwei Jahren flächendeckend eingeführt werden sollen. In den existierenden Studien zur pädagogischen Prozessqualität zeigt sich jedoch immer wieder, dass eine Anpassung der pädagogischen Angebote an den Entwicklungsstand der Kinder und

⁹⁰ Vgl. z. B. Bayern: <https://www.gesetze-bayern.de/Content/Document/BayKiBiG-21>; Berlin: <https://gesetze.berlin.de/bsbe/document/jlr-KitaF%C3%B6GVBEV4P17>; Hamburg: <https://www.hamburg.de/resource/blob/35340/7340ef6a2002e74726beb978f47cbo8a/landesrahmen-vertrag-neu-data.pdf>; Niedersachsen: <https://bildungsportal-niedersachsen.de/fruehkindliche-bildung/finanzhilfe-foerderprogramme/finanzhilfe/besondere-finanzhilfe-zur-sprachfoerderung> (aufgerufen am 20.02.2026).

⁹¹ Vgl. z. B. Hessen: https://laghessen.de/wp-content/uploads/2019/01/5_2016_kifoeq_gesamt_final.pdf; Nordrhein-Westfalen: <https://www.mkjfgfi.nrw/system/files/media/document/file/kibiz-mit-stand-vom-01.08.2022.pdf>; Schleswig-Holstein: <https://www.gesetze-rechtsprechung.sh.juris.de/bssh/document/jlr-KTagStGSH2020V25P16b>; Sachsen: <https://www.kinder-staerken-sachsen.de/esf-plus-programm> (aufgerufen am 20.02.2026).

⁹² Vgl. <https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/landesrechtliche-regelungen/personal-und-einrichtungen/bildungsplaene-1-1> und einzelne aktuellere Fassungen (aufgerufen am 20.02.2026).

⁹³ Vgl. <https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/landesrechtliche-regelungen/personal-und-einrichtungen/uebersicht-der-regelungen-2> und einzelne aktuellere Fassungen (aufgerufen am 20.02.2026).



eine Förderung in der Zone der nächsten Entwicklung selten stattfinden (Anders & Oppermann, 2024). Auch scheint es eine Überforderung zu sein, die zum Teil umfangreichen Daten aus Beobachtungssystemen für die Nutzung in der pädagogischen Arbeit mit dem einzelnen Kind zu übersetzen. Darüber hinaus werden nicht alle Kinder mit besonderem Förderbedarf identifiziert und erhalten ergänzende Förderung. Die Rolle der vielfältigen beobachtungsbasierten Dokumentationsinstrumente für die Weiterentwicklung der pädagogischen Angebote sowie die Einleitung ergänzender Fördermaßnahmen bleibt dabei häufig unklar und unverbindlich. Gerade dadurch, dass das Potenzial für die pädagogische Arbeit nicht ausgenutzt wird, können Beobachtung und Dokumentation aus Sicht der Fachkräfte eine zusätzliche Belastung und keinen arbeitserleichternden Mehrwert darstellen. Diese Aspekte werden aktuell insbesondere für die Weiterentwicklung der Qualität der sprachlichen Bildung und Förderung diskutiert, da die Anzahl der Kinder mit Sprachförderbedarf aufgrund gesellschaftlicher Entwicklungen ansteigt und etwa 20 bis 30 Prozent aller Kinder am Ende der Grundschulzeit die Mindeststandards im Lesen und Schreiben nicht erreichen (Wittig & Schneider, 2022).

Viele der gesichteten Rahmen- und Orientierungspläne thematisieren explizit, dass die in der Kita dokumentierten Entwicklungsdaten im Sinne anschlussfähiger Bildungsprozesse auch an die aufnehmende Grundschule (z. B. Berlin, Baden-Württemberg, Nordrhein-Westfalen, Schleswig-Holstein, Thüringen) und teilweise auch den aufnehmenden Hort (z. B. Mecklenburg-Vorpommern) weitergegeben werden sollen oder dürfen, und auch KMK und JFMK sehen den Austausch individueller Daten als Teil einer guten Kommunikation zwischen frühkindlicher Bildung und Primarbereich. An Ressourcen und Kompetenzen der Kinder aus der frühkindlichen Bildung solle angeknüpft und auf bereits erkannte Förderbedarfe in der Schule ohne Brüche eingegangen werden (Beschluss JFMK & KMK, 13.10.2023). In der Regel wird dafür das Einverständnis der Eltern gefordert, oder die Weitergabe der Entwicklungsdokumentation erfolgt direkt durch die Eltern. Teilweise findet auch ein direkter Austausch zwischen Erzieher:innen und Grundschullehrkräften zu den Daten statt (vgl. z. B. das Hamburger Vorstellungsverfahren für Viereinhalbjährige). Eine Weitergabe von Entwicklungsdaten an die Grundschule wird aber auch sehr kritisch gesehen und ist mit Ängsten vor Stigmatisierung der Kinder verbunden.

6.1.4 Erziehungs- und Bildungspartnerschaften

Bildungs- und Erziehungspartnerschaft bezeichnet im frühkindlichen Bereich (zum schulischen Kontext vgl. Kapitel 4) eine strukturell verankerte, dialogisch gestaltete Zusammenarbeit zwischen Kindertageseinrichtung und Familie, die auf offenen Austausch, gegenseitige Anerkennung von Expertise und geteilte Verantwortung für Erziehung, Bildung und Betreuung zielt und die gemeinsame Förderung des Kindes unterstützt (Deutsches Jugendinstitut, 2011; Friederich, 2011). Nach Stange zeichnen sich Erziehungs- und Bildungspartnerschaften auch dadurch aus, dass Eltern, Kindertageseinrichtungen, Schulen, Jugendhilfe und weitere Akteure im Sozialraum systematisch und verbindlich zusammenarbeiten (Stange, 2012).



In den landesrechtlichen Regelungen zur Kindertagesbetreuung⁹⁴ und/oder den landesweiten Bildungs-, Erziehungs- oder Rahmenplänen⁹⁵ ist im Sinne einer solchen Erziehungs- und Bildungspartnerschaft ein regelmäßiger Austausch zwischen den pädagogischen Fachkräften und den Erziehungsberechtigten verankert. In der Regel sollen mindestens jährlich oder zweimal jährlich gemeinsame Entwicklungsgespräche stattfinden; bei jüngeren Kindern wird teilweise eine häufigere Durchführung empfohlen. Die Länder unterscheiden sich darin, wie explizit dabei systematische Daten zum Entwicklungsstand der Kinder zum Gegenstand des Gesprächs gemacht werden müssen oder sollen. Die Sprachentwicklung wird am ehesten als thematisierbarer Kompetenzbereich bezeichnet. In Baden-Württemberg müssen die Kindertagesstätten den Eltern ein Entwicklungsgespräch anbieten, wenn Befunde aus den Schuleingangsuntersuchungen in mindestens einem Kompetenzbereich einen intensiven Förderbedarf aufzeigen. Zur tatsächlichen Häufigkeit und erlebten Qualität von Angeboten der Elternzusammenarbeit liegen inzwischen erste repräsentative Hinweise aus Elternperspektive vor: Die DJI-Kinderbetreuungsstudie (KiBS 2019–2020)⁹⁶ zeigt, dass ein relevanter Anteil der Eltern die Frequenz von Entwicklungsgesprächen als zu gering einschätzt (je nach Altersgruppe ca. ein Viertel bis knapp ein Drittel) und dass sich Bewertungen der Beziehung zu Fachkräften systematisch nach sozialen Lagen und sprachlichen Ressourcen unterscheiden können (Lippert et al., 2022). Zusammenfassend zeigt die kursorische Sichtung der gesetzlichen Grundlagen sowie der Bildungs-, Erziehungs- und Rahmenpläne, dass vor allem mehr oder weniger systematische Beobachtungsdaten, aber weniger weitere standardisierte Entwicklungsmaße systematisch zum Gegenstand der Entwicklungsgespräche gemacht werden.

6.2 Konzepte und Maßnahmen

6.2.1 Weiterentwicklung des Systemmonitorings

Weiterentwicklung der bestehenden Monitoring-Instrumente

Aus der Analyse des Status quo ergeben sich eine Reihe möglicher Maßnahmen zur Weiterentwicklung der datengestützten Qualitätsentwicklung und Steuerung auf den verschiedenen Ebenen des Systems der Frühen Bildung. Auf der Systemebene sind Zugang und Teilhabe zu sichern. Hierfür sind Daten zur Bevölkerungsentwicklung und Geburtenvorausberechnungen notwendig. Ferner besteht das Ziel, überregionale Handlungsbedarfe für die Qualitätsentwicklung und Steuerung zu identifizieren, etwa zur Anpassung von Bildungsplänen oder zur Ressourcensteuerung übergreifender Qualitätsinitiativen. Handlungsoptionen liegen zunächst in der Weiterführung und der Verstetigung aller relevanten nationalen Monitoringsysteme für die pädagogische Qualität in der Kindertagesbetreuung, insbesondere ERiK. Ferner würde eine Weiterentwicklung der Kinder- und Jugendhilfestatistik im Sinne einer Ergänzung relevanter Daten, die die

⁹⁴ Vgl. <https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/landesrechtliche-regelungen/personal-und-einrichtungen/uebersicht-der-regelungen-2> und einzelne aktuellere Fassungen (aufgerufen am 20.02.2026).

⁹⁵ Vgl. <https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/landesrechtliche-regelungen/personal-und-einrichtungen/bildungsplaene-1-1> (aufgerufen am 20.02.2026).

⁹⁶ https://www.dji.de/kibs_19-20 (aufgerufen am 20.02.2026).



Belastungen der Kinder und Familien abbilden, sowie der Entwicklung der Cell-Key-Methode das Potenzial zur Qualitätsentwicklung und Steuerung im Sinne der Identifikation regionaler Entwicklungsbedarfe erhöhen. Besonders relevant ist hier auch die Verknüpfung von Kita- und Trägerdaten. Auch wenn Daten zu besonderen Förderbedarfen der Kinder vorliegen und die Statistik kleinräumige Auswertungen ermöglicht, fehlen wichtige Informationen über die ökonomische Lage der Kinder, den Sozialraum sowie die Ausstattung der Kitas mit spezifischen Fachkräften. Zudem gestaltet sich die systematische Erfassung der Kita-Träger als herausfordernd, da diese bislang nicht umfassend in die Statistik integriert sind und es keine Möglichkeit gibt, die Daten mit denen der Kitas zu verknüpfen.

Die Daten, die durch das DJI zur Verfügung gestellt werden, haben großes Potenzial zur Ressourcenplanung und Steuerung politischer Initiativen und Gesetze zur Qualitätsentwicklung, zur Fachkräftegewinnung und -sicherung sowie zur Verteilung von Ressourcen in der frühkindlichen Bildung. Sie werden bislang aber noch nicht voll ausgeschöpft. Dementsprechend könnte eine größere Verbindlichkeit bei der Orientierung an real mit Daten dokumentierten Entwicklungsbedarfen bei der Verteilung von Ressourcen sowie bei der Entwicklung von Initiativen der Qualitätsentwicklung eine zielgenauere Weiterentwicklung befördern.

Bundesweites Monitoring der Prozess- und Interaktionsqualität

Die zentrale Qualitätskomponente im Hinblick auf das Bildungspotenzial der frühkindlichen Bildung ist die pädagogische Prozess- und Interaktionsqualität (z. B. Anders & Oppermann, 2024; Ulferts et al., 2019). Diese lässt sich valide durch etablierte Beobachtungsinstrumente wie z. B. die *Kindergarteneinschätzungsskala (KES-RZ/KES-E)*, Roßbach et al., 2018; Tietze et al., 2017), das *Classroom Assessment Scoring System (CLASS)*, Pianta et al., 2008) oder die deutsche Übersetzung der *Sustained-Shared-Thinking-Emotional-Well-Being-Skala (SSTEWS)*, Siraj et al., 2015) für unterschiedliche Altersklassen erfassen.

Wiederholte Qualitätserhebungen haben gezeigt, dass gezielte Entwicklungsmaßnahmen die Prozessqualität messbar verbessern können, was die Einführung eines regelmäßigen Prozessmonitorings als notwendig erscheinen lässt (Kluczniok, Faas et al., 2024). Unterschiedliche Stakeholder aus verschiedenen Bereichen haben im Rahmen der Machbarkeitsstudie von Kluczniok, Faas et al. (2024) die Bedeutung eines bundesweiten Qualitätsmonitorings hervorgehoben, um eine datengestützte Steuerung des Kita-Systems zu ermöglichen. An anderen Stellen der Praxis werden aber auch immer wieder Vorbehalte gegenüber quantitativ orientierten Messansätzen deutlich. Große Herausforderungen für die Umsetzung eines solchen Monitorings umfassen die Notwendigkeit einer langfristigen Finanzierung, die Berücksichtigung regionaler Unterschiede sowie die Schaffung eines gemeinsamen Qualitätsverständnisses. Internationale Beispiele zeigen, dass ein Prozessmonitoring in föderalen Systemen erfolgreich implementiert werden kann, wie z. B. in England oder den Niederlanden (Kluczniok, Faas et al., 2024). In den Niederlanden wird im Rahmen eines nationalen Qualitätsmonitorings im Auftrag des Ministeriums für Soziales jährlich die Interaktionsqualität in einer Stichprobe von Kinderbetreuungseinrichtungen mit validierten Beobachtungsinstrumenten untersucht, die auch internationale Vergleiche ermöglichen.⁹⁷ Die Grundschule startet für viele Kinder mit vier, verpflichtend dann nach dem fünften Geburtstag. Hier sind alle vier Jahre Inspektionen

⁹⁷ <https://issa.nl/news/national-quality-monitor-child-care-netherlands> (aufgerufen am 20.02.2026).



vorgesehen, die auch die Beobachtung der Unterrichtsqualität einschließen (Inspectorate of Education, 2025). In Großbritannien ist Ofsted (Office for Standards in Education, Children's Services and Skills) für die Inspektion von Schulen und Einrichtungen der Frühen Bildung und Betreuung zuständig.⁹⁸ Einrichtungen der Frühen Bildung werden mindestens alle 6 Jahre inspiziert auf Basis ausformulierter Standards,⁹⁹ bei Auffälligkeiten auch häufiger,¹⁰⁰ Schulen (ab 4 bis 5 Jahre, teilweise mit Angeboten für 2- bis 3-jährige Kinder) in der Regel alle 4 Jahre.¹⁰¹ Es gibt unterschiedliche Möglichkeiten der Ausgestaltung eines bundesweiten Qualitätsmonitorings. Es bedarf jedoch einheitlicher Standards und einer länderübergreifenden, systematischen Datenerfassung.

Ein bundesweites Qualitätsmonitoring würde zudem politischen Entscheidungsträger:innen relevante Daten liefern, die als Grundlage für die Entwicklung und Anpassung politischer Initiativen und Programme dienen können. Die evidenzbasierten Ergebnisse könnten in die Formulierung qualitativ hochwertiger Standards und die Optimierung gesetzlicher Rahmenbedingungen einfließen, um die frühkindliche Bildung nachhaltig zu verbessern (Kluczniok, Faas et al., 2024). Ein weiteres wichtiges Potenzial liegt in der Vergleichbarkeit und Transparenz der erhobenen Daten. Durch einheitliche Erhebungsmethoden und Standards lassen sich Daten zwischen verschiedenen Einrichtungen und Ländern vergleichen, was die Transparenz im Bildungssystem fördert und den Austausch von Best Practices erleichtert. Schließlich kann ein solches Monitoring als langfristiges Steuerungsinstrument dienen, das die Entwicklung des frühkindlichen Bildungssystems nachhaltig unterstützt.

Insgesamt bietet ein bundesweites Monitoring der Prozessqualität in der Frühen Bildung die Möglichkeit, eine datengestützte, gerechte und qualitativ hochwertige Bildungslandschaft zu schaffen, die den Bedürfnissen von Kindern und Familien gerecht wird.

Potenzial der Erfassung der Kompetenzen der Kinder

Im System der frühkindlichen Betreuung, Bildung und Erziehung gibt es, anders als im Schulsystem, in Deutschland keine regelmäßigen nationalen oder internationalen Studien zu den Kompetenzen der Kinder. Trotzdem wird – teilweise perspektivisch – eine Reihe von Daten erhoben, die eine ähnliche Funktion erfüllen könnten. So wird mit StarS (vgl. Kapitel 1) ein Instrumentarium für Kinder am Ende des ersten Grundschuljahres entwickelt, das sich perspektivisch auch für das Ende der Kita-Zeit anpassen und weiterentwickeln ließe. Schon jetzt sehen viele Länder vor Beginn der Schulpflicht Schuleingangsuntersuchungen vor. Sie liegen in der Verantwortung der Gesundheitsämter, umfassen aber neben vielen Maßen der körperlichen Entwicklung auch Maße der sprachlich-kognitiven und sozial-emotionalen Entwicklung. Auch die Sprachstandserhebungen, die gerade sukzessive in den Ländern eingeführt werden, haben – einheitliche und validierte Erhebungsinstrumente vorausgesetzt – das Potenzial, das landes- und bundesweite

⁹⁸ <https://www.gov.uk/government/organisations/ofsted/about#our-responsibilities> (aufgerufen am 20.02.2026).

⁹⁹ https://assets.publishing.service.gov.uk/media/68c15c7f8c6d992f23edd81a/Early_years_inspection_toolkit.pdf (aufgerufen am 20.02.2026).

¹⁰⁰ <https://www.gov.uk/government/publications/early-years-inspection-toolkit-operating-guide-and-information/early-years-inspection-operating-guide-for-inspectors-for-use-from-november-2025> (aufgerufen am 20.02.2026).

¹⁰¹ <https://www.gov.uk/government/publications/school-inspection-toolkit-operating-guide-and-information/inspection-information-for-state-funded-schools-for-use-from-november-2025> (aufgerufen am 20.02.2026).



Monitoring zu ergänzen, nationale und regionale Entwicklungsbedarfe aufzuzeigen und Ansatzpunkte für eine sinnvolle Ressourcensteuerung zu liefern.

6.2.2 Datengestützte Steuerung und organisationale Qualitätsentwicklung

Ein bundesweites Monitoring der Prozessqualität in der Frühen Bildung, das auf systematisch erhobenen Beobachtungsdaten basiert, könnte ein zentrales Instrument zur Förderung der Bildungsqualität und Chancengerechtigkeit in Deutschland sein (Kluczniok, Faas et al., 2024). Ein solches System bietet erhebliche Potenziale – sowohl für die zielgenaue Verteilung von Ressourcen (z. B. Stellenanteile) als auch für die kontinuierliche Qualitätsentwicklung in Kindertageseinrichtungen (Kitas). Durch die datengestützte Analyse von Bildungsprozessen lassen sich spezifische Bedarfe von Einrichtungen, Regionen und Kindern identifizieren und gezielt adressieren.

Ein zentraler Vorteil eines flächendeckenden Qualitätsmonitorings liegt in der bedarfsgerechten Allokation von Mitteln. Bisher werden Ressourcen oft ausschließlich nach pauschalen und sozioökonomischen Kriterien (z. B. Gruppengröße oder Sozialindex) verteilt, ohne Qualitätsmerkmale und Qualitätsentwicklungsbedarfe der Einrichtungen zu berücksichtigen. Ein datengestützter Ansatz ermöglicht es hingegen, finanzielle Mittel, fachliche strukturelle Unterstützungsangebote und Fort- und Weiterbildung dort einzusetzen, wo sie am dringendsten benötigt werden, um regionale Ungleichheiten in der frühkindlichen Bildung zu verringern (Kluczniok, Grad et al., 2024) und Personal- und Zeitressourcen (z. B. für sprachliche Förderung oder Fortbildungen) gezielt zu steuern. Ein aktuelles Pilotprojekt des Trägers Fröbel zeigt beispielsweise, dass die Herkunftssprache von Kindern nur moderat mit ihrem tatsächlichen sprachlichen Förderbedarf korreliert (Anders et al., i. E.). Dies unterstreicht, wie wichtig es ist, entwicklungsbezogene Daten ergänzend zu sozioökonomischen Indikatoren in die Qualitätsentwicklung und -steuerung einzubeziehen – statt sich auf pauschale Indikatoren wie den Migrationshintergrund zu verlassen. Eine solche evidenzbasierte Steuerung könnte die Effizienz und Wirksamkeit von Fördermaßnahmen deutlich erhöhen.

Wenn die Ergebnisse an Träger und Einrichtungen zurückgemeldet werden, können sie auch auf organisationaler Ebene ihr volles Potenzial entfalten. Die regelmäßige Erhebung und Auswertung von Daten der Prozess- und Interaktionsqualität schaffen die Grundlage für eine kontinuierliche Qualitätsentwicklung in Kitas. Durch die Identifikation konkreter Stärken und Schwächen in den Bildungsangeboten können Träger und Einrichtungen gezielte Maßnahmen zur Verbesserung ergreifen (z. B. Fortbildungen, Mentoring-Programme), pädagogische Konzepte an die Bedürfnisse und Entwicklungsbedarfe der Kinder anpassen und damit Bildungschancen langfristig erhöhen.

Eine besondere Rolle spielen dabei Evaluationssysteme, die sowohl intern (durch die Einrichtungen selbst) als auch extern (durch unabhängige Stellen) durchgeführt werden. Während interne Evaluationen die Selbstreflexion der Teams fördern, weisen externe Evaluationen oft eine höhere Validität und Objektivität auf. Eine Verbindung beider Ansätze – kombiniert mit einer verbindlichen Verankerung in Qualitätsmanagement-Systemen – könnte die Datenqualität und deren Nutzung weiter verbessern.

Auch die Sprachstandserhebungen weisen Nutzungspotenziale auf Ebene einzelner Träger und Einrichtungen auf. Exemplarisch ist hier ein Ansatz von Fröbel e.V. zu nennen (Anders et al., i. E.). In diesem Projekt



wurde auf Basis der *Begleitenden alltagsintegrierten Sprachentwicklungsbeobachtung in Kindertageseinrichtungen (BaSiK)* (Zimmer, 2019) ein Schwellenwert zur Feststellung eines ergänzenden Sprachförderbedarfs inhaltlich entwickelt und empirisch an Daten aus Einrichtungen des Trägers erprobt. Zusätzlich wurden Praxisperspektiven auf Entwicklungsstandserhebungen in einer qualitativen Teilstudie untersucht sowie die Steuerungsmöglichkeiten und -potenziale auf Trägerebene reflektiert und diskutiert. Eine große Rolle spielten hier auch die Möglichkeiten, die durch die Online-Plattform KITALINO entstehen, denn der Algorithmus zur Erfassung eines ergänzenden Sprachförderbedarfs könnte problemlos in Online-Plattformen integriert werden, sofern die Daten digital erfasst werden. Darüber hinaus kann die Auswertung von Entwicklungsdaten auf Einrichtungsebene dazu genutzt werden, spezifische Angebote der Elternbildung zu konzipieren und vorzuhalten.

Damit datengestützte Qualitätsentwicklung ihr volles Potenzial entfalten kann, sind drei zentrale Stell-schrauben entscheidend:

- *Verbindlichkeit in der Qualitätsentwicklung und Steuerung:* Aktuell fehlt es oft an einer verpflichtenden Kopplung zwischen (strukturellen) Qualitätsdaten (wie der Personalschlüssel) und der Verteilung von Mitteln. (Strukturelle) Qualitätsdaten weisen den Qualitätsentwicklungsbedarf aus. Eine gesetzliche Verankerung auf Landesebene, die Träger und Kommunen verpflichtet, Ressourcen evidenz- und datengestützt einzusetzen, könnte die Zielgenauigkeit politischer Maßnahmen deutlich erhöhen.
- *Transparenz durch die Veröffentlichung von Evaluationsergebnissen:* Eltern sollten Zugang zu Qualitätsdaten der Einrichtungen haben, um informierte Entscheidungen bei der Kita-Wahl treffen zu können. Eine öffentliche Berichterstattung (z. B. in Form von Qualitätsberichten) stärkt zudem die Rechenschaftspflicht der Träger und fördert den Wettbewerb um Qualität.
- *Passgenaue Unterstützungsangebote für Fachkräfte:* Qualitätsdaten können genutzt werden, um die Weiterentwicklung der Ausbildung, Fortbildungen und Fachberatungsangebote präzise auf die Bedarfe der Teams zuzuschneiden. Darüber hinaus können Qualitätsdaten genutzt werden, um die Strukturqualität in Regionen und Einrichtungen mit besonderen Herausforderungen passgenau zu stärken. Entwicklungsstandsdaten (insbesondere zum Sprachstand der Kinder) können hierbei ebenfalls wichtige Informationen über die besonderen pädagogischen Herausforderungen sowie die strukturellen und fachlichen Unterstützungsbedarfe von Einrichtungen oder Regionen liefern. Besonders wichtig ist dabei die mittelbare pädagogische Arbeitszeit, die für Reflexion, Teamabsprachen und Konzeptentwicklung genutzt werden sollte.

Ein bundesweites Qualitätsmonitoring in der Frühen Bildung, das Prozessqualität, die Nutzung von Daten aus der Erfassung kindlicher Entwicklungsstände und Ressourcenverteilung systematisch verknüpft, könnte einen Paradigmenwechsel in der Kita-Landschaft einleiten. Durch die zielgenaue Steuerung von Mitteln, die kontinuierliche Qualitätsentwicklung und eine größere Transparenz ließen sich nicht nur Bildungsungleichheiten verringern, sondern auch die Wirksamkeit öffentlicher Investitionen deutlich steigern. Entscheidend ist dabei, dass die erhobenen Daten nicht nur gesammelt, sondern auch konsequent genutzt werden – sei es in der politischen Planung, der Trägersteuerung oder der pädagogischen Praxis.



6.2.3 Bildungs- und Förderangebote auf Basis von Entwicklungsstandsdaten und Übergang in die Grundschule

Wie oben beschrieben, wird in den Ländern eine Reihe von Verfahren eingesetzt, um den Entwicklungsstand der Kinder zu erfassen und zu dokumentieren. Ähnlich wie bei den Sprachstandserhebungen steht auch hier eine Fokussierung auf eine überschaubare Anzahl qualitätsgeprüfter Verfahren aus. Oftmals werden nicht nur in denselben Ländern, sondern auch bei denselben Trägern unterschiedliche Verfahren eingesetzt. Ein bislang unzureichend ausgeschöpftes Potenzial liegt darin, die Dokumentations- und Entwicklungsstandsdaten auch verbindlich für die individuelle Planung pädagogischer Angebote zu nutzen. Im Sinne einer ganzheitlichen Bildung gilt dies für alle Fähigkeits- und Fertigkeitsbereiche sowie motivationalen Voraussetzungen und Interessen der Kinder. Im Sinne einer durchgängigen Unterstützung der Entwicklung basaler Kompetenzen gilt es jedoch besonders dort, wo die Erhebungen des Entwicklungsstands einen individuellen, ergänzenden Förderbedarf in den sprachlichen, mathematischen oder sozial-emotionalen Kompetenzen erkennen lassen. Das SWK-Gutachten zur Grundschule (2022a) enthält eine Vielzahl von Empfehlungen für Bildungsangebote und die ergänzende Förderung sprachlicher, mathematischer und sozial-emotionaler Kompetenzen. Entscheidend ist, im Sinne einer zirkulären Fallarbeit Anamnese, Erfassung des Entwicklungsstands, Bildungs- und Förderangebote und Evaluation systematisch zu verknüpfen. In den angloamerikanischen Systemen der Frühen Bildung ist eine solche professionelle Fallarbeit, die durch datengestützte Förderentscheidungen gekennzeichnet ist, für die frühe sprachliche Bildung weitverbreitet (vgl. *Response to Intervention* oder *Recognition & Response*; Buysse & Peisner-Feinberg, 2010; Egert et al., 2025) und in ihrer Wirksamkeit nachgewiesen (Buysse et al., 2016). Auch für die frühe mathematische Bildung (Methe & van der Heyden, 2011; Mononen et al., 2015) gibt es diese Ansätze. Evidenzbasierte mehrstufige Programme zur Förderung sozial-emotionaler Fähigkeiten wie *School-wide* oder *Program-wide Positive Behaviour Support (SWPBS/PWPBS)* folgen derselben Logik. Sie sehen auf der ersten Stufe eine universelle Förderung aller Kinder, auf der zweiten Stufe eine gezielte Förderung weniger Kinder mit dem Risiko der Entwicklung von Verhaltensauffälligkeiten und auf der dritten Stufe eine intensive Unterstützung von Kindern mit komplexen Problemlagen vor (Hemmeter et al., 2007; Piedade et al., 2026).

Werden mit validen Testinstrumenten erfasste Entwicklungsstandsdaten an die Grundschule weitergegeben, eröffnet dies die Möglichkeit, ohne größere Brüche an den Entwicklungsstand angepasste Bildungsangebote ab dem Schuleintritt fortzusetzen. Voraussetzungen hierfür sind allerdings valide und reliable Erhebungsinstrumente, entsprechend qualifizierte Fachkräfte sowie vertrauensvolle Partnerschaften zwischen Eltern, Fachkräften und Grundschullehrkräften und eine entsprechende Qualifikation der Grundschullehrkräfte, wie es auch Kapitel 1.2.2 zur diagnosebasierten Förderung im Unterricht vorsieht. All das könnte Sorgen entgegenwirken, dass Kinder damit mit dauerhaften Etikettierungen in die Grundschule starten.

6.2.4 Datengestützte Erziehungs- und Bildungspartnerschaften

Wie bereits in Kapitel 4 für das Schulalter dargestellt wurde, können Erziehungs- und Bildungspartnerschaften datengestützt so gestärkt werden, dass positive Effekte auf die Entwicklung der Kinder zu erwarten sind. Nationale und internationale Studien zur Wirksamkeit früher Bildungsangebote zeigen ebenfalls,



dass frühe Bildungsprogramme dann besonders stark auf die kindliche Entwicklung wirken, wenn neben den Kindern auch intensiv mit den Eltern gearbeitet wird (Anders, 2013). Einzelne Studien aus Deutschland weisen ebenfalls auf die spezifischen förderlichen Effekte einzelner Unterstützungsangebote und Anregungen frühpädagogischer Fachkräfte für Eltern hin (Lehrl et al., 2020). Die förderlichen Effekte lassen sich vor dem Hintergrund verstehen, dass ein hoher Anregungsgehalt in der Familie die Entwicklung der Kinder positiv beeinflusst (Lehrl, 2018).

Die oben dargestellten Formen der Datenerhebung und Dokumentation in Kindertageseinrichtungen eignen sich grundsätzlich für den Austausch mit den Erziehungsberechtigten: Entwicklungsdaten bieten wertvolle Einblicke in den Entwicklungsstand und die Fortschritte der Kinder, insbesondere wenn auf der Basis individuelle Entwicklungsberichte erstellt und regelmäßig aktualisiert werden. Diese Berichte dokumentieren Fortschritte in verschiedenen Bereichen wie Sprache, Motorik und sozial-emotionale Fähigkeiten. Entwicklungsportfolios und andere Dokumentationen, die die Fortschritte und besonderen Meilensteine eines Kindes festhalten, können visuell ansprechend gestaltet werden, um den Eltern einen klaren Überblick über die Entwicklung ihres Kindes zu bieten.

Digitale Plattformen spielen eine zunehmend wichtige Rolle. In den letzten Jahren ist eine Reihe von Kita-Apps auf den Markt gekommen, die entweder allein der Bildungs- und Entwicklungsdokumentation oder der Kommunikation mit den Eltern dienen oder als Komplettlösung beides miteinander vereinen. Laut zwei bayernweiten Kitaleitungsbefragungen hat sich der Einsatz von Kita-Apps zwischen 2021 und 2024 von 26 auf 53 Prozent verdoppelt, wobei offenbleibt, ob diese vor allem der Kommunikation oder auch der Entwicklungsdokumentation dienen (Reichert-Garschhammer et al., 2025). Der Einsatz solcher Online-Tools zur Dokumentation und zum Austausch von Entwicklungsdaten kann es Eltern ermöglichen, regelmäßig über den Fortschritt ihrer Kinder informiert zu werden. Interaktive Elemente auf diesen Plattformen fördern zudem die direkte Rückmeldung der Eltern und bieten ihnen die Möglichkeit, aktiv die Entwicklung ihrer Kinder zu unterstützen. So können alters- und entwicklungsangepasste Elemente eingebettet werden, die Eltern Beispiele und Anregungen geben, wie sie ihre Kinder in ihrer Entwicklung fördern können (Jelley & Sylva, 2018; York et al., 2019). Solche Apps können auch mit spezifischen Aufgaben oder verlinkten Anschauungsvideos angereichert werden. So zeigt eine Untersuchung aus Deutschland zum Potenzial einer digitalen Eltern-App für Familien mit Kleinkindern (3 bis 24 Monate) positive Effekte der App-Nutzung auf die sprachbezogene elterliche Selbstwirksamkeit sowie auf sprachanregende Aktivitäten, wie beispielsweise die gemeinsame Betrachtung von Bilderbüchern (Prokupek et al., 2025). Datenschutz und Datensicherheit sind dabei zu berücksichtigen (Gruber et al., 2022).

Die verschiedenen Formen der datengestützten Rückmeldungen dienen als Grundlage für Gespräche zwischen Erzieher:innen und Eltern. Eltern können ihre Beobachtungen und Erfahrungen aus dem Alltag einbringen und, wenn sie wollen, auch Informationen aus weiteren Quellen wie den pädiatrischen Vorsorgeuntersuchungen einbringen, was zu einer umfassenderen Sicht auf die Entwicklung des Kindes führt. Solche Gespräche bieten den Eltern die Möglichkeit, den Entwicklungsstand ihres Kindes zu reflektieren, gemeinsam mit den Fachkräften realistische Ziele zu setzen und Vereinbarungen zu treffen, wie in Kita und Familie darauf hingewirkt wird. Idealerweise mündet dies in die Entwicklung individueller Bildungs- oder Förderpläne für jedes Kind, die auf den gesammelten Entwicklungsdaten und weiteren Beobachtungen basieren. Diese Pläne werden in enger Zusammenarbeit mit den Eltern erstellt, wodurch eine aktive Beteiligung gefördert



wird. Regelmäßige Feedbackgespräche zwischen Erzieher:innen und Eltern können die Reflexion über die kindliche Entwicklung stärken und eine kontinuierliche Anpassung der Bildungsangebote ermöglichen.

Grundsätzlich bieten darüber hinaus Elternabende, Workshops und Kurse die Möglichkeit, Eltern über die spezifischen Entwicklungsphasen ihrer Kinder zu informieren, ihnen praktische Tipps zu geben, wie sie die Entwicklung zu Hause unterstützen können und dabei auch kindspezifische Bedarfe aufgreifen. Große Effekte auf die Kompetenzentwicklung der Kinder lassen sich durch intensive und gezielte Interventionen sowie engen Kontakt zwischen Fachkräften und Familien erzielen (Axford et al., 2019; X. Ma et al., 2016), sodass die pädagogischen Fachkräfte bei Bedarf die Eltern über weiterführende Elternbildungsprogramme informieren können. Eine enge Vernetzung zwischen Kitas und anderen Anbietern von Frühen Hilfen, Familienunterstützungsprogrammen und Elternbildungsansätzen sollte hier angestrebt werden. Auch eine systematische Verkettung der verschiedenen Unterstützungsangebote ist von Vorteil (Schütte et al., 2024). Familienzentren können einen idealen Ort der Verankerung bieten. Veranstaltungen und Feste in der Einrichtung können ebenfalls genutzt werden, um den Kindern die Gelegenheit zu geben, Projekte, Spiele, Lieblingsbücher etc. vorzustellen, während die Einbeziehung der Eltern in gemeinsame Aktivitäten die Partnerschaft zwischen Eltern und Erzieher:innen weiter festigt.

Schließlich ist die professionelle Entwicklung der frühpädagogischen Fachkräfte von großer Bedeutung. Fortbildungen zur Interpretation und Kommunikation von Entwicklungsdaten unterstützen die Fachkräfte dabei, effektive Partnerschaften mit Eltern aufzubauen (Betz et al., 2019). Der regelmäßige Austausch mit Fachleuten ermöglicht es Fachkräften, bewährte Methoden zur Nutzung von Entwicklungsdaten in der Zusammenarbeit mit Eltern zu identifizieren und anzuwenden. Die gezielte Nutzung von Entwicklungsdaten in Kindertageseinrichtungen kann nicht nur das Verständnis und die Unterstützung der kindlichen Entwicklung verbessern, sondern auch das Engagement der Eltern fördern.

6.3 Schlussfolgerungen und Empfehlungen

Eine systematische Nutzung valider Daten auf den verschiedenen Ebenen des Systems der Frühen Bildung kann entscheidend zur Weiterentwicklung der pädagogischen Qualität beitragen. Auf der Systemebene von Bund, Ländern und Kommunen liegen mit der Kinder- und Jugendhilfestatistik sowie den laufenden Erhebungen des Deutschen Jugendinstituts reichhaltige Datensätze vor. Diese sind jedoch bislang nur unzureichend aufeinander abgestimmt, sodass ihr Analysepotenzial nicht voll ausgeschöpft werden kann. Darüber hinaus sind notwendige lokale und regionale Analysen der Herausforderungslagen aufgrund von Datenschutzbestimmungen bislang nicht vollumfänglich möglich. Daten aus Entwicklungsstandserhebungen haben ein zusätzliches Potenzial für die Identifikation von Einrichtungen mit besonderen Unterstützungs- und Entwicklungsbedarfen, das weiterentwickelt werden sollte. Eine solche Optimierung der Datenlage in Richtung eines systematischen Qualitätsmonitorings schafft eine Grundlage dafür, politische Initiativen zur Qualitätsverbesserung und die Ressourcenzuweisung auf Landes- und kommunaler Ebene passgenau auf die identifizierten Bedarfe abzustimmen.

Zentral für das Bildungspotenzial der frühkindlichen Bildungseinrichtungen ist eine hohe Prozess- und Interaktionsqualität. Daher zielen politische Programme und Initiativen zur Qualitätsentwicklung in der



Regel implizit oder explizit auf ihre Verbesserung ab. Die globale/emotionale Qualität ist in der Regel relativ hoch, aber im Bereich der frühen bereichsspezifischen Förderung und Bildung (z. B. Sprache, Mathematik) ist sie oft (zu) gering. Besonders gering ist sie in Kitagruppen mit einem hohen Anteil von Kindern mit Zuwanderungsgeschichte, wo von einem hohen Anteil von Kindern mit Förderbedarfen auszugehen ist. Es fehlt darüber hinaus an flächendeckenden Monitoring-Daten zur Prozess- und Interaktionsqualität in den Kindertageseinrichtungen, wie sie in anderen Staaten bereits regelmäßig erhoben werden. Sie können in den nationalen Bildungsbericht einfließen und an Einrichtungen und Familien zurückgemeldet werden. Solche Daten können eine zentrale Grundlage für Qualitätsentwicklungsprozesse in jeder einzelnen Einrichtung bilden.

Die Kita-Aufsicht hat auf der Ebene der einzelnen Kindertageseinrichtungen und Träger einen Fokus auf die Einhaltung gesetzlicher Bestimmungen zur Errichtung und auf den laufenden Betrieb. Die Fachberatungen der Träger unterstützen die Umsetzung und Weiterentwicklung der pädagogischen Konzepte sowie der Qualität. Die Qualitätsentwicklung der Einrichtungen soll in den meisten Ländern durch regelmäßige interne Evaluationen, teilweise auch durch externe Evaluationen, vorangetrieben werden, deren Ergebnisse jedoch nicht überall systematisch transparent gemacht werden. Es existiert oft eine Vielzahl von Evaluationsinstrumenten, die auch trägerspezifisch gestaltet sind.

Auch wenn der Verbindlichkeitsgrad variiert, wird zur Dokumentation der pädagogischen Arbeit, der Interessen und Motivationen sowie der Entwicklung der Kinder und Sprachstandserhebungen in den Kindertageseinrichtungen eine Vielzahl von Systemen eingesetzt, die in der Regel auf Beobachtungen durch einzelne Fachkräfte basieren. Wissenschaftliche Expertisen verweisen vielfach auf mangelnde Güte der eingesetzten Dokumentationssysteme. Darüber hinaus sind diese für die Fachkräfte mit hohem Fortbildungsbedarf und bei der Durchführung mit hohem Zeitaufwand verbunden. Auch für die sukzessive eingeführten landesweiten Sprachstandserhebungen werden sehr viele unterschiedliche Instrumente verwendet und weiterentwickelt. Damit bestehen große Weiterentwicklungsbedarfe in einer Validierung und Verschlan- kung der Instrumente zur Erfassung und Dokumentation der pädagogischen Arbeit und Qualität sowie der kindlichen Entwicklungsstände. Mit Blick auf die pädagogische Qualität besteht auch Bedarf an einer größeren Transparenz der Ergebnisse für und ihrer Nutzung durch Eltern, Träger u. a. m.

Mit den Schuleingangsuntersuchungen liegen weitere Daten zum Entwicklungsstand vor. Eine Herausforderung besteht weiterhin darin, die vielfältigen Entwicklungsstandserhebungen verbindlich in sinnvolle pädagogische Anschlussprozesse zu überführen. Ein offener Punkt ist darüber hinaus die Möglichkeit der Weitergabe der Daten zwischen Kindertageseinrichtung und Grundschule zur optimalen Gestaltung eines anschlussfähigen Übergangs unter Einbindung der Eltern.

Die Bildungspläne (Rahmen- und Orientierungspläne) und Ausführungsgesetze der Länder sehen zumeist regelmäßige Elterngespräche vor, um Dokumentationen der pädagogischen Arbeit, kindliche Interessen sowie Ergebnisse aus Entwicklungsstandserhebungen zu besprechen. Darüber hinaus bestehen vielfältige Möglichkeiten zur Gestaltung der Erziehungs- und Bildungspartnerschaft sowie zur Verankerung von Bildungsangeboten, die datengestützt angereichert werden können.

Eine datengestützte Qualitätsentwicklung auf allen Ebenen des Systems der Frühen Bildung setzt voraus, dass die pädagogischen Fach- und Leitungskräfte sowie Trägervertreter:innen, Fachberatungen und das



Unterstützungssystem durch Aus-, Fort- und Weiterbildung dafür qualifiziert und Zeitressourcen dafür geschaffen sowie nachhaltig systematisch verankert werden.

Vor diesem Hintergrund empfiehlt die SWK:

Empfehlung 18) Eine Gesamtstrategie zum Systemmonitoring in der Frühen Bildung entwickeln; dabei

1. bestehende nationale Datensätze (ERiK, AID:A, KJH-Statistik) für Berichterstattung und Monitoring weiterentwickeln und aufeinander abstimmen,
2. Analysepotenziale der Kinder- Jugendhilfestatistik erhöhen, um regionale Qualitätsentwicklungs- und Unterstützungsbedarfe identifizieren zu können (Cell-Key-Methode, Verknüpfung von Träger- und Kita-Daten),
3. landesweite Daten (z. B. zur Strukturqualität) zur Abbildung des regionalen Qualitätsentwicklungsbedarfs und zur Steuerung übergreifender Qualitätsinitiativen verbindlich nutzen, damit die Qualität an den Stellen weiterentwickelt wird, an denen ein tatsächlicher Bedarf besteht,
4. perspektivisch ein nationales, flächendeckendes beobachtungs-basiertes Monitoring der Prozess- und Interaktionsqualität in allen Kindertageseinrichtungen einführen; hierfür sollte auf validierte Verfahren (z. B. KES-R-E; CLASS, SSTEW) zurückgegriffen werden und Umsetzungsmodi wie in erfolgreichen internationalen Beispielen (z. B. Niederlande, England) gewählt werden,
5. das Potenzial von Schuleingangsuntersuchungen, StarS und der Sprachstandserhebungen zu Monitoringzwecken des frühkindlichen Bildungssystems weiterentwickeln.

Empfehlung 19) Datengestützte Qualitätsentwicklung von Kindertageseinrichtungen und Steuerung auf der organisationalen Ebene weiterentwickeln; dabei

1. Verbindlichkeit interner Evaluationen erhöhen und Evaluationsergebnisse sowie ggf. einrichtungsspezifische Monitoringergebnisse (vgl. Empfehlung 18) transparent an die Eltern kommunizieren,
2. vorhandene Qualitätsdaten (Strukturqualität und perspektivisch Prozess- und Interaktionsqualität) zur Identifikation und Begründung des Entwicklungsbedarfs bei der Zuweisung von Ressourcen in länderübergreifenden und länderspezifischen oder regionalen Initiativen der Qualitätsentwicklung verbindlich nutzen,
3. Sprachstandserhebungen mit validierten Instrumenten in allen Ländern verpflichtend durchführen und Erhebungssysteme innerhalb der Länder vereinheitlichen,
4. Entwicklungsstandsdaten der Kinder bei der regionalen und übergreifenden Zuweisung von Ressourcen zur Qualitätsentwicklung ergänzend berücksichtigen.



Empfehlung 20) Bildung und Förderung der Kinder sowie ihren Übergang in die Grundschule datengestützt stärken; dabei

1. Verfahren zur Erfassung des Entwicklungsstands, der kindlichen Interessen und Motivationen sowie der pädagogischen Arbeit unter Nutzung des Potenzials digitaler Instrumente auf eine überschaubare Zahl qualitätsgeprüfter Verfahren reduzieren,
2. Dokumentationen der pädagogischen Arbeit und Entwicklungsstandsdaten zur Planung und Durchführung individueller pädagogischer Angebote verbindlich nutzen,
3. basale Kompetenzen (Sprache, Mathematik, sozial-emotionale Fähigkeiten) auf Basis der Erhebung des Entwicklungsstands alltagsintegriert und ergänzend fördern; die pädagogischen Anschlussprozesse hierbei verbindlich klären und priorisieren,
4. die Möglichkeit einer auf einer vertrauensvollen Partnerschaft von Kindertageseinrichtung, Eltern und Grundschulen fußenden Weitergabe individueller Ergebnisse der Entwicklungsstandserhebungen der Kinder an die Grundschulen im Sinne einer anschlussfähigen Bildung prüfen.

Empfehlung 21) Erziehungs- und Bildungspartnerschaften datengestützt stärken; dabei

1. Eltern über Ergebnisse aus Dokumentationen der pädagogischen Arbeit, der Interessen und Motivationen der Kinder sowie der Entwicklungsstandserhebungen im Rahmen strukturierter Eltern- und Entwicklungsgespräche regelmäßig informieren,
2. Eltern zu Bildungsaktivitäten, Förder- und Entwicklungsmöglichkeiten auf Basis der Ergebnisse aus Dokumentationen der pädagogischen Arbeit, der Interessen und Motivationen der Kinder sowie Entwicklungsstandserhebungen beraten.

Empfehlung 22) Strukturelle und fachliche Voraussetzungen für die Weiterentwicklung der datengestützten Qualitätsentwicklung in der Frühen Bildung auf allen Ebenen schaffen; dabei

1. datengestützte Entwicklung von Lern- und Bildungsangeboten sowie datengestützte Erziehungs- und Bildungspartnerschaften verbindlich in der Ausbildung verankern,
2. Funktionsstellen zur Unterstützung der Dokumentation der pädagogischen Arbeit, der kindlichen Interessen und Motivationen und Entwicklungsstandserhebungen sowie der daran anschließenden Bildungs- und Förderprozesse schaffen (vgl. SWK-Gutachten zur Grundschule, 2022a),
3. Zeitressourcen für datengestützte Qualitätsentwicklung im Rahmen der mittelbaren pädagogischen Arbeit für frühpädagogische Fachkräfte schaffen,
4. Zeitressourcen für datengestützte Qualitätsentwicklung für Kitaleitungen schaffen,
5. kontinuierliche Fort- und Weiterbildung und fachliche Unterstützung für Kitaleitungen, frühpädagogische Fachkräfte, Fachberatungen und Trägervertreter:innen anbieten.



Hinweise zur Umsetzung der Empfehlungen

Datengestützte Entwicklung und Steuerung in Schulen und der frühkindlichen Bildung erfordern einen systemischen Ansatz, der sowohl die Erfassung valider und relevanter Daten sowie deren adressatengerechte Aufbereitung und Rückmeldung als auch die Unterstützung der Datennutzung durch evidenzbasierte Maßnahmen umfasst. Eine systematische Implementation dieses Ansatzes erfordert gemeinsame Anstrengungen unterschiedlicher Akteure. Länder, Bund, Kommunen und Wissenschaft müssen dabei in unterschiedlichen Konstellationen zusammenwirken.

Folgende Maßnahmen sollten **im Schulbereich** prioritär eingeleitet werden, damit datengestützte Entwicklung und Steuerung in Lern- und Entwicklungsfortschritte von Kindern und Jugendlichen münden.

1. *Ländergemeinsam* sollten, ergänzend zur bereits laufenden Entwicklung von StarS sowie den Basiskompetenztests in der Grundschule, standardbasierte Tests vor allem für die Jahrgangsstufen 4 und 9 anhand der bereits etablierten Instrumente in der Verantwortung des IQB weiterentwickelt werden. Diese Tests dienen der Überprüfung der Wirksamkeit der Maßnahmen, die auf der Basis der Ergebnisse von VERA 3 und 8 getroffen wurden.
2. Um überfachliche Kompetenzen als Bildungsziele entsprechend dem internationalen Stand stärker auf allen Systemebenen zu verankern, ist die *ländergemeinschaftliche* Entwicklung eines Rahmenmodells erforderlich, das überfachliche Kompetenzen und sozial-emotionale Merkmale des Lernens beschreibt und mit einem fokussierten Set valider Indikatoren operationalisiert.
3. Initiativen zur systematischen Entwicklung lernprozessbegleitender Diagnosematerialien, verknüpft mit evidenzbasierten Fördermaterialien für fachliche und überfachliche Kompetenzen, sollten von den *Ländern gemeinsam mit dem Bund* vorangetrieben werden. Dazu ist neben einer gezielten, qualitätsgesicherten Programmforschung eine nachhaltig finanzierte Infrastruktur für Forschung und Entwicklung erforderlich, die – in Anlehnung an die Organisation von Translationsprozessen in der Medizin – in Form digitaler Zentren umgesetzt werden sollte.
4. Kompetenzen der diagnose- und materialbasierten Unterrichtsentwicklung und Förderung sowie der Data Literacy müssen in den *ländergemeinsamen* Standards für die Lehrkräftebildung verankert werden, damit sie in die Module der ersten und zweiten Phase verbindlich aufgenommen werden. Im gegenwärtig laufenden Prozess der Überarbeitung der Lehrkräftebildungsstandards ist es deshalb sicherzustellen, dass die für eine datengestützte Unterrichtsentwicklung und Lernförderung grundlegenden Kompetenzerwartungen berücksichtigt werden und ein Alignment mit den Bildungsstandards für Schüler:innen erfolgt.
5. Die Entwicklung und dauerhafte Bewirtschaftung einer *ländergemeinsamen* Plattform, auf der qualitätsgeprüfte, evidenzbasierte diagnostische Instrumente sowie Materialien zur Unterrichtsentwicklung und Lernförderung für Lehrkräfte niedrigschwellig zur Verfügung gestellt werden, sollte mit Mitteln des Bundes ermöglicht werden.

Diese Entwicklungen sind grundlegend für alle adressatenspezifischen Empfehlungen des Gutachtens: für eine datengestützte Unterrichtsentwicklung und Lernförderung, eine datengestützte Schulentwicklung, eine datengestützte Schulaufsicht, datengestützte Bildungspartnerschaften mit Eltern und pädagogischen Fachkräften sowie eine datengestützte Steuerung des Gesamtsystems.



Im Bereich der **Frühen Bildung** sollten folgende Maßnahmen mit höchster Priorität vorangetrieben werden, um die pädagogische Qualität datengestützt weiterzuentwickeln und zu verbessern:

1. Bestehende nationale Datensätze (ERiK, AID:A, KJH-Statistik) sollten für die Berichterstattung und das Monitoring weiterentwickelt, aufeinander abgestimmt und insbesondere die Analysepotenziale der Kinder- und Jugendhilfestatistik erhöht werden (Cell-Key-Methode, Verknüpfung von Träger- und Kita-Daten). So kann eine solide Datenbasis für die Identifikation von (regionalen) Qualitätsentwicklungs- und Unterstützungsbedarfen geschaffen werden.
2. Zur Stärkung der Qualitätsentwicklung auf der organisationalen Ebene sollte sowohl die Verbindlichkeit interner Evaluationen als auch die Transparenz der Kommunikation der Ergebnisse an die Eltern erhöht werden. Hierfür sollte trägerübergreifend ein Verständnis von Evaluationen zur Qualitätsentwicklung entwickelt und auch die Relevanz von Evaluationen als Teil von Qualitätsmanagement- und Qualitätsentwicklungssystemen aufgegriffen werden. In diesem Rahmen sollten die fachlichen Unterstützungs- und Qualifizierungsbedarfe von Führungskräften zur Nutzung von Evaluationsergebnissen für die Organisationsentwicklung adressiert werden.
3. Die Verfahren zur Erfassung des Entwicklungsstands, der kindlichen Interessen und Motivationen sowie zur Dokumentation der pädagogischen Arbeit sollten auf eine überschaubare Zahl qualitätsgeprüfter Verfahren reduziert und Sprachstandserhebungen in allen Bundesländern – mit innerhalb der Länder vereinheitlichten Verfahren – durchgeführt werden. Die pädagogischen Anschlussprozesse nach der Erfassung kindlicher Entwicklungsstände sollten verbindlich geklärt werden, sodass individuelle pädagogische Angebote für die Kinder auf Basis der Daten – unter Einbindung der Eltern – geplant und durchgeführt werden. Die fachlichen Unterstützungs- und Qualifizierungsbedarfe der frühpädagogischen Fachkräfte sollten in diesem Prozess adressiert werden.
4. Datengestützte Entwicklung und Durchführung pädagogischer Angebote und datengestützte organisationale Qualitätsentwicklung benötigen Zeit. Daher sollte die aktuell sich andeutende Entspannung bei der Personalsituation dafür genutzt werden, zusätzliche Zeitressourcen für datengestützte Qualitätsentwicklung in der mittelbaren pädagogischen Arbeitszeit der frühpädagogischen Fachkräfte sowie den verfügbaren Leitungsstunden der Kitaleitungen zu schaffen. Zusätzliche Funktionsstellen für die datengestützte Qualitätsentwicklung sollten geschaffen werden.



Literatur

- Adanne, E. F. (2024). A meta-analysis of data-driven school leaders and school effectiveness in the 21st century. *Journal of Human Resource and Sustainability Studies*, 12, 204–225. <https://doi.org/10.4236/jhrss.2024.121011>
- Aderet-German, T. & Ben-Peretz, M. (2020). Using data on school strengths and weaknesses for school improvement. *Studies in Educational Evaluation*, 64, Artikel 100831. <https://doi.org/10.1016/j.stueduc.2019.100831>
- Aleven, V. & McLaughlin, E. (2016). Instruction Based on Adaptive Learning Technologies. In R. E. Mayer & P. A. Alexander (Hrsg.), *Handbook of research on learning and instruction* (2nd edition, S. 522–560). Routledge.
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y. & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333(6045), 1034–1037. <https://doi.org/10.1126/science.1207998>
- Altrichter, H. (2019). Transformationen der Steuerungspolitik und das Schulprogramm als Koordinierungsinstrument. In U. Steffens & P. Posch (Hrsg.), *Lehrerprofessionalität und Schulqualität. Grundlagen der Qualität von Schule 4* (S. 287–314). Waxmann.
- Altrichter, H., Hautz, H. & Krainz, U. (2022). Schulentwicklungsberatung in Österreich – Entfaltung eines Arbeitsfeldes im schulischen Unterstützungssystem. *Pädagogische Horizonte*, 6(2), 59–84. <https://doi.org/10.17883/PA-HO-2022-02-06>
- Altrichter, H. & Maag Merki, K. (Hrsg.). (2016). *Handbuch Neue Steuerung im Schulsystem* (2. Aufl.). Springer VS. <https://doi.org/10.1007/978-3-531-18942-0>
- Ammann, M., Wiesner, C. & Schratz, M. (2020). Über das Zusammenwirken von Schulaufsicht und Schulleiterinnen und Schulleitern: Einblicke in die Qualitätsinitiative Schulqualität Allgemeinbildung am Beispiel des Bilanz- und Zielvereinbarungsgesprächs. In E. D. Klein & N. Bremm (Hrsg.), *Unterstützung – Kooperation – Kontrolle. Zum Verhältnis von Schulaufsicht und Schulleitung in der Schulentwicklung* (S. 191–216). Springer VS. https://doi.org/10.1007/978-3-658-28177-9_10
- Anders, Y. (2013). Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung. *Zeitschrift für Erziehungswissenschaft*, 16(2), 237–275. <https://doi.org/10.1007/s11618-013-0357-5>
- Anders, Y., Hafner, F. & Hummel, T. G. (i. E.). *Bericht Projekt „SteBas“. Studie zur Erprobung der Potenziale von BaSiK für Verbesserung der pädagogischen Arbeit sowie zur bedarfsgerechten Steuerung von Ressourcen* [Unveröffentlichter Forschungsbericht]. Universität Bamberg.
- Anders, Y. & Oppermann, E. (2024). Frühpädagogische Qualität in Kindertageseinrichtungen: Eine Erweiterung des Struktur-Prozess-Modells. *Zeitschrift für Erziehungswissenschaft*, 27(2), 551–577. <https://doi.org/10.1007/s11618-024-01218-7>
- Anders, Y. & Roßbach, H.-G. (2019). Qualität in der Kindertagesbetreuung. In O. Köller, M. Hasselhorn, F. W. Hesse, K. Maaz, J. Schrader, H. Solga, C. K. Spieß & K. Zimmer (Hrsg.), *Das Bildungswesen in Deutschland. Bestand und Potenziale* (S. 441–470). Klinkhardt.
- Andersen, S. C. & Nielsen, H. S. (2020). Learning from performance information. *Journal of Public Administration Research and Theory*, 30(3), 415–431. <https://doi.org/10.1093/jopart/muz036>



- Antony-Newman, M. (2024). Parent engagement and the teaching profession: A policy framework. *School Community Journal*, 34(2), 109–128.
- Arndt, A.-K. & Werning, R. (2013). Unterrichtsbezogene Kooperation von Regelschullehrkräften und Lehrkräften für Sonderpädagogik. In R. Werning & A.-K. Arndt (Hrsg.), *Inklusion: Kooperation und Unterricht entwickeln* (S. 12–40). Klinkhardt.
- Arndt, A.-K. & Werning, R. (2016). Unterrichtsbezogene Kooperation von Regelschullehrkräften und Sonderpädagog/innen im Kontext inklusiver Schulentwicklung. Implikationen für die Professionalisierung. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. Zeitschrift für Pädagogik*, 62. Beiheft (S. 160–174). Beltz.
- Aust, L., Linker, J.-C., Eichholz, L., Schiffer, J., Nührenböcker, M., Selter, C. & Souvignier, E. (2025). Implementing formative assessment into school practice: A matter of structuring the intervention? *Journal of Teacher Education*, 77(2), 138–153. <https://doi.org/10.1177/00224871251350680>
- Autor:innengruppe Bildungsberichterstattung. (2018). *Bildung in Deutschland 2018. Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*. wbv Publikation. <https://doi.org/10.3278/6001820fw>
- Autor:innengruppe Bildungsberichterstattung. (2024). *Bildung in Deutschland 2024. Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung*. wbv Publikation. <https://doi.org/10.3278/6001820iw>
- Avenarius, H. (2007). Widersprüche bei der Steuerung des Schulwesens in Deutschland. Einige Bemerkungen aus rechtlicher Sicht. In P. Molt & H. Dickow (Hrsg.), *Kulturen und Konflikte im Vergleich. Festschrift für Theodor Hanf* (S. 875–884). Nomos.
- Avenarius, H., Ditton, H., Döbert, H., Klemm, K. & Klieme, E. (2003). *Bildungsbericht für Deutschland. Erste Befunde*. Leske + Budrich. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_01_01-Bildungsbericht-erste-Befunde.pdf
- Avenarius, H. & Füssel, H.-P. (2008). *Schulrecht im Überblick*. Wissenschaftliche Buchgesellschaft.
- Avenarius, H. & Füssel, H.-P. (2010). *Schulrecht. Ein Handbuch für Praxis, Rechtsprechung und Wissenschaft* (8., neubearb. Aufl.). Carl Link.
- Axford, N., Berry, V., Lloyd, J., Moore, D., Rogers, M., Hurst, A., Blockley, K., Durkin, H. & Minton, J. (2019). *How can schools support parents' engagement in their children's learning? Evidence from research and practice*. Education Endowment Foundation (EEF). <https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/parental-engagement/>
- Bae, S. (2018). Redesigning systems of school accountability: A multiple measures approach to accountability and support. *Education Policy Analysis Archives*, 26, 8. <https://doi.org/10.14507/epaa.26.2920>
- Baker, A., Perreault, D., Reid, A. & Blanchard, C. M. (2013). Feedback and organizations: Feedback is good, feedback-friendly culture is better. *Canadian Psychology / Psychologie canadienne*, 54(4), 260–268. <https://doi.org/10.1037/a0034691>
- Barger, M. M., Kim, E. M., Kuncel, N. R. & Pomerantz, E. M. (2019). The relation between parents' involvement in children's schooling and children's adjustment: A meta-analysis The relation between parents' involvement in children's schooling and children's adjustment: A meta-analysis. *Psychological Bulletin*, 145(9), 855–890. <https://doi.org/10.1037/bul0000201>



- Barkmann, C. & Schulte-Markwort, M. (2012). Prevalence of emotional and behavioural disorders in German children and adolescents: A meta-analysis. *Journal of Epidemiology and Community Health*, 66(3), 194–203. <https://doi.org/10.1136/jech.2009.102467>
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich*. Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiss, M. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Leske + Budrich. <https://doi.org/10.1007/978-3-322-83412-6>
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O. & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Leske + Budrich. <https://doi.org/10.1007/978-3-322-95096-3>
- Becker-Mrotzek, M., Roth, H.-J., Griebach, J., von Dewitz, N. & Schöneberger, C. (Hrsg.). (2022). *Sprachliche Bildung im Transfer: Konzepte der Sprach- und Schriftsprachförderung weitergeben*. Kohlhammer.
- Behörde für Schule und Berufsbildung Hamburg (2022). *Bildungsplan Grundschule, Stadtteilschule, Gymnasium. Allgemeiner Teil. Bildung und Erziehung an Hamburgs Schulen*. <https://www.hamburg.de/resource/blob/123132/f4ec5d1f1348e96f16dc22d676c53ea2/a-teil-dl-data.pdf>
- Behrens, U., Weirich, S. & Vohr, A. (2014). Ein Beobachtungs- und Bewertungsraster für mündliche Seminarpräsentationen in der Lehramtsausbildung. In E. Grundler & C. Spiegel (Hrsg.), *Konzeptionen des Mündlichen – Wissenschaftliche Perspektiven und didaktische Konsequenzen* (S. 186–209). hep Verlag.
- Bernasconi, T., Casale, G., Grünke, M. & Scherger, A. (2025). Interdisziplinäre Konkretisierungsbedarfe bei der Feststellung sonderpädagogischer Unterstützungsbedarfe. In J. Brandenburg, J. Gerken, A. Hußmann, J. Kuhl, O. Kunina-Habenicht, K. Limbach, F. Mörike, S. Schulze & S. Weigelt (Hrsg.), *Sonderpädagogik intra- und interdisziplinär: Perspektiven aus Theorie, Forschung und Praxis* (S. 15–26). Klinkhardt
- Besa, K.-S. & Gesang, J. (2020). Potenziale und Grenzen datengestützter Schulentwicklung an Hamburger Schulen. In M. Warmt, M. Pietsch, S. Graw-Krausholz & S. Tosana (Hrsg.), *Schulinspektion in Hamburg. Der zweite Zyklus 2012-2020: Perspektiven aus Theorie, Empirie und Praxis* (S. 115–128). Wissenschaftlicher Verlag Berlin.
- Betebner, D. W. & Linn, R. L. (2009). *Growth in student achievement: Issues of measurement, longitudinal data analysis, and accountability* [Paper presented at: Exploratory Seminar: Measurement challenges within the race to the top agenda]. <https://www.ets.org/Media/Research/pdf/BetebnerandLinnPresenterSession1.pdf>
- Betz, T. (2015). *Das Ideal der Bildungs- und Erziehungspartnerschaft. Kritische Fragen an eine verstärkte Zusammenarbeit zwischen Kindertageseinrichtungen, Grundschulen und Familien*. Bertelsmann Stiftung. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Studie_WB_Bildungs- und Erziehungspartnerschaft_2015.pdf
- Betz, T., Bischoff, S., Eunicke, N., Kayser, L. B. & Zink, K. (2017). *Partner auf Augenhöhe? Forschungsbefunde zur Zusammenarbeit von Familien, Kitas und Schulen mit Blick auf Bildungschancen*. Verlag Bertelsmann Stiftung. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/imported/leseprobe/LP_978-3-86793-816-7_1.pdf



- Betz, T., Bischoff-Pabst, S., Eunicke, N. & Menzel, B. (2019). *Kinder zwischen Chancen und Barrieren. Zusammenarbeit zwischen Kita und Familie: Perspektiven und Herausforderungen*. Bertelsmann Stiftung. <https://doi.org/10.11586/2019043>
- Bezirksregierung Düsseldorf (2024). *Leitfaden für den Umgang mit Schulabsentismus in der Bezirksregierung Düsseldorf*. https://www.brd.nrw.de/system/files/media/document/2024-05/20240514_4_41_schulabsentismus_leitfaden.pdf
- Bijlsma, H. J. E., Visscher, A. J., Dobbelaer, M. J. & Veldkamp, B. P. (2019). Does smartphone-assisted student feedback affect teachers' teaching quality? *Technology, Pedagogy and Education*, 28(2), 217–236. <https://doi.org/10.1080/1475939X.2019.1572534>
- Björk, L. G., Kowalski, T. J. & Browne-Ferrigno, T. (2014). *The school district superintendent in the United States of America*. https://ecommons.udayton.edu/cgi/viewcontent.cgi?article=1011&context=eda_fac_pub
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Black, P. & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Hrsg.), *Assessment and learning* (S. 81–100). SAGE.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Bock-Famulla, K., Berg, E., Baierl, A., Hornung, H. & Kapella, O. (2024). *Ländermonitoring Frühkindliche Bildungssysteme 2024. Profile der Bundesländer. Transparenz schaffen – Governance stärken*. Bertelsmann Stiftung. https://www.laendermonitor.de/fileadmin/files/laendermonitor/laenderprofil/2024/RP_Laenderprofil_2024.pdf
- Böhm-Kasper, O., Selders, O. & Lambrecht, M. (2016). Schulinspektion und Schulentwicklung – Ergebnisse der quantitativen Schulleitungsbefragung. In O. Böhm-Kasper, T. Brüsemeister, F. Dietrich, L. Gromala, M. Heinrich, M. Lambrecht, B. Preuß, M. Rürup, O. Selders & J. Wissinger (Hrsg.), *Schulinspektion als Steuerungsimpuls? Ergebnisse aus Forschungsprojekten* (S. 1–50). Springer VS.
- Bokhove, C., Jerrim, J. & Sims, S. (2023). Are some school inspectors more lenient than others? *School Effectiveness and School Improvement*, 34(4), 419–441. <https://doi.org/10.1080/09243453.2023.2240318>
- Bonanati, M. (2015). Selbsteinschätzung in Lehrer-Schüler-Eltern-Gesprächen. In H. de Boer & M. Bonanati (Hrsg.), *Gespräche über Lernen – Lernen im Gespräch* (S. 177–193). Springer VS. https://doi.org/10.1007/978-3-658-09696-0_10
- Bonsen, M. & Frey, K. A. (2024). Qualitätsentwicklung in Schulen. In T. Betz, T. Feldhoff, P. Bauer, U. Schmidt & B. Schmidt-Hertha (Hrsg.), *Handbuch Qualität in pädagogischen Feldern. Diskurse. Theoretische Grundlagen. Empirische Beiträge. Kritische Einwürfe* (S. 1–19). Springer VS. https://doi.org/10.1007/978-3-658-40471-0_14-1
- Boonk, L., Gijsselaers, H. J., Ritzen, H. & Brand-Gruwel, S. (2018). A review of the relationship between parental involvement indicators and academic achievement. *Educational Research Review*, 24, 10–30. <https://doi.org/10.1016/J.EDUREV.2018.02.001>
- Bormann, I., Killus, D., Niedlich, S. & Würbel, I. (2022). Home-school interaction: A vignette study of parents' views on situations relevant to trust. *European Education*, 53(3-4), 137–151. <https://doi.org/10.1080/10564934.2022.2081084>



- Born, A., Klaudy, E. K., Micheel, B., Risse, T. & Stöbe-Blossey, S. (2019). *Familienzentren an Grundschulen. Abschlussbericht zur Evaluation in Gelsenkirchen* (IAQ-Forschung 2019-04). Universität Duisburg-Essen. <https://doi.org/10.17185/dupublico/49074>
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R. & Walther, G. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Waxmann.
- Brandenburg, J., Gerken, J., Hußmann, A., Kuhl, J., Kunina-Habenicht, O., Limbach, K., Mörike, F., Schulze, S. & Weigelt, S. (Hrsg.). (2025). *Sonderpädagogik intra- und interdisziplinär: Perspektiven aus Theorie, Forschung und Praxis*. Klinkhardt.
- Brauckmann, S. (2012). Schulleitungshandeln zwischen deconcentration, devolution und delegation (3D) – empirische Annäherungen aus internationaler Perspektive. *Empirische Pädagogik*, 26(1), 78–102. <https://doi.org/10.25656/01:5872>
- Breiter, A. & Light, D. (2006). Data for school improvement: Factors for designing effective information systems to support decision-making in schools. *Educational Technology & Society*, 9(3), 206–217. <http://www.jstor.org/stable/jeductechsoci.9.3.206>
- Bromme, R. & Prenzel, M. (Hrsg.). (2014). *Von der Forschung zur evidenzbasierten Entscheidung. Die Darstellung und das öffentliche Verständnis der empirischen Bildungsforschung*. Zeitschrift für Erziehungswissenschaft, 27. Sonderheft. Springer VS.
- Bronnert, H. (2024). *Führung, Schulkultur und Elternarbeit. Eine gemischtmethodische Analyse der Zusammenhänge* [Dissertation]. Technische Universität Dortmund. <https://doi.org/10.17877/DE290R-25721>
- Brookhart, S. (2011). *Grading and learning: Practices that support student achievement*. Solution Tree Press.
- Brown, M., Gardezi, S., del Castillo Blanco, L., Simeonova, R., Parvanova, Y., McNamara, G., O'Hara, J. & Kechri, Z. (2021). School self-evaluation an international or country specific imperative for school improvement? *International Journal of Educational Research Open*, 2, Artikel 100063. <https://doi.org/10.1016/j.ijedro.2021.100063>
- Burkard, C. (2004). Funktionen und Schwerpunkte von Schulprogrammen aus Sicht der Schulaufsicht. In H. G. Holtappels (Hrsg.), *Schulprogramme – Instrumente der Schulentwicklung. Konzeptionen, Forschungsergebnisse, Praxisempfehlungen* (S. 137–153). Juventa.
- Burkhardt, H. & Schoenfeld, A. (2018). Assessment in the service of learning: challenges and opportunities or Plus ça Change, Plus c'est la même Chose. *ZDM – Mathematics Education*, 50(4), 571–585. <https://doi.org/10.1007/s11858-018-0937-1>
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79(4), 474–482. <https://doi.org/10.1037/0022-0663.79.4.474>
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58(1), 1–14. <https://doi.org/10.1111/j.2044-8279.1988.tb00874.x>
- Buysse, V. & Peisner-Feinberg, E. (2010). Recognition & Response: Response to intervention for PreK. *Young Exceptional Children*, 13(4), 2–13. <https://doi.org/10.1177/1096250610373586>



- Buysse, V., Peisner-Feinberg, E., Soukakou, E., Fettig, A., Schaaf, J. & Burchinal, M. (2016). Using Recognition & Response (R&R) to improve children's language and literacy skills: Findings from two studies. *Early Childhood Research Quarterly*, 36(3), 11–20. <https://doi.org/10.1016/j.ecresq.2015.11.005>
- Caputo, A. & Rastelli, V. (2014). School improvement plans and student achievement: Preliminary evidence from the Quality and Merit Project in Italy. *Improving Schools*, 17(1), 72–98. <https://doi.org/10.1177/1365480213515800>
- Carstensen, B. & Klusmann, U. (i. E.). Eine Analyse des Fort- und Weiterbildungsangebots zur Implementation sozial-emotionaler Lernziele im deutschen Schulsystem.
- Casale, G., Herzog, M. & Volpe, R. J. (2023). Measurement Efficiency of a Teacher Rating Scale to Screen for Students at Risk for Social, Emotional, and Behavioral Problems. *Journal of Intelligence*, 11(3). <https://doi.org/10.3390/jintelligence11030057>
- Casale, G., Hövel, D. C., Hennemann, T. & Hillenbrand, C. (2018). Prävention und psychische Gesundheitsförderung in der Schule. In H. Christiansen, D. Ebert & B. Röhrle (Hrsg.), *Prävention und psychische Gesundheitsförderung Bd. VI. Entwicklungen und Perspektiven* (Bd. 24, S. 245–286). DGVT-Verlag.
- Chamakalayil, L., Ivanova-Chessex, O., Leutwyler, B. & Scharathow, W. (Hrsg.). (2022). *Eltern und pädagogische Institutionen. Macht- und ungleichheitskritische Perspektiven*. Beltz Juventa. <https://doi.org/10.25656/01:29141>
- Chapman, C. & Sammons, P. (2013). *School self-evaluation for school improvement: what works and why?* <https://eric.ed.gov/?id=ed546801>
- Charalambous, C. Y. & Praetorius, A.-K. (2025). Are we doing okay? Making the case for critically examining small and inconsistent effects in research on teaching quality. *School Effectiveness and School Improvement*, 36(2), 153–163. <https://doi.org/10.1080/09243453.2025.2482574>
- Chernyshenko, O. S., Kankaraš, M. & Drasgow, F. (2018). *Social and emotional skills for student success and well-being: Conceptual framework for the OECD study on social and emotional skills. OECD Education Working Papers No. 173*. <https://doi.org/10.1787/db1d8e59-en>
- Choi, Á. (2019). *Education inspection: which models work best?* Fundació Jaume Bofill; Institut Català d'Avaluació de Polítiques Públiques. https://ivalua.cat/sites/default/files/2019-11/29_01_2019_15_37_55_QUE_FUNCIONA_13_ANGLES.PDF
- Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J. C. & Durlak, J. A. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. *Child Development*, 94(5), 1181–1204. <https://doi.org/10.1111/cdev.13968>
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität*. Waxmann.
- Confrey, J. & Maloney, A. (2015). A design research study of a curriculum and diagnostic assessment system for a learning trajectory on equipartitioning. *ZDM – Mathematics Education*, 47(6), 919–932. <https://doi.org/10.1007/s11858-015-0699-y>
- Copeland, M. (2016). Data-driven leadership: Determining your indicators and building your dashboards. *Independent School*, 76(1), 35–40.
- Corcoran, R. P., Cheung, A. C. K., Kim, E. & Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educational Research Review*, 25, 56–72. <https://doi.org/10.1016/j.edurev.2017.12.001>



- Council of Ministers of Education (2026). *PCAP 2023. Report on the Pan-Canadian Assessment of Science, Reading, and Mathematics*. https://www.cmec.ca/Publications/Lists/Publications/Attachments/447/PCAP2023_Public_report_FINAL_EN.pdf
- Cremers-van Wees, L. M. C. M., Rekveld, I., Brandsma, H. P. & Bosker, R. J. (1996). *Instrumenten voor kwaliteitszorg: Beschrijving van 31 instrumenten*. University of Twente.
- Curran, F. C., Carlo, S. & Harris-Walls, K. (2024). Making the data visible: A systematic review of systems-level data dashboards for leadership and policy in education. *Review of Educational Research*, 96(1), 3–52. <https://doi.org/10.3102/00346543241288249>
- Dalehefte, I. M., Wendt, H., Köller, O., Wagner, H., Pietsch, M., Döring, B., Fischer, C. & Bos, W. (2014). Bilanz von neun Jahren SINUS an Grundschulen in Deutschland. Evaluation der mathematikbezogenen Daten im Rahmen von TIMSS 2011. *Zeitschrift für Pädagogik*, 60(2), 245–263. <https://doi.org/10.25656/01:14656>
- Danzig, A. B., Borman, K. M., Jones, B. A. & Wright, W. F. (Hrsg.). (2013). *Learner-centered leadership. Research, policy, and practice*. Routledge.
- Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational researcher*, 44(2), 132–137. <https://doi.org/10.3102/0013189X15575346>
- Data Quality Campaign (2008). *What does role-based access look like? Examples from states*. <https://files.eric.ed.gov/fulltext/ED539324.pdf>
- Data Quality Campaign (2019). *Growth data: It matters, and it's complicated*. <https://dataqualitycampaign.org/wp-content/uploads/2019/04/DQC-Growth-Data-Resources.pdf>
- Datnow, A. & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change*, 17(1), 7–28. <https://doi.org/10.1007/s10833-015-9264-2>
- de Hoyos, R., Garcia-Moreno, V. A. & Patrinos, H. A. (2017). The impact of an accountability intervention with diagnostic feedback: Evidence from Mexico. *Economics of Education Review*, 58, 123–140. <https://doi.org/10.1016/j.econedurev.2017.03.007>
- Dearing, E., Kreider, H., Simpkins, S. & Weiss, H. B. (2006). Family involvement in school and low-income children's literacy: Longitudinal associations between and within families. *Journal of Educational Psychology*, 98(4), 653–664. <https://doi.org/10.1037/0022-0663.98.4.653>
- Dedering, K., Kamarianakis, E. & Racherbäumer, K. (2022). Schulentwicklungsberatung. Begrifflich-konzeptionelle Grundlegung, empirische Betrachtung und (kritische) Perspektivierung. *Die Deutsche Schule*, 114(4), 345–362. <https://doi.org/10.31244/dds.2022.04.02>
- Dedering, K., Tillmann, K.-J., Goecke, M. & Rauh, M. (2013). *Wenn Experten in die Schule kommen. Schulentwicklungsberatung – empirisch betrachtet*. Springer VS. <https://doi.org/10.1007/978-3-658-01402-5>
- Demetriou, D. & Kyriakides, L. (2012). The impact of school self-evaluation upon student achievement: A group randomisation study. *Oxford Review of Education*, 38(2), 149–170. <https://doi.org/10.1080/03054985.2012.666032>
- Demski, D. (2017). *Evidenzbasierte Schulentwicklung. Empirische Analyse eines Steuerungsparadigmas*. Springer VS. <https://doi.org/10.1007/978-3-658-18078-2>



- Demski, D. (2021). Vergleichsarbeiten – Unterstützung für Schulen oder Kontrollinstrument? In T. Webs & V. Manitius (Hrsg.), *Unterstützungssysteme für Schulen. Konzepte, Befunde und Perspektiven* (S. 143–160). wbv Publikation.
- Dent, A. L. & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, 28(3), 425–474. <https://doi.org/10.1007/s10648-015-9320-8>
- Detzel, J. (2024). *Vergleichsarbeiten (VERA): Eine empirische Untersuchung zur Akzeptanz bei Lehrkräften* [Dissertation]. Rheinland-Pfälzische Universität Kaiserslautern-Landau.
- Deutscher Bundestag – Wissenschaftliche Dienste (2020). *Schulautonomie in den Landesgesetzen (WD 8 - 3000 - 018/20)*. Deutscher Bundestag. <https://www.bundestag.de/resource/blob/691806/3577812bfaofa54206d9e4efd255c526/WD-8-018-20-pdf-data.pdf>
- Deutsches Jugendinstitut (2011). *Zusammenarbeit mit Eltern. Grundlagen für die kompetenzorientierte Weiterbildung*. (WiFF Wegweiser Weiterbildung Nr. 3). <https://doi.org/10.25656/01:28570>
- Dickhäuser, O. & Rheinberg, F. (2003). Bezugsnormorientierung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (2. Auflage, S. 41–48). Beltz.
- Dittmann, E., Metzdorf, A., Müller, H. & Schmolke, R. (2018). *Qualifizierte Kooperation von Jugendhilfe und Schule. Eine Arbeitshilfe für die Praxis*. (Überarbeitete Neuauflage). Institut für Sozialpädagogische Forschung Mainz (ism). https://www.ism-mz.de/fileadmin/uploads/Publikationen/Qualifizierte_Kooperation_von_Jugendhilfe_und_Schule_2018.pdf
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule. Zeitschrift für Pädagogik*, 41. Beiheft (73-92). Beltz. <https://doi.org/10.25656/01:8486>
- Dollinger, S. (2019). Lernunterstützende Rückmeldung in Lernentwicklungsgesprächen. *Zeitschrift für Grundschulforschung*, 12(1), 197–212. <https://doi.org/10.1007/s42278-019-00040-z>
- Dollinger, S. & Hartinger, A. (2020). *Lernentwicklungsgespräche in der Grundschule. Erprobte Praxisbausteine*. Cornelsen.
- Donker, A. S., de Boer, H, Kostons, D., Dignath van Ewijk, C. C. & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1–26. <https://doi.org/10.1016/j.edurev.2013.11.002>
- Doss, C., Fahle, E. M., Loeb, S. & York, B. N. (2019). More than just a nudge: Supporting kindergarten parents with differentiated and personalized text-messages. *Journal of Human Resources*, 54(3), 567–603. <https://doi.org/10.3368/jhr.54.3.0317-8637R>
- Drachsler, H. (2023). *Towards highly informative learning analytics*. Open Universiteit Heerlen. <https://doi.org/10.25656/01:26787>
- Dreher, A. & Leuders, T. (2021). Fachspezifität von Unterrichtsqualität – aus der Perspektive der Mathematikdidaktik. *Unterrichtswissenschaft*, 49(2), 285–292. <https://doi.org/10.1007/s42010-021-00116-9>
- Drinck, B., Diegmann, D., Schmidt, M. & Flagmeyer, D. (2013). *RuN. Rezeption und Nutzung von Ergebnissen der externen Evaluation an sächsischen Grundschulen, Mittelschulen und Gymnasien*. Abschlussbericht. Sächsisches Bildungsinstitut.



- Duckworth, A. L. & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational researcher*, 44(4), 237–251. <https://doi.org/10.3102/0013189X15584327>
- Dumont, H., Trautwein, U., Nagy, G. & Nagengast, B. (2014). Quality of parental homework involvement: Predictors and reciprocal relations with academic functioning in the reading domain. *Journal of Educational Psychology*, 106(1), 144–161. <https://doi.org/10.1037/a0034100>
- Dunaway, D. M., Kim, D.-H. & Szad, E. R. (2012). Perceptions of the purpose and value of the school improvement plan process. *The Educational Forum*, 76(2), 158–173. <https://doi.org/10.1080/00131725.2011.652490>
- Durlak, J. A., Mahoney, J. L. & Boyle, A. E. (2022). What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research. *Psychological Bulletin*, 148(11-12), 765–782. <https://doi.org/10.1037/bul0000383>
- Dusenbury, L., Calin, S., Domitrovich, C. & Weissberg, R. P. (2015). *What does evidence-based instruction in social and emotional learning actually look like in practice? A brief on findings from CASEL's program reviews*. Collaborative for Academic, Social, and Emotional Learning (CASEL). https://drc.case-l.org/uploads/sites/3/2019/02/What-Does-Evidence-Based-Instruction-in-Social-and-Emotional-Learning-Actually-Look-Like-in-Practice_.pdf
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
- Earl, L. M. & Katz, S. (2006). *Leading schools in a data-rich world. Harnessing data for school improvement*. Corwin Press.
- Education Endowment Foundation (2018). *Working with parents to support children's learning: Guidance report*. <https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/supporting-parents>
- Egert, F., Körner, F. & Sachse, S. (2025). Was passiert in Kindertageseinrichtungen, wenn ein Kind sprach-auffälliges Verhalten zeigt? Explorative Studie zur professionellen, zirkulären Fallarbeit. *Frühe Bildung*, 14(2), 79–86. <https://doi.org/10.1026/2191-9186/a000708>
- Ehren, M. & Bachmann, R. (2020). Accountability to build school and system improvement capacity. In M. Ehren & J. Baxter (Hrsg.), *Trust, accountability and capacity in education system reform. Global Perspectives in Comparative Education* (S. 102–123). Routledge. <https://doi.org/10.4324/9780429344855-5>
- Ehren, M. & Shackleton, N. (2016a). Mechanisms of change in Dutch inspected schools: Comparing schools in different inspection treatments. *British Journal of Educational Studies*, 64(2), 185–213. <https://doi.org/10.1080/00071005.2015.1019413>
- Ehren, M. & Shackleton, N. (2016b). Risk-based school inspections: Impact of targeted inspection approaches on Dutch secondary schools. *Educational Assessment, Evaluation and Accountability*, 28(4), 299–321. <https://doi.org/10.1007/s11092-016-9242-0>
- Elsing, S. & van Ackeren, I. (2017). Orientierungsrahmen zur Schulqualität im nationalen Vergleich. Eine deskriptive Sichtung unter besonderer Berücksichtigung der Wirkungsdimension und ausgewählter internationaler Ansätze. In P. Dobbstein, B. Groot-Wilken & S. Koltermann (Hrsg.), *Referenzsysteme zur Unterstützung von Schulentwicklung* (S. 25–62). Waxmann.
- Emslander, V., Levy, J., Scherer, R. & Fischbach, A. (2022). Value-added scores show limited stability over time in primary school. *PLoS ONE*, 17(12), Artikel e0279255. <https://doi.org/10.1371/journal.pone.0279255>



- Enderle, T. & Vollmar, M. (2019). Geheimhaltung in der Hochschulstatistik. *WISTA – Wirtschaft und Statistik*(6), 87–99. https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2019/06/geheimhaltung-hochschulstatistik-062019.pdf?__blob=publicationFile
- Epstein, J. L. (2018). School, family, and community partnerships in teachers' professional work. *Journal of Education for Teaching*, 44(3), 397–406. <https://doi.org/10.1080/02607476.2018.1465669>
- Erdem, C. & Kaya, M. (2020). A meta-analysis of the effect of parental involvement on students' academic achievement. *Journal of Learning for Development*, 7(3), 367–383. <https://doi.org/10.56059/jl4d.v7i3.417>
- Ertl, S., Klippel, E. & Hartinger, A. (2021). Lernentwicklungsgespräche qualitativ gestalten. In G. Stückl & M. Wilhelm (Hrsg.), *Lehren und Lernen in der bayerischen Grundschule. Kommentare und Unterrichtshilfen zum LehrplanPLUS Grundschule. Loseblatt-Ausgabe* (36.05 1-12). Carl Link.
- Ertl, S., Kücherer, B. & Hartinger, A. (2022). Lernentwicklungsgespräche und die Entwicklung motivationaler Aspekte des Lernens. *Zeitschrift für Grundschulforschung*, 15(1), 221–236. <https://doi.org/10.1007/s42278-021-00134-7>
- Ertl, S., Weißenhorn, S., Kücherer, B. & Hartinger, A. (2022). Qualität von Zielvereinbarungen in Lernentwicklungsgesprächen. In E. Gläser, J. Poschmann, P. Büker & S. Miller (Hrsg.), *Reflexion und Reflexivität im Kontext Grundschule. Perspektiven für Forschung, Lehrer:innenbildung und Praxis* (S. 155–162). Klinkhardt. <https://doi.org/10.25656/01:25562>
- Escobar, I. H. G. (2019). School improvement plans, a tool to improve the quality of education. *New Trends and Issues Proceedings on Humanities and Social Sciences*, 6(1), 440–450. <https://doi.org/10.18844/prosoc.v6i1.4197>
- Ettinger, R. S. (2015). *Shifting from a plan to a process: School improvement plans in the Cambridge public schools* [Dissertation]. Harvard Graduate School of Education. <https://dash.harvard.edu/server/api/core/bitstreams/7312037d-8403-6bd4-e053-0100007fdf3b/content>
- Europäische Union (2016). *Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung)*. *Amtsblatt der Europäischen Union* L 119. https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=serv%3AOJ.L_.2016.119.01.0001.01.DEU&toc=OJ%3AL%3A2016%3A119%3ATOC
- Eyles, A. & Machin, S. (2019). The introduction of academy schools to England's education. *Journal of the European Economic Association*, 17(4), 1107–1146. <https://doi.org/10.1093/jeea/jvy021>
- Fabel-Lamla, M. (2022). Schule. In M. K. W. Schweer (Hrsg.), *Facetten des Vertrauens und Misstrauens. Herausforderungen für das soziale Miteinander* (S. 175–200). Springer VS. https://doi.org/10.1007/978-3-658-29047-4_10
- Faddar, J., Vanhoof, J. & de Maeyer, S. (2017). Instruments for school self-evaluation: Lost in translation? A study on respondents' cognitive processing. *Educational Assessment, Evaluation and Accountability*, 29(4), 397–420. <https://doi.org/10.1007/s11092-017-9270-4>
- Farrell, C. C. (2015). Designing school systems to encourage data use and instructional improvement. *Educational Administration Quarterly*, 51(3), 438–471. <https://doi.org/10.1177/0013161X14539806>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>



- Fauth, B., Wagner, W., Bertram, C., Göllner, R., Roloff, J., Lüdtke, O., Polikoff, M. S., Klusmann, U. & Trautwein, U. (2020). Don't blame the teacher? The need to account for classroom characteristics in evaluations of teaching quality. *Journal of Educational Psychology*, 112(6), 1284–1302. <https://doi.org/10.1037/edu0000416>
- Fend, H. (1986). „Gute Schulen - schlechte Schulen“. Die einzelne Schule als paedag. Handlungseinheit. *Die Deutsche Schule*, 78(3), 275–293.
- Fend, H. (1998). *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung*. Juventa.
- Fernandez, K. E. (2011). Evaluating school improvement plans and their affect on academic performance. *Educational Policy*, 25(2), 338–367. <https://doi.org/10.1177/0895904809351693>
- Fiege, C. (2013). *Faire Vergleiche in der Schulleistungsforschung – Methodologische Grundlagen und Anwendung auf Vergleichsarbeiten* [Dissertation]. Friedrich-Schiller-Universität Jena. https://www.db-thueringen.de/servlets/MCRFileNodeServlet/dbt_derivate_00028843/Diss/ChristianeFiege.pdf
- Fiege, C. (2016). Faire Vergleiche bei Vergleichsarbeiten: Möglichkeiten und Grenzen. In B. Groot-Wilken, K. Isaac & J.-P. Schräpler (Hrsg.), *Sozialindices für Schulen. Hintergründe, Methoden und Anwendung* (S. 71–96). Waxmann.
- Fiege, C., Reuther, F. & Nachtigall, C. (2011). Faire Vergleiche? – Berücksichtigung von Kontextbedingungen des Lernens beim Vergleich von Testergebnissen aus deutschen Vergleichsarbeiten. *Zeitschrift für Bildungsforschung*, 1(2), 133–149. <https://doi.org/10.1007/s35834-011-0009-x>
- Figlio, D. & Loeb, S. (2011). School accountability. In E. A. Hanushek, S. Machin & L. Wößmann (Hrsg.), *Handbook of the economics of education* (Bd. 3, S. 383–421). North-Holland.
- Filderman, M. J., Toste, J. R., Didion, L. & Peng, P. (2022). Data literacy training for K–12 teachers: A meta-analysis of the effects on teacher outcomes. *Remedial and Special Education*, 43(5), 328–343. <https://doi.org/10.1177/07419325211054208>
- Filiasov, S. & Sweetman, A. (2023). Low-stakes standardized tests in British Columbia, Canada: system accountability and/or individual feedback? *Education Economics*, 31(2), 145–165. <https://doi.org/10.1080/09645292.2022.2091113>
- Forell, M. (2023). Zur theoretischen Verfasstheit des schulischen Sozialraums. Kartierungen sozialräumlicher Dimensionen von Schule. In M. Forell, G. Bellenberg, L. Gerhards & L. Schleenbecker (Hrsg.), *Schule als Sozialraum im Sozialraum. Theoretische und empirische Erkundung sozialräumlicher Dimensionen von Schule* (S. 13–25). Waxmann.
- Forell, M. & Schuchardt, J. (o. D.). *Mikrokosmos schulischer Nahraum: Sozialraumanalyse mit der ALSO-App*. Zugriff am 29. Januar 2026, verfügbar unter <https://www.kommunales-bildungsmonitoring.de/themen/mikrokosmos-sozialraum/interview-sozialraumanalyse-mit-der-also-app>
- Förster, N., Forthmann, B., Urbach, O. & Souvignier, E. (2024). Lernverlaufsdiagnostik im Lesen mit quop. In T. Richter & W. Lenhard (Hrsg.), *Diagnose und Förderung des Lesens im digitalen Kontext* (S. 117–136). Hogrefe.
- Förster, N. & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, 32, 91–100. <https://doi.org/10.1016/j.learninstruc.2014.02.002>



- Friederich, T. (2011). *Zusammenarbeit mit Eltern – Anforderungen an frühpädagogische Fachkräfte. Eine Expertise der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF)* (WiFF Expertisen Nr. 22). Deutsches Jugendinstitut. https://www.weiterbildungsinitiative.de/fileadmin/Redaktion/Publikationen/WiFF_Expertise_Friederich.pdf
- Fuchs, A., Radkowitz, A. & Sommerhoff, D. (2025). Using learning progress monitoring to promote academic performance? A meta-analysis of the effectiveness. *Educational Research Review*, 46, Artikel 100648. <https://doi.org/10.1016/j.edurev.2024.100648>
- Gal, I. (2004). Statistical literacy. Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Hrsg.), *The challenge of developing statistical literacy, reasoning and thinking* (S. 47–78). Kluwer Academic Publisher. https://doi.org/10.1007/1-4020-2278-6_3
- Gallimore, R., Ermeling, B., Saunders, W. & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *The Elementary School Journal*, 109(5), 537–553. <https://doi.org/10.1086/597001>
- Gärtner, H. (2016). Schulinspektion – Zum gegenwärtigen und zukünftigen Verhältnis von Inspektionsforschung und -praxis. In Bundesministerium für Bildung und Forschung (Hrsg.), *Bildungsforschung 2020. Zwischen wissenschaftlicher Exzellenz und gesellschaftlicher Verantwortung. Tagung des Bundesministeriums für Bildung und Forschung vom 27. bis 28. März 2014 in Berlin* (S. 62–74). https://www.bmbf.de/SharedDocs/Publikationen/DE/3/31184_Bildungsforschung_Band_42.pdf?__blob=publicationFile&v=5
- Gärtner, H. (2021). Wie kann externe Evaluation/Schulinspektion zur Schulentwicklung beitragen. In T. Webs & V. Manitus (Hrsg.), *Unterstützungssysteme für Schulen. Konzepte, Befunde und Perspektiven* (S. 177–192). wbv Publikation.
- Gärtner, H. (2025). Nutzung von Lernstandserhebungen zur Förderung adaptiven Unterrichts. In A. Albers & N. Jude (Hrsg.), *Blickpunkt Bildungsmonitoring – Bilanzen und Perspektiven* (S. 142–157). Beltz Juventa.
- Gärtner, H. & Brauckmann, S. (2014, 10. März). *Professionalisierung von Schulleitungshandeln zwischen Tradition und Innovation*. Institut für Schulqualität der Länder Berlin und Brandenburg e. V. (ISQ). DGfE-Kongress, Humboldt Universität zu Berlin. https://www.researchgate.net/profile/holger-gaertner/publication/276936398_gartner_h_brauckmann_s_2014_professionalisierung_von_schulleitungshandeln_zwischen_tradition_und_innovation_dgfe_berlin/links/555c5ce508ae6aeao817523e/gaertner-h-brauckmann-s-2014-professionalisierung-von-schulleitungshandeln-zwischen-tradition-und-innovation-dgfe-berlin.pdf
- Gärtner, H., Thiel, F., Kellermann, C. & Nachbauer, M. (2022). Ein theoriebasierter Schülerfragebogen für Unterrichtsevaluation. *Journal for Educational Research Online*, 14(1), 147–173. <https://doi.org/10.25656/01:24876>
- Gärtner, H., Thiel, F., Nachbauer, M. & Kellermann, C. (2021). Unterrichtsentwicklung durch Unterrichtsfeedback. Erste Entwicklungsschritte eines Beobachtungsbogens für Schulleitungen. *Zeitschrift für Bildungsforschung*, 11(3), 529–547. <https://doi.org/10.25656/01:27696>
- Gärtner, H., Wurster, S. & Pant, H. A. (2014). The effect of school inspections on school improvement. *School Effectiveness and School Improvement*, 25(4), 489–508. <https://doi.org/10.1080/09243453.2013.811089>
- Geary, D. C., Nicholas, A., Li, Y. & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology*, 109(5), 680–693. <https://doi.org/10.1037/edu0000159>



- Gebhardt, M. (2023). *Pädagogische Diagnostik. Leistung, Kompetenz und Entwicklung messen, bewerten und für individuelle Förderung interpretieren*. Münchner Beiträge zur Inklusions- und Sonderpädagogik, Vol. 3, Version 0.3. Ludwig-Maximilians-Universität München. <https://doi.org/10.5282/UBM/EPUB.110013>
- Gehlbach, H. & Hough, H. (2018). *Measuring social emotional learning through student surveys in the CORE districts: A pragmatic approach to validity and reliability*. PACE. Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/SEL_VValidity_May-2018.pdf
- Gerecht, M., Steinert, B., Klieme, E. & Döbrich, P. (2007). *Skalen zur Schulqualität. Dokumentation der Erhebungsinstrumente. Pädagogische Entwicklungsbilanzen mit Schulen (PEB)*. Gesellschaft zur Förderung Pädagogischer Forschung; Deutsches Institut für Internationale Pädagogische Forschung. <https://doi.org/10.25656/01:3121>
- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R. & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/docs/practice-guide/rti_math_pg_042109.pdf
- Gesel, S. A., LeJeune, L. M., Chow, J. C., Sinclair, A. C. & Lemons, C. J. (2021). A meta-analysis of the impact of professional development on teachers' knowledge, skill, and self-efficacy in data-based decision-making. *Journal of learning disabilities*, 54(4), 269–283. <https://doi.org/10.1177/0022219420970196>
- Giar, K., Hohlstein, F., Wipke, M. & Scharnagl, A. (2023). Konzeption eines statistischen Bildungsverlaufsregisters in Deutschland: Entwicklungen bis 2023 und Ausgestaltungsoptionen. *WISTA – Wirtschaft und Statistik*(3), 51–62. https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2023/03/konzeption-bildungsverlaufsregister-032023.pdf?__blob=publicationFile&v=1
- Göllner, R., Wagner, W., Klieme, E., Lüdtke, O., Nagengast, B. & Trautwein, U. (2016). Erfassung der Unterrichtsqualität mithilfe von Schülerurteilen: Chancen, Grenzen und Forschungsperspektiven. In Bundesministerium für Bildung und Forschung (Hrsg.), *Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments* (S. 63–82). <https://doi.org/10.25656/01:12674>
- Goodall, J. & Montgomery, C. (2014). Parental involvement to parental engagement: A continuum. *Educational Review*, 66(4), 399–410. <https://doi.org/10.1080/00131911.2013.781576>
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5), 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Gordon, R. (1987). An operational classification of disease prevention. In J. A. Steinberg & M. M. Silverman (Hrsg.), *Preventing mental disorders. A research perspective* (S. 20–26). National Institute of Mental Health.
- Gore, J., Miller, D., Fray, L. & Patfield, S. (2023). *Building capacity for quality teaching in Australian schools 2018-2023*. Final Report to the Paul Ramsay Foundation. University of Newcastle, Teachers and Teaching Research Centre. https://www.newcastle.edu.au/_data/assets/pdf_file/0008/956231/Building-Capacity-for-Quality-Teaching-in-Australian-Schools-Final-Report-2023.pdf
- Gottlebe, K., Dietrich, S., Latzko, B. & Liebers, K. (2020). *Bericht zur Normierung der Aufgabenpakete ILeA plus Deutsch im Schuljahr 2018/19*. Universität Leipzig. <https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-711284>



- Government of Alberta. (2023). *Superintendent leadership quality standard*. Alberta Education. <https://open.alberta.ca/dataset/3900e9e4-7ddf-4dad-9234-6b8307eb9e8/resource/2b861b35-59e6-48b9-9638-8fd6c29cdco4/download/educ-superintendent-leadership-quality-standard-2023.pdf>
- Grabarek, J. & Kallemeyn, L. M. (2020). Does teacher data use lead to improved student achievement? A review of the empirical evidence. *Teachers College Record: The Voice of Scholarship in Education*, 122(12), 1–42. <https://doi.org/10.1177/016146812012201201>
- Graham, S., Hebert, M. & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, 115(4), 523–547. <https://doi.org/10.1086/681947>
- Graham-Clay, S. (2024). Communicating with parents 2.0: Strategies for teachers. *School Community Journal*, 34(1), 9–60. <https://www.adi.org/journal/SS2024/Graham-Clay1.pdf>
- Grob, U., Maag Merki, K. & Büeler, X. (2003). Young adult survey. Theoretische Begründung und empirische Befunde zur Validierung eines Indikatorensystems zu überfachlichen Kompetenzen. *Schweizerische Zeitschrift für Bildungswissenschaften*, 25(2), 309–330. <https://doi.org/10.25656/01:3666>
- Gross, S., Hankeln, C., Rösike, K.-A. & Prediger, S. (2025). How do expert and novice teachers monitor and enhance student understanding? Qualitative comparisons informing the design of a digital formative assessment platform. *Technology, Knowledge and Learning*, 30(2), 991–1020. <https://doi.org/10.1007/s10758-024-09755-0>
- Groß Ophoff, J., Koch, U. & Hosenfeld, I. (2019). Vergleichsarbeiten in der Grundschule von 2004 bis 2015: Trends in der Akzeptanz und Auseinandersetzung mit Rückmeldungen. In J. Zuber, H. Altrichter & M. Heinrich (Hrsg.), *Bildungsstandards zwischen Politik und schulischem Alltag* (S. 205–228). Springer VS.
- Gruber, M., Höfig, C., Golla, M., Urban, T. & Große-Kampmann, M. (2022). “We may share the number of diaper changes”: A privacy and security analysis of mobile child care applications. *Proceedings on Privacy Enhancing Technologies*, 2022(3), 394–414. <https://doi.org/10.56553/popets-2022-0078>
- Grüter, S. (2025). *Eltern inklusive? Studien zur Zusammenarbeit von inklusiven Sekundar- und Gesamtschulen und Eltern von Schüler:innen mit und ohne sonderpädagogischen Unterstützungsbedarf* [Rahmenschrift zur kumulativen Dissertation]. Universität Bielefeld. <https://doi.org/10.4119/unibi/3002769>
- Grützmacher, L., Herbert, B., Holzberger, D., Naumann, A., Steffensky, M. & Vieluf, S. (2025). Critical examination of the measurement quality of student outcomes: A systematic review. *School Effectiveness and School Improvement*, 36(2), 211–241. <https://doi.org/10.1080/09243453.2025.2482579>
- Guskey, T. R. & Jung, L. A. (2013). *Answers to essential questions about standards, assessments, grading, and reporting*. Corwin.
- Häbig, J. (2018). *Lernentwicklungsgespräche aus der Sicht von Schülerinnen und Schülern. Eine Mixed-Methods-Studie an Gymnasien*. Springer VS. <https://doi.org/10.1007/978-3-658-21632-0>
- Hahn, S., Schulte, K., Schulze, P. & Thonke, F. (2025). Welche Story steckt in den Daten einer Schule? Das Format der Datenkonferenzen als multiprofessionelles Setting zur Standortbestimmung einer Schule. In A. Albers & N. Jude (Hrsg.), *Blickpunkt Bildungsmonitoring – Bilanzen und Perspektiven* (S. 199–212). Beltz Juventa.
- Hamilton, L., Halverson, R., Jackson, S., Mandinach, E. B., Supovitz, J. & Wayman, J. (2009). *Using student achievement data to support instructional decision making*. National Center for Education Evaluation and Regional Assistance. <https://files.eric.ed.gov/fulltext/ED506645.pdf>



- Hankeln, C., Kroehne, U., Voss, L., Gross, S. & Prediger, S. (2025). Developing digital formative assessment for deep conceptual learning goals: Which topic-specific research gaps need to be closed? *Educational Technology Research and Development*, 73(4), 1953–1973. <https://doi.org/10.1007/s11423-025-10486-x>
- Harks, B., Rakoczy, K., Hattie, J., Besser, M. & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: The role of feedback's perceived usefulness. *Educational Psychology*, 34(3), 269–290. <https://doi.org/10.1080/01443410.2013.785384>
- Hattie, J. (2008). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. In L. H. Meyer, S. Davidson, H. Anderson, P. M. Johnston & M. Rees (Hrsg.), *Tertiary assessment and higher education student outcomes: Policy, practice and research* (S. 259–275): Victoria University; Ako Aotearoa: National Centre for Tertiary Teaching Excellence.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hawlotschek, P. & Henschel, S. (2025). Promoting teacher knowledge and self-efficacy for data-based decision-making: Dimensional structure, sensitivity to professional development, and relationships with teachers' intended data use. *Zeitschrift für Erziehungswissenschaft*, 28(6), 1577–1603. <https://doi.org/10.1007/s11618-025-01321-3>
- Hawlotschek, P., Henschel, S., Richter, D. & Stanat, P. (2024). The relationship between teachers' and principals' use of results from nationwide achievement tests: The mediating role of teacher attitudes and data use culture. *Studies in Educational Evaluation*, 80, Artikel 101317. <https://doi.org/10.1016/j.stueduc.2023.101317>
- Hawlotschek, P., Henschel, S., Schnitzler, C. & Stanat, P. (2025). Datengestützte Unterrichtsentwicklung mit Vergleichsarbeiten. Akzeptanz und wahrgenommene Nützlichkeit von VERA aus Sicht von Lehrkräften im IQB-Bildungstrend. In A. Albers & N. Jude (Hrsg.), *Blickpunkt Bildungsmonitoring – Bilanzen und Perspektiven* (S. 111–128). Beltz Juventa.
- Heckel, H. & Avenarius, H. (2000). *Schulrechtskunde. Ein Handbuch für Praxis, Rechtsprechung und Wissenschaft* (7. Aufl.). Luchterhand.
- Heikkinen, S., Saqr, M., Malmberg, J. & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions – a systematic literature review. *Education and Information Technologies*, 28(3), 3059–3088. <https://doi.org/10.1007/s10639-022-11281-4>
- Hejtmanek, R. A., Hahn, S., Schulte, K. & Klein, E. D. (2024). *Data Richness in Schulen unterstützen. Eine Handreichung für Akteur*innen aus dem Unterstützungssystem*. Technische Universität Dortmund, Institut für Bildungsmonitoring und Qualitätsentwicklung. https://www.bosch-stiftung.de/sites/default/files/publications/pdf/2025-01/TUDortmund_DataRichnessinSchulen_2024_barrierearm_CC-BY-NC.pdf
- Helmke, A., Helmke, T., Lenske, G., Pham, G., Praetorius, A.-K., Schrader, F.-W. & Ade-Thurow, M. (2011). *EMU – Unterrichtsdiagnostik. Version 3.2*. Universität Koblenz-Landau; Kultusministerkonferenz (KMK). https://lehrerfortbildung-bw.de/s_sueb/allgschulen/veranstalt/tt-4b-2011/f9/prae/Broschuere_203_13022011.pdf



- Helmke, A. & Hosenfeld, I. (2005). Standardbezogene Unterrichtsevaluation. In G. Brägger, B. Bucher, N. Landwehr & W. Böttcher (Hrsg.), *Schlüsselfragen zur externen Schulevaluation* (S. 127–151). hep Verlag.
- Hemmeter, M. L., Fox, L., Jack, S. & Broyles, L. (2007). A program-wide model of positive behavior support in early childhood settings. *Journal of Early Intervention*, 29(4), 337–355. <https://doi.org/10.1177/105381510702900405>
- Hendriks, M. A., Doolaard, S. & Bosker, R. J. (2001). School self-evaluation in the Netherlands: Development of the ZEBO-instrumentation. *Prospects*, 31(4), 503–518. <https://doi.org/10.1007/BF03220036>
- Henschel, S. & Stanat, P. (2019). Bildungsstandards als Element der Qualitätssicherung und -entwicklung im deutschen Schulsystem. In E. Kiel, B. Herzig, U. Maier & U. Sandfuchs (Hrsg.), *Handbuch Unterrichten an allgemeinbildenden Schulen* (S. 374–384). Klinkhardt.
- Herman, R., Christianson, K., Nguyen, P., Bush-Mecenas, S., Moini, J. S., DiNicola, S. E., Wrabel, S. L., Daramola, Eupha Jeanne, Holmes, Pierrece, Wolters, N., Gamse, B. C., Foley, N. F., Abbady, R., McCoy, L. L. & Zelazny, S. (2024). *Evaluation of the networks for school improvement initiative – Networks and Intermediaries. Interim Report*. RAND. https://www.rand.org/pubs/research_reports/RRA242-1.html
- Hernandez, D. J. (2012). *PreK-3rd: Next Steps for State Longitudinal Data Systems* (PreK-3rd Policy to Action Brief Nr. 8). <https://files.eric.ed.gov/fulltext/ED542862.pdf>
- Hertweck, F. (2024). *Datengestützt zum Erfolg. Internationale Beispiele für eine evidenzbasierte Transformation des Bildungssystems*. Vodafone Stiftung Deutschland. <https://www.vodafone-stiftung.de/bildungsdaten-studie-2024/>
- Hertweck, F., Isphording, I. E., Matthewes, S. H., Schneider, K. & Spieß, C. K. (2023). *Für einen besseren Datenzugang im Bereich Bildung. Stellungnahme der Arbeitsgruppe Bildungsdaten*. Verein für Socialpolitik. https://www.socialpolitik.de/sites/default/files/2023-11/5_Stellungnahme%20der%20UAG_Bildungsdaten.pdf
- Hilbrich, C. & Walter, K.-H. (2003). *Schulprogramme in Brandenburg – Selbsteinschätzungen und Fremdbeurteilungen. Textanalyse ausgewählter Schulprogramme und Schulleiterbefragungen an 100 Schulen* (Schulforschung in Brandenburg Nr. 4). Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. https://mbjs.brandenburg.de/sixcms/media.php/140/schulforschung_heft4_schulforschung_bbg.pdf
- Hill, N. E. & Tyson, D. F. (2009). Parental involvement in middle school: A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology*, 45(3), 740–763. <https://doi.org/10.1037/a0015362>
- Hillmayr, D., Reinhold, F., Zierwald, L. & Reiss, K. (2017). *Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe. Einsatzmöglichkeiten, Umsetzung und Wirksamkeit*. Waxmann.
- Hillmayr, D., Täschner, J., Brockmann, L. & Holzberger, D. (2021). *Elternbeteiligung im schulischen Kontext. Potenzial zur Förderung des schulischen Erfolgs von Schülerinnen und Schülern*. Waxmann. <https://doi.org/10.31244/9783830993667>
- Hofbauer, E. & Bernhard, R. (2023). Schulentwicklungsberatung und Schulentwicklungsberater*innen in kulturell diversen Settings in Österreich: Empirische Einsichten und warum wir mehr Unterrichtsentwicklung benötigen. *R&E-SOURCE*, 10(3), 136–158. <https://doi.org/10.53349/re-source.2023.i3.a1134>



- Hofer, S. I., Holzberger, D. & Reiss, K. (2020). Evaluating school inspection effectiveness: A systematic research synthesis on 30 years of international research. *Studies in Educational Evaluation*, 65, Artikel 100864. <https://doi.org/10.1016/j.stueduc.2020.100864>
- Hoffmann, C. & Pfister, A. (2019). Führen mit Zielen. In E. Lippmann, A. Pfister & U. Jörg (Hrsg.), *Handbuch angewandte Psychologie für Führungskräfte. Führungskompetenz und Führungswissen* (5., vollst. überarb. Aufl., S. 675–724). Springer.
- Hofman, R. H., Dijkstra, N. J. & Adriaan Hofman, W. H. (2009). School self-evaluation and student achievement. *School Effectiveness and School Improvement*, 20(1), 47–68. <https://doi.org/10.1080/09243450802664115>
- Holtappels, H. G. & Müller, S. (2004). Inhalte von Schulprogrammen – Ergebnisse einer Inhaltsanalyse Hamburger Schulprogrammtexte. In H. G. Holtappels (Hrsg.), *Schulprogramme – Instrumente der Schulentwicklung. Konzeptionen, Forschungsergebnisse, Praxisempfehlungen* (S. 79–102). Juventa.
- Hoogland, I., Schildkamp, K., van der Kleij, F., Heitink, M., Kippers, W., Veldkamp, B. & Dijkstra, A. M. (2016). Prerequisites for data-based decision making in the classroom: Research evidence and practical illustrations. *Teaching and Teacher Education*, 60, 377–386. <https://doi.org/10.1016/j.tate.2016.07.012>
- Horwood, M., Parker, P. D., Marsh, H. W., Guo, J. & Dicke, T. (2022). School autonomy policies lead to increases in principal autonomy and job satisfaction. *International Journal of Educational Research*, 115, Artikel 102048. <https://doi.org/10.1016/j.ijer.2022.102048>
- Hosenfeld, I., Zimmer-Müller, M. & Strasser, J. (2024). Datenbasierte Unterrichtsentwicklung mit VERA: (Wie) kann das funktionieren? In M. Schiefner-Rohs, S. Hofhues & A. Breiter (Hrsg.), *Datafizierung (in) der Bildung. Kritische Perspektiven auf digitale Vermessung in pädagogischen Kontexten* (S. 283–290). transcript Verlag. <https://doi.org/10.14361/9783839465820-017>
- Hough, H., Kalogrides, D. & Loeb, S. (2017). *Using surveys of students' social-emotional learning and school climate for accountability and continuous improvement*. PACE. Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/SEL-CC_report.pdf
- Huber, D. J. & Conway, J. M. (2015). The effect of school improvement planning on student achievement. *Planning and changing*, 46, 56–70.
- Hutchinson, C. & Young, M. (2011). Assessment for learning in the accountability era: Empirical evidence from Scotland. *Studies in Educational Evaluation*, 37(1), 62–70. <https://doi.org/10.1016/j.stueduc.2011.03.007>
- Ikemoto, G. S. & Marsh, J. A. (2007). Cutting through the “data-driven” mantra: Different conceptions of data-driven decision making. *Yearbook of the National Society for the Study of Education*, 106(1), 105–131. <https://doi.org/10.1111/j.1744-7984.2007.00099.x>
- Inspectorate of Education (2025). *Inspection Framework 2021 for the supervision of preschool education and primary education*. <https://english.onderwijsinspectie.nl/site/binaries/site-content/collections/documents/2025/09/04/inspection-framework-primary-education-2021/Inspection+framework+2025+for+primary+education.pdf>
- Institut für Bildungsmonitoring und Qualitätsentwicklung (2024a). *Zwischenbericht zur Schulqualität in Hamburg. 3. Zyklus Schulinspektion von Oktober 2019 bis Juli 2024*. https://ifbq.hamburg.de/wp-content/uploads/sites/803/2025/01/Bericht_Qualitaet-Hamburger-Schulen.pdf



- Institut für Bildungsmonitoring und Qualitätsentwicklung (2024b). *Handreichung SCHNABEL. Einführung in das individualdiagnostische Verfahren zur Erfassung von Rechtschreibkompetenzen*. https://ifbq.hamburg.de/wp-content/uploads/sites/803/2024/11/Handreichung-SCHNABEL-Klasse-1-10_Stand-April-2024.pdf
- Jackson, C. K. (2023). *When does school autonomy improve student outcomes?* (EdWorking Paper: 23-808). Annenberg Institute at Brown University. <https://doi.org/10.26300/cdj7-rg41>
- Jaekel, A.-K. (2021). *Die Nutzung von Schülerurteilen zur Erfassung von Unterrichtsqualität in Forschung und Schulpraxis* [Dissertation]. Eberhard Karls Universität Tübingen. <https://doi.org/10.15496/PUBLIKATION-61126>
- Jakobsen, M. L., Baekgaard, M., Moynihan, D. P. & van Loon, N. (2018). Making sense of performance regimes: Rebalancing external accountability and internal learning. *Perspectives on Public Management and Governance*, 1(2), 127–141. <https://doi.org/10.1093/ppmgov/gvx001>
- Jelley, F. & Sylva, K. (2018). *EasyPeasy. Evaluation in Newham: Findings from the Sutton Trust Parental Engagement Fund (PEF) Project*. <https://www.suttontrust.com/wp-content/uploads/2019/12/EasyPeasyNewham-FINAL.pdf>
- Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education*, 40(3), 237–269. <https://doi.org/10.1177/0042085905274540>
- Jeynes, W. H. (2014). A Meta-Analysis: The relationship between father involvement and student academic achievement. *Urban Education*, 50(4), 387–423. <https://doi.org/10.1177/0042085914525789>
- Jimerson, J. B., Garry, V., Poortman, C. L. & Schildkamp, K. (2021). Implementation of a collaborative data use model in a United States context. *Studies in Educational Evaluation*, 69, Artikel 100866. <https://doi.org/10.1016/j.stueduc.2020.100866>
- John, O. P., Naumann, L. P. & Soto, Christopher, J. (2010). Paradigm shift to the integrative Big Five Trait Taxonomy. History, measurement, and conceptual issues. In O. P. John, R. W. Robins & L. A. Pervin (Hrsg.), *Handbook of personality. Theory and research* (3rd ed., S. 114–158). Guilford Press.
- Johnson, D. A., Johnson, C. M. & Dave, P. (2024). Performance feedback in organizations: Understanding the functions, forms, and important features. In D. A. Johnson & C. M. Johnson (Hrsg.), *Handbook of Organizational Performance, Volume I. Foundations and Advances* (S. 146–171). Routledge. <https://doi.org/10.4324/9781032708270-7>
- Jones, S. M., Brush, K. E., Ramirez, T., Xinyi Mao, Z., Marenus, M., Wettje, S., Finney, K., Raisch, N., Podoloff, N., Kahn, J., Barnes, S., Stickle, L., Brion-Meisels, G., McIntyre, J., Cuartas, J. & Bailey, R. (2021). *Navigating SEL from the inside out. Looking inside and across 33 leading SEL programs: A practical resource for schools and OST providers. Preschool & elementary focus*. Harvard Graduate School of Education; The Wallace Foundation. <https://doi.org/10.59656/YD-OS5671.001>
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130–144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Jost, J. (2025). Datengestützte Unterrichtsentwicklung sollte nicht erst in der Schule beginnen! *Didaktik Deutsch*, 30(59), 19–24. <https://doi.org/10.21248/dideu.830>
- Jugend- und Familienministerkonferenz & Kultusministerkonferenz (2022). *Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen (Beschluss der JMK vom 13./14.05.2004 und Beschluss der KMK vom 03./04.06.2004 i. d. F. vom 06.05.2021 (JFMK) und 24.03.2022 (KMK)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_06_03-Fruehe-Bildung-Kindertageseinrichtungen.pdf



- Jugend- und Familienministerkonferenz & Kultusministerkonferenz (2023). *Ergebnisniederschrift der gemeinsamen Sitzung von Jugend- und Familienministerkonferenz (JFMK) und Kultusministerkonferenz (KMK) am 13. Oktober 2023 in Berlin*. https://jfmk.de/wp-content/uploads/2025/07/Ergebnisniederschrift-JFMK_KMK-Sitzung-am-13.10.2023_barr.pdf
- Junger, R. & Liebers, K. (2024). Digitale formative Diagnoseverfahren in der Grundschule – Begründungen und Überblick über vorliegende Entwicklungen am Beispiel Deutsch. *Zeitschrift für Grundschulforschung*, 17(1), 3–19. <https://doi.org/10.1007/s42278-024-00190-9>
- Jungjohann, J. & Gebhardt, M. (2022). Bezugsnormorientierung im Unterricht. In M. Gebhardt, D. Scheer & M. Schurig (Hrsg.), *Handbuch der sonderpädagogischen Diagnostik. Grundlagen und Konzepte der Statusdiagnostik, Prozessdiagnostik und Förderplanung* (S. 25–32). Universität Regensburg.
- Kaliisa, R., Misiejuk, K., López-Pernas, S., Khalil, M. & Saqr, M. (2024). Have learning analytics dashboards lived up to the hype? A systematic review of impact on students' achievement, motivation, participation and attitude. In The Association for Computing Machinery, *Proceedings of the 14th Learning Analytics and Knowledge Conference*, Kyoto, Japan.
- Karaman, P. (2021). The effect of formative assessment practices on student learning: A meta-analysis study. *International Journal of Assessment Tools in Education*, 8(4), 801–817. <https://doi.org/10.21449/ijate.870300>
- Karst, K., Bonefeld, M., Dotzel, S., Fehringer, B. C. O. F. & Steinwascher, M. (2022). Data-based differentiated instruction: The impact of standardized assessment and aligned teaching material on students' reading comprehension. *Learning and Instruction*, 79, Artikel 101597. <https://doi.org/10.1016/j.learninstruc.2022.101597>
- Karst, K., Yendell, O., Marx, A., Lettau, W.-D. & Hawlitschek, P. (2024). Die Etablierung von Evidenzteams in SchuMaS – Eine Strategie zur systematischen Nutzung von Daten für die Schul- und Unterrichtsentwicklung. In K. Maaz & A. Marx (Hrsg.), *SchuMaS – Schule macht stark. Sozialraumorientierte Schul- und Unterrichtsentwicklung an Schulen in schwierigen Lagen. Aufbau und erste Arbeitsergebnisse des Forschungsverbunds* (S. 225–243). Waxmann.
- Kautz, T., Heckman, J., Diris, R., Weel, B. ter & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success*. NBER Working Paper 20749. <https://doi.org/10.3386/w20749>
- Keitel, J. (2024). *Schulleitungen im institutionellen Zielvereinbarungsprozess. Handlungsleitende Orientierungen von Schulleiter*innen zwischen alten und neuen Steuerungsvorstellungen*. Springer VS. <https://doi.org/10.1007/978-3-658-47047-0>
- Kelchtermans, G. (2006). Teacher collaboration and collegiality as workplace conditions. A review. *Zeitschrift für Pädagogik*, 52(2), 220–237. <https://doi.org/10.25656/01:4454>
- Kellermann, C., Nachbauer, M., Gärtner, H. & Thiel, F. (2022). Effekte eines Unterrichtsfeedbacks von Schulleitungen auf die selbst eingeschätzte Entwicklung der Unterrichtskompetenz von Lehrkräften – Ergebnisse einer Interventionsstudie unter Berücksichtigung der Zielorientierungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 25(5), 1267–1291. <https://doi.org/10.1007/s11618-022-01087-y>
- Kellermann, C., Nachbauer, M., Gärtner, H. & Thiel, F. (2023). The effect of school leaders' feedback on teaching quality perceived by students – results of an intervention study. *School Effectiveness and School Improvement*, 34(1), 24–42. <https://doi.org/10.1080/09243453.2022.2075018>



- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980. <https://doi.org/10.3102/0034654315626800>
- Keune, M. S. (2014). *Schulinspektion unter besonderer Berücksichtigung externer Zielvereinbarungen. Eine explorative ‚mixed methods‘-Studie am Beispiel der hessischen Schulinspektion* [Dissertation]. Westfälische Wilhelms-Universität Münster. <https://d-nb.info/1137470267/34>
- Keuning, T., van Geel, M. & Visscher, A. J. (2017). Why a data-based decision-making intervention works in some schools and not in others. *Learning Disabilities Research & Practice*, 32(1), 32–45. <https://doi.org/10.1111/ldrp.12124>
- Keuning, T., van Geel, M., Visscher, A. J. & Fox, J.-P. (2019). Assessing and validating effects of a data-based decision-making intervention on student growth for mathematics and spelling. *Journal of Educational Measurement*, 56(4), 757–792. <https://doi.org/10.1111/jedm.12236>
- Killus, D. & Paseka, A. (2021). Kooperation zwischen Eltern und Schule: eine Orientierung im Themenfeld. *Die Deutsche Schule*, 113(3), 253–266. <https://doi.org/10.31244/ddS.2021.03.02>
- Killus, D. & Paseka, A. (2022). Erziehungs- und Bildungspartnerschaft mit Eltern. *Pädagogik*(8), 38–42. <https://doi.org/10.3262/paed2208038>
- Kingston, N. & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37. <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- Klein, E. D. & Bremm, N. (Hrsg.). (2020). *Unterstützung – Kooperation – Kontrolle. Zum Verhältnis von Schulaufsicht und Schulleitung in der Schulentwicklung*. Springer VS. <https://doi.org/10.1007/978-3-658-28177-9>
- Kleine, J. (2012). *Rezeption und Nutzung von Ergebnissen aus Schulinspektionen. Umgang von Schulen und Schulaufsicht mit Ergebnissen der Qualitätsanalyse NRW* [Dissertation]. Westfälische Wilhelms-Universität Münster. <https://www.uni-muenster.de/Ebooks/index.php/series/catalog/book/205>
- Klieme, E. (2018). Unterrichtsqualität. In M. Harring, C. Rohlf's & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 393–408). Waxmann.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Deutsches Institut für Internationale Pädagogische Forschung (DIPF). http://www.ak-mathe.muc.kobis.de/Download/Arbeitskreis/expertise_bildungsstandards.pdf
- Klieme, E. & Baumert, J. *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*. Bundesministerium für Bildung und Forschung (BMBF). http://digital.bib-bvb.de/view/bvb_single/single.jsp?dvs=1772795116966-353&locale=de&VIE-WER_URL=/view/bvb_single/single.jsp?&DELIVERY_RULE_ID=39&bfe=view/action/singleViewer.do?dvs=&frameld=1&usePid1=true&usePid2=true
- Klopsch, B. & Sliwka, A. (2020). Schulqualität als Resultat einer Verschränkung von Systemebenen. Daten-gestützte Schulentwicklung in der Provinz Alberta, Kanada. In D. Fickermann, V. Manitiuis & M. Karcher (Hrsg.), „*Neue Steuerung*“ – *Renaissance der Kybernetik? Die Deutsche Schule*, 15. Beiheft (S. 58–73). Waxmann. <https://doi.org/10.25656/01:24143>
- Kluczniok, K., Faas, S., von Hülsen, K., Schneider, M., Fitzner, J., Koch, C. & Aden, H. (2024). *Machbarkeitsstudie zur Ansprache und Umsetzung eines bundesweiten Monitorings zur Prozessqualität – Abschlussbericht*. pädquis. <https://kitaqualitaetsmonitor.de>



- Kluczniok, K., Grad, T., Schneider, M. & Faas, S. (2024). *Expertise „Auswirkungen von Kindertagesbetreuung auf die kindliche Entwicklung“*. pädquis. https://www.fruehe-chancen.de/fileadmin/user_upload/PDF-Dateien/Expertisen_und_Studien/Expertise_Auswirkungen_von_Kindertagesbetreuung_auf_die_kindliche_Entwicklung_p%C3%A4dquis.pdf
- Klusmann, U., Aldrup, K., Roloff, J., Lüdtke, O. & Hamre, B. K. (2022). Does instructional quality mediate the link between teachers' emotional exhaustion and student outcomes? A large-scale study using teacher and student reports. *Journal of Educational Psychology*, 114(6), 1442–1460. <https://doi.org/10.1037/edu0000703>
- Knoop-van Campen, C. A. N., Wise, A. & Molenaar, I. (2023). The equalizing effect of teacher dashboards on feedback in K-12 classrooms. *Interactive Learning Environments*, 31(6), 3447–3463. <https://doi.org/10.1080/10494820.2021.1931346>
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A. & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66(2), 165–191. <https://doi.org/10.1007/s11577-014-0253-x>
- Koch, U. (2011). *Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten? Datenkompetenz von Lehrkräften und die Nutzung von Ergebnissrückmeldungen aus Vergleichsarbeiten*. Waxmann.
- Kochmanski, N. & Cobb, P. (2023). Identifying and negotiating productive instructional improvement goals in one-on-one mathematics coaching. *Journal of Teacher Education*, 74(5), 437–450. <https://doi.org/10.1177/00224871221143124>
- Koedel, C., Mihaly, K. & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47, 180–195. <https://doi.org/10.1016/j.econedurev.2015.01.006>
- Koh, E. & Tan, J. P.-L. (2017). *Teacher-actionable insights in student engagement: A learning analytics taxonomy*. International Conference on Computers in Education. <https://library.apsce.net/index.php/ICCE/article/view/2300>
- Kokoç, M., Bütüner, S. Ö. & Güler, M. (2025). A meta-analysis on the effect of learning analytics interventions on students' academic performance. *Journal of Research on Technology in Education*, 1–24. <https://doi.org/10.1080/15391523.2025.2536571>
- Köller, O. (2025). Text-Bewertung mit Large Language Models. Chancen und Herausforderungen beim Einsatz von generativer KI zur Bewertung schriftlicher Leistungen. *Diagnostica*, 71(3-4), 105–108. <https://doi.org/10.1026/0012-1924/a000356>
- Kraft, M. A., Blazar, D. & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Kraft, M. A. & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review*, 47, 49–63. <https://doi.org/10.1016/j.econedurev.2015.04.001>
- Kreis, A. & Staub, F. C. (2008). Praxislehrpersonen als Unterrichtscoachs und als Mediatoren in der Rekontextualisierung unterrichtsbezogenen Wissens. *Beiträge zur Lehrerbildung*, 26(2), 198–210. <https://doi.org/10.25656/01:13675>
- Kreitz-Sandberg, S., Backlund, Å., Fredriksson, U., Isaksson, J., Rasmusson, M. & Gren Landell, M. (2022). Recording and reporting school attendance and absence: International comparative views on attendance statistics in Sweden, Germany, England, and Japan. *Orbis Scholae*, 16(2-3), 187–212. <https://doi.org/10.14712/23363177.2023.9>



- Kroupa, A., Tarkian, J., Füssel, H.-P., Schewe, C. & Thiel, F. (2019). Bildungsrechtliche Grundlagen datenbasierter Qualitätssicherung und -entwicklung. In F. Thiel, J. Tarkian, E.-M. Lankes, N. Maritzen, T. Riecke-Baulecke & A. Kroupa (Hrsg.), *Datenbasierte Qualitätssicherung und -entwicklung in Schulen. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland* (231-312). Springer VS. https://doi.org/10.1007/978-3-658-23240-5_7
- Kruse, S. D. (2001). Creating communities of reform: continuous improvement planning teams. *Journal of Educational Administration*, 39(4), 359–383. <https://doi.org/10.1108/EUM000000005496>
- Kuhn, A., Schleper, H., Pfafferott, M. & Zorn, D. (2025). *Bildungsverlaufsregister und Schüler-ID in Deutschland. Status quo und Gelingensbedingungen*. Bertelsmann Stiftung. <https://doi.org/10.11586/2025054>
- Kuhn, H.-J. (2019). *Angst vor dem Urteil? Eine Bestandsaufnahme der externen Schulevaluation in Deutschland* (böll.brief Teilhabegesellschaft Nr. 10). https://www.boell.de/sites/default/files/boell.brief_tg10_angst_vor_dem_urteil.pdf
- Kultusministerkonferenz (1997). *Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland. Konstanzer Beschluss (Beschluss der Kultusministerkonferenz vom 24.10.1997)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf
- Kultusministerkonferenz (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. LinkLuchterhand. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/Beschluesse_Veroeffentlichungen/Bildungsmonitoring_Broschuere_Endf.pdf
- Kultusministerkonferenz (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Carl Link. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf
- Kultusministerkonferenz (2018). *Bildung und Erziehung als gemeinsame Aufgabe von Eltern und Schule (Beschluss der Kultusministerkonferenz vom 11.10.2018)*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2018/2018_10_11-Empfehlung-Bildung-und-Erziehung.pdf
- Kultusministerkonferenz (2020a). *Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen (Beschluss der Kultusministerkonferenz vom 15.10.2020)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2020/2020_10_15-Laendervereinbarung.pdf
- Kultusministerkonferenz (2020b). *Politische Vorhaben zur „Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen“ vom 15.10.2020 (Beschluss der KMK vom 15.10.2020)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2020/2020_10_15-Politische-Vorhaben-LV.pdf
- Kultusministerkonferenz (2022). *Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004 i. d. F. vom 07.10.2022)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- Kultusministerkonferenz (2024a). *Empfehlung der Kultusministerkonferenz zur Stärkung der mathematisch-informatisch-naturwissenschaftlich-technischen Bildung (Beschluss der KMK vom 07.05.2009 i. d. F. vom 13.06.2024)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2009/2009_05_07-Empf-MINT.pdf



- Kultusministerkonferenz (2024b). *Eckpunktepapier zur Verankerung von Medienbildung und informatischer Bildung in der Grundschule (Beschluss der Kultusministerkonferenz vom 28.11.2024)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_11_28-Medien-inform-Bildung-Grundschule.pdf
- Kultusministerkonferenz (2024c). *Bericht zum Stand der Umsetzung der Förderstrategie für leistungsschwächere Schülerinnen und Schüler (Beschluss der Kultusministerkonferenz vom 25.04.2024)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_04_25-Umsetzung-Foerderstrategie.pdf
- Kultusministerkonferenz (2024d). *Politische Vorhaben zur „Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen“ vom 15.10.2020 (Beschluss der KMK vom 15.10.2020). Bericht über die Umsetzung der Vorhaben*.
- Kultusministerkonferenz (2025). *Zielbild zur Rolle und Arbeit der Schulaufsicht (Beschluss der Kultusministerkonferenz vom 16.10.2025 für die Kultusministerkonferenz)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2025/2025_10_16-Schulaufsicht.pdf
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Ferdinand Schöningh. <https://doi.org/10.36198/9783838538952>
- Kunter, M. & Voss, T. (2013). The model of instructional quality in COACTIV: A multicriteria analysis. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project* (S. 97–124). Springer. https://doi.org/10.1007/978-1-4614-5149-5_6
- Kunze, E., Wietstock, F. & Krelle, M. (i. E.). Digitale Lautleseformate mit LeOn (Leseraum Online) im Grundschulunterricht. Ein didaktischer Ansatz zur Verbindung von Mündlichkeit und Digitalisierung im Lesenlernen. In U. Behrens, M.-O. Carl, J. Kreuz & M. Magirus (Hrsg.), *Poteniale mündlicher Kommunikation im Deutschunterricht – quer durch alle Kompetenzbereiche*.
- Kyriakides, L. & Campbell, R. J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation*, 30(1), 23–36. [https://doi.org/10.1016/S0191-491X\(04\)90002-8](https://doi.org/10.1016/S0191-491X(04)90002-8)
- Lai, M. K. & McNaughton, S. (2013). Analysis and discussion of classroom and achievement data to raise student achievement. In K. Schildkamp, M. K. Lai & L. M. Earl (Hrsg.), *Data-based decision making in education. Challenges and opportunities* (S. 23–47). Springer. https://doi.org/10.1007/978-94-007-4816-3_3
- Lai, M. K., McNaughton, S., Amituanai-Tolosa, M., Turner, R. & Hsiao, S. (2009). Sustained acceleration of achievement in reading comprehension: The New Zealand experience. *Reading Research Quarterly*, 44(1), 30–56. <https://doi.org/10.1598/RRQ.44.1.2>
- Landeshauptstadt München (2022). *Münchner Bildungsbericht 2022*. <https://www.pi-muenchen.de/wp-content/uploads/2023/03/Muenchner-Bildungsbericht-2022-3.pdf>
- Landeshauptstadt München (2024). *Münchner Bildungsbericht Berufliche Bildung 2024*. https://stadt.muenchen.de/dam/jcr:f912fb1a-1e17-4516-a4e0-fo099b54affc/muenchner_bildungsbericht_berufliche_bildung_2024.pdf
- Landesinstitut für Schule und Medien Berlin-Brandenburg (2021). *ILeA plus. Handbuch für Lehrerinnen und Lehrer*. https://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/unterricht/lernbegleitende_Diagnostik/ilea_plus/211102-ILeAplus.pdf



- Lange, H. (2003a). Schulaufsicht zwischen normativen Anforderungen und faktischen Wirkungsmöglichkeiten. In H.-P. Füssel & P. M. Roeder (Hrsg.), *Recht – Erziehung – Staat. Zur Genese einer Problemkonstellation und zur Programmatik ihrer zukünftigen Entwicklung*. Zeitschrift für Pädagogik, 47. Beiheft (S. 137–155). Beltz. <https://doi.org/10.25656/01:3973>
- Lange, H. (2003b). Schulaufsicht nach PISA. *Recht der Jugend und des Bildungswesens*, 51(4), 485–491.
- Lange, H. (2008). Vom Messen zum Handeln: „empirische Wende“ der Bildungspolitik? *Recht der Jugend und des Bildungswesens*, 56(1), 7–15.
- Lankes, E.-M. & Huber, F. (2014). *Leitfaden für die Erstellung von Zielvereinbarungen*. Staatsinstitut für Schulqualität und Bildungsforschung München. https://www.schulentwicklung.isb.bayern.de/fileadmin/user_upload/Schulentwicklung/QmbS/5_iE_Dateien/leitfaden_zv_2014.pdf
- Larsen, J. N., Maxwell, K. M. D. & Khalil, M. (2025). Exploring the link between motivational regulation and learning design with learning analytics: A systematic literature review. *Journal of Learning Analytics*, 12(1), 215–237. <https://doi.org/10.18608/jla.2025.8469>
- Lawal, A. (2021). How teachers can communicate effectively with parents who speak a different language. *International Journal of Linguistics, Literature and Translation*, 4(9), 166–178. <https://doi.org/10.32996/ijlt.2021.4.9.17>
- Lee, H., Chung, H. Q., Zhang, Y., Abedi, J. & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education*, 33(2), 124–140. <https://doi.org/10.1080/08957347.2020.1732383>
- Lee, J [Jaewook], Smith, D., Woodhead, S. & Lan, A. (2024). *Math multiple choice question generation via Human-Large Language Model collaboration*. Proceedings of the 17th International Conference on Educational Data Mining. <https://dblp.org/rec/conf/edm/0006Sol24>
- Lee, J [Jihyun], Alonzo, D., Beswick, K., Abril, J. M. V., Chew, A. W. & Oo, C. Z. (2024). Dimensions of teachers' data literacy: A systematic review of literature from 1990 to 2021. *Educational Assessment, Evaluation and Accountability*, 36(2), 145–200. <https://doi.org/10.1007/s11092-024-09435-8>
- Lee, J [Jiwon] & Lee, J [Jeongmin] (2025). Generative AI chatbot for teachers' data-informed decision-making: Effects and insights. *Educational Technology & Society*, 28(3), 298–317. [https://doi.org/10.30191/ETS.202507_28\(3\).TP06](https://doi.org/10.30191/ETS.202507_28(3).TP06)
- Lehrl, S. (2018). *Qualität häuslicher Lernumwelten im Vorschulalter. Eine empirische Analyse zu Konzept, Bedingungen und Bedeutung*. Springer VS. <https://doi.org/10.1007/978-3-658-20184-5>
- Lehrl, S., Flöter, M., Wieduwilt, N. & Anders, Y. (2020). Direkte und indirekte Bedeutsamkeit der Zusammenarbeit mit Familien für die kindliche Sprachentwicklung. In K. Blatter, K. Groth & M. Hasselhorn (Hrsg.), *Evidenzbasierte Überprüfung von Sprachförderkonzepten im Elementarbereich* (S. 129–152). Springer VS. https://doi.org/10.1007/978-3-658-26438-3_6
- Levatino, A., Verger, A., Camphuijsen, M., Termes, A. & Parcerisa, L. (2024). School governance through performance-based accountability: A comparative analysis of its side effects across different regulatory regimes. *Review of Research in Education*, 48(1), 248–286. <https://doi.org/10.3102/0091732X241270672>
- Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O. & Reiss, K. (Hrsg.). (2023). *PISA 2022. Analyse der Bildungsergebnisse in Deutschland*. Waxmann. <https://doi.org/10.25656/01:28666>



- Liebers, K., Junger, R., Kanold, E., Latzko, B., Gottlebe, K. & Dietrich, S. (2020). *Bericht zum Teilprojekt 2 „Erprobung der Aufgabenpakete ILeA plus Deutsch“ (überarbeitete und redaktionell ergänzte Version vom 27.03.2020)*. Universität Leipzig. <https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-706980>
- Light, D., Honey, M., Heinze, J., Brunner, C., Wexler, D., Mandinach, E. B. & Fasca, C. (2005). *Linking data and learning – The Grow Network Study: Summary report*. Education Development Center. <https://eric.ed.gov/?id=ED485607>
- Linberg, T. & Taut, S. (2022, 9. März). *Externe Schulevaluation: Liefern wir valide Daten zur Unterrichtsqualität?* [Symposiumsbeitrag]. Otto-Friedrich-Universität Bamberg & Leibniz-Institut für Bildungsverläufe. 9. GEBF-Tagung, Bamberg. <https://www.gebf-ev.de/men%C3%BC/gebf-kongresse/2022-9-gebf-tagung-bamberg/>
- Lipowsky, F. & Rzejak, D. (2012). Lehrerinnen und Lehrer als Lerner – Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen. In L. Criblez, D. Bosse & T. Hascher (Hrsg.), *Reform der Lehrerbildung in Deutschland, Österreich und der Schweiz. Teil 1: Analysen, Perspektiven und Forschung* (S. 235–254). Prolog.
- Lippert, K., Jähner, A. & Kuger, S. (2022). *Einschätzung zu Fachkräften und Angebote für Familien in der Kindertagesbetreuung: Die Perspektive der Eltern. DJI-Kinderbetreuungsreport 2021*. Deutsches Jugendinstitut. https://www.dji.de/fileadmin/user_upload/KiBS/Kinderbetreuungsreport_2021_Studie6_Angebote_und_Fachkraefte.pdf
- Liu, Y [Yan], Wang, W. & Xu, E. (2025). The effectiveness of learning analytics-based interventions in enhancing students' learning effect: A meta-analysis of empirical studies. *SAGE Open*, 15(2), Artikel 21582440251336707. <https://doi.org/10.1177/21582440251336707>
- Liu, Y [Yunting], Bhandari, S. & Pardos, Z. A. (2025). Leveraging LLM respondents for item evaluation: A psychometric analysis. *British Journal of Educational Technology*, 56(3), 1028–1052. <https://doi.org/10.1111/bjet.13570>
- Locke, E. A. & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *American Psychologist*, 57(9), 705–717. <https://doi.org/10.1037/0003-066x.57.9.705>
- Loeb, S., Christian, M. S., Hough, H., Meyer, R. H., Rice, A. B. & West, M. R. (2019). School differences in social-emotional learning gains: Findings from the first large-scale panel survey of students. *Journal of Educational and Behavioral Statistics*, 44(5), 507–542. <https://doi.org/10.3102/1076998619845162>
- Loeb, S. & Strunk, K. (2007). Accountability and local control: Response to incentives with and without authority over resource generation and allocation. *Education Finance and Policy*, 2(1), 10–39. <https://doi.org/10.1162/edfp.2007.2.1.10>
- Lokhande, M., Hoeft, M. & Wendt, H. (2014). *Eltern als Bildungspartner: Wie Beteiligung an Grundschulen gelingen kann*. Forschungsbereich beim Sachverständigenrat deutscher Stiftungen für Integration und Migration (SVR). https://www.svr-migration.de/wp-content/uploads/2014/12/SVR-FB_Elternbeteiligung.pdf
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, 34(2), 120–131. <https://doi.org/10.1016/j.cedpsych.2008.12.001>



- Luginbuhl, R., Webbink, D. & de Wolf, I. F. (2009). Do inspections improve primary school performance? *Educational Evaluation and Policy Analysis*, 31(3), 221–237. <https://doi.org/10.3102/0162373709338315>
- Lynch, K., Hill, H. C., Gonzalez, K. E. & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis*, 41(3), 260–293. <https://doi.org/10.3102/0162373719849044>
- Ma, W., Adesope, O. O., Nesbit, J. C. & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918. <https://doi.org/10.1037/a0037123>
- Ma, X., Shen, J., Krenn, H. Y., Hu, S. & Yuan, J. (2016). A meta-analysis of the relationship between learning outcomes and parental involvement during early childhood education and early elementary education. *Educational Psychology Review*, 28(4), 771–801. <https://doi.org/10.1007/s10648-015-9351-1>
- Mack, W. (2014). Bildung in Schule und Jugendhilfe: Soziale Bedingungen von Bildung als Herausforderung und Chance für die Kooperation von Schule und Jugendhilfe. *Die Deutsche Schule*, 106(1), 62–71. <https://doi.org/10.25656/01:25845>
- Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K. & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist*, 76(7), 1128–1142. <https://doi.org/10.1037/amp0000701>
- Maier, J. L., Panetta, T., Philipp, J., Ruth-Herbein, E. & Fauth, B. (2021). *Unterrichtsfeedbackbogen Tiefenstrukturen. Präpilottierung und Validierung: Erste Ergebnisse der wissenschaftlichen Begleitung*. https://ibbw-bw.de/site/pbs-bw-rebrush2024/get/documents_E-1283939021/KULTUS.Dachmandant/KULTUS/Dienststellen/ibbw/Empirische%20Bildungsforschung/Programme-und-Projekte/Unterrichtsfeedbackbogen/UFB_Ergebnisbericht_Pr%C3%A4p%2BVal.pdf
- Maier, U. (2008). Rezeption und Nutzung von Vergleichsarbeiten aus der Perspektive von Lehrkräften. *Zeitschrift für Pädagogik*, 54(1), 95–117. <https://doi.org/10.25656/01:4338>
- Maier, U. (2010). Vergleichsarbeiten im Spannungsfeld zwischen formativer und summativer Leistungsmessung. *Die Deutsche Schule*, 102(1), 60–69. <https://doi.org/10.25656/01:25617>
- Maier, U., Bohl, T., Kleinknecht, M. & Metz, K. (2011). Einflüsse von Merkmalen des Testsystems und Schulkontextfaktoren auf die Akzeptanz und Rezeption von zentralen Testrückmeldungen durch Lehrkräfte. *Journal for Educational Research Online*, 3(2), 62–93. <https://doi.org/10.25656/01:5625>
- Mandinach, E. B. & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366–376. <https://doi.org/10.1016/j.tate.2016.07.011>
- Mandinach, E. B., Gummer, E. S. & Muller, R. D. (2011). *The complexities of integrating data-driven decision making into professional preparation in schools of education: It's harder than you think*.
- Mandinach, E. B., Honey, M., Light, D. & Brunner, C. (2008). A conceptual framework for data-driven decision-making. In E. B. Mandinach & M. Honey (Hrsg.), *Data-driven school improvement: Linking data and learning* (S. 13–31). Teachers College Press.
- Mandinach, E. B. & Schildkamp, K. (2021). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation*, 69, Artikel 100842. <https://doi.org/10.1016/j.stueduc.2020.100842>



- Mapp, K. L. & Kuttner, P. J. (2013). *Partners in education. A dual capacity-building framework for family-school partnerships*. SEDL. <https://www.ed.gov/sites/ed/files/documents/family-community/partners-education.pdf>
- Maritzen, N. (2014). Glanz und Elend der KMK-Strategie zum Bildungsmonitoring. Versuch einer Bilanz und eines Ausblicks. *Die Deutsche Schule*, 106(4), 398–413. <https://doi.org/10.25656/01:25903>
- Marrone, R., Lam, B., Barthakur, A., Brinkman, S., Dawson, S. & Gabriel, F. (2024). The relationship between wellbeing and academic achievement: A comprehensive cross-sectional analysis of system wide data from 2016-2019. *Journal of Learning Analytics*, 11(3), 123–141. <https://doi.org/10.18608/jla.2024.8357>
- Marsh, J. A., Sloan McCombs, J. & Martorell, F. (2010). How instructional coaches support data-driven decision making. *Educational Policy*, 24(6), 872–907. <https://doi.org/10.1177/0895904809341467>
- Marzano, R. J. & Waters, T. (2009). *District leadership that works: Striking the right balance*. Solution Tree Press.
- Matsumura, L. C., Correnti, R., Walsh, M., Bickel, D. D. & Zook-Howell, D. (2019). Online content-focused coaching to improve classroom discussion quality. *Technology, Pedagogy and Education*, 28(2), 191–215. <https://doi.org/10.1080/1475939X.2019.1577748>
- McKown, C. (2019). Challenges and opportunities in the applied assessment of student social and emotional learning. *Educational Psychologist*, 54(3), 205–221. <https://doi.org/10.1080/00461520.2019.1614446>
- McNaughton, S., Lai, M. K. & Hsiao, S. (2012). Testing the effectiveness of an intervention model based on data use: A replication series across clusters of schools. *School Effectiveness and School Improvement*, 23(2), 203–228. <https://doi.org/10.1080/09243453.2011.652126>
- Methe, S. & van der Heyden, A. M. (2011). Response to intervention for early mathematics. In V. Buysse & E. Peisner-Feinberg (Hrsg.), *Handbook of response to intervention in early childhood* (S. 169–184). Paul H. Brookes Publishing Co.
- Meyer, R. H., Pier, L., Mader, J., Christian, M., Rice, A., Loeb, S., Fricke, H. & Hough, H. (2019). *Can we measure classroom supports for social-emotional learning?* PACE. Policy Analysis for California Education. <https://files.eric.ed.gov/fulltext/ED600449.pdf>
- Meyers, C. V. & van Gronigen, B. A. (2019). A lack of authentic school improvement plan development: Evidence of principal satisficing behavior. *Journal of Educational Administration*, 57(3), 261–278. <https://doi.org/10.1108/JEA-09-2018-0154>
- Ministerium für Bildung, Wissenschaft und Kultur Schleswig-Holstein (2022). *Konzept zum Schulabsentismus*. https://www.schleswig-holstein.de/DE/landesregierung/ministerien-behoerden/III/Service/Broschueren/Bildung/Absentismus_Konzept.pdf?__blob=publicationFile&v=3
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2024). *Handreichung zu den Ziel- und Leistungsvereinbarungen an allgemein bildenden Schulen*.
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2021). *Fachberatung in der Schulaufsicht: Fachberaterinnen und Fachberater für die Fachkoordination Deutsch und Mathematik*. Runderlass 10-32 Nr. 51.2. <https://bass.schule.nrw/19546.htm>



- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2024). *Richtlinien. Bildungs- und Erziehungsgrundsätze für die allgemeinbildenden Schulen in Nordrhein-Westfalen*. https://lehrplannavigator.nrw.de/system/files/media/document/file/rili_allgb_schulen_2024_10_04_0.pdf
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2025a). *Schulkompass NRW 2030. Klare Ziele. Klarer Fokus*. https://www.schulministerium.nrw/system/files/media/document/file/schulkompass_nrw_2030_dqe_broschuere.pdf
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2025b). *Informationen zur Qualitätsanalyse*. https://www.schulministerium.nrw/system/files/media/document/file/informationen_qualitaetsanalyse_nrw_250801.pdf
- Mintrop, H. & MacLellan, A. M. (2002). School improvement plans in elementary and middle schools on probation. *The Elementary School Journal*, 102(4), 275–300. <https://doi.org/10.1086/499704>
- Mohr, I. (2006). *Analyse von Schulprogrammen. Eine Arbeit im Rahmen der Internationalen Grundschul-Lese-Untersuchung (IGLU)*. Waxmann.
- Mononen, R., Aunio, P., Koponen, T. & Aro, M. (2015). A review of early numeracy interventions for children at risk in mathematics. *International Journal of Early Childhood Special Education*, 6(1), 25–54. <https://doi.org/10.20489/intjecse.14355>
- Moroni, S., Dumont, H., Trautwein, U., Niggli, A. & Baeriswyl, F. (2015). The need to distinguish between quantity and quality in research on parental involvement: The example of parental help with homework. *The Journal of Educational Research*, 108(5), 417–431. <https://doi.org/10.1080/00220671.2014.901283>
- Moser Opitz, E. (2022). Diagnostisches und didaktisches Handeln verbinden: Entwicklung eines Prozessmodells auf der Grundlage von Erkenntnissen aus der pädagogischen Diagnostik und der Förderdiagnostik. *Journal für Mathematik-Didaktik*, 43, 205–230. <https://doi.org/10.1007/s13138-022-00201-1>
- MSB NRW (2024). *Gemeinsames Gutachten zum Wissenschaftlichen Prüfauftrag zur steigenden Anzahl der Schülerinnen und Schüler mit Bedarf an sonderpädagogischer Unterstützung*. https://www.schulministerium.nrw/system/files/media/document/file/langfassung_wissenschaftlicher_pruefauftrag_sonderpaedagogische_foerderung.pdf
- Müller, S. (2015). Zukunft der externen Evaluation – Trends und Perspektiven in Deutschland. In M. Pietsch, B. Scholand & K. Schulte (Hrsg.), *Schulinspektion in Hamburg. Der erste Zyklus 2007–2013: Grundlagen, Befunde, Perspektiven* (S. 369–384). Waxmann.
- National Policy Board for Educational Administration (2015). *Professional standards for educational leaders*. https://www.npbea.org/wp-content/uploads/2017/06/Professional-Standards-for-Educational-Leaders_2015.pdf
- Nationale Akademie der Wissenschaften Leopoldina (2024). *Förderung der Selbstregulationskompetenzen von Kindern und Jugendlichen in Kindertageseinrichtungen und Schulen*. https://doi.org/10.26164/leopoldina_03_01157
- Nelson, R., Ehren, M. & Godfrey, D. (2015). *Literature review on internal evaluation*. Institute of Education, University College London. <https://discovery.ucl.ac.uk/id/eprint/1475833/1/Literature-review-internal-evaluation.pdf>
- Neumann, P. & Lütje-Klose, B. (2020). Diagnostik in inklusiven Schulen – zwischen Stigmatisierung, Etikettierungs-Ressourcen-Dilemma und förderorientierter Handlungsplanung. In C. Gresch, P. Kuhl, M. Grosche, C. Sälzer & P. Stanat (Hrsg.), *Schüler*innen mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Einblicke und Entwicklungen* (S. 3–28). Springer VS



- O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72(3), 293–329. <https://doi.org/10.17763/haer.72.3.021q742t8182h238>
- OECD (2001). *Knowledge and skills for life. First Results from PISA 2000*. <https://doi.org/10.1787/9789264195905-en>
- OECD (2012). *Let's read them a story! The parent factor in education*. <https://doi.org/10.1787/9789264176232-en>
- OECD (2013). *Synergies for better learning. An international perspective on evaluation and assessment*. <https://doi.org/10.1787/9789264190658-en>
- OECD (2015). *Education at a glance 2015. OECD Indicators*. <https://doi.org/10.1787/eag-2015-en>
- OECD (2019a). *PISA 2018 Results (Volume III). What School Life Means for Students' Lives*. <https://doi.org/10.1787/acd78851-en>
- OECD (2019b). *Providing quality early childhood education and care. Results from the Starting Strong Survey 2018*. <https://doi.org/10.1787/301005d1-en>
- OECD (2020). *PISA 2018 results (Volume V). Effective policies, successful schools*. <https://doi.org/10.1787/ca768d40-en>
- OECD (2021). *Beyond academic learning. First results from the survey of social and emotional skills*. <https://doi.org/10.1787/92a11084-en>
- OECD (2023). *Schools as hubs for social and emotional learning – Are schools and teachers ready?* (OECD Education Spotlights Nr. 4). <https://doi.org/10.1787/f6d12db7-en>
- OECD (2025). *Results from TALIS Starting Strong 2024. Strengthening early childhood education and care*. <https://doi.org/10.1787/20afo8co-en>
- Olk, T. & Speck, K. (2012). Kooperation von Jugendhilfe und Schule. In W. Thole (Hrsg.), *Grundriss Soziale Arbeit. Ein einführendes Handbuch* (4. Aufl., S. 355–360). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-94311-4_18
- Otterpohl, N. & Wild, E. (2017). Kooperation zwischen Elternhaus und Schule im Kontext der schulischen Leistungsentwicklung. In B. Kracke & P. Noack (Hrsg.), *Handbuch Entwicklungs- und Erziehungspsychologie* (S. 1–15). Springer. https://doi.org/10.1007/978-3-642-54061-5_13-1
- Panadero, E., Andrade, H. & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher*, 45(1), 13–31. <https://doi.org/10.1007/s13384-018-0258-y>
- Panadero, E. & Jonsson, A. (2013). The use of rubrics for formative assessment revisited: A review. *Educational Research Review*, 9, 129–144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A. & Siegle, T. (Hrsg.). (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Waxmann.
- Paredes, M. C. (2011). *Parent involvement as an instructional strategy: Academic parent-teacher teams* [Dissertation]. Arizona State University. <https://www.cpedinitiative.org/assets/docs/paredes.pdf>
- Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B. & Pachan, M. (2008). *The positive impact of social and emotional learning for Kindergarten to eighth-grade students. Findings from three scientific reviews*. Collaborative for Academic, Social, and Emotional Learning (CASEL). <https://files.eric.ed.gov/fulltext/ED505370.pdf>



- Penny, A. R. & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74(2), 215–253. <https://doi.org/10.3102/00346543074002215>
- Pianta, R. C., La Paro, K. M. & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Paul H. Brookes Publishing Co.
- Piedade, F., Ferreira, T., Lemos, A., Guedes, C., Alves, D., Grande, C., Leal, T., Vatou, A. & Cadima, J. (2026). Integrating the PERMA model and SWPBS approach: Effects on portuguese early childhood teachers' well-being and children's self-regulation. *Early Childhood Research Quarterly*, 74, 189–198. <https://doi.org/10.1016/j.ecresq.2025.09.008>
- Pierce, R. & Chick, H. (2011). Teachers' beliefs about statistics education. In C. Batanero, G. Burrill & C. Reading (Hrsg.), *Teaching statistics in school mathematics-challenges for teaching and teacher education. A joint ICMI/IASE Study: the 18th ICMI Study* (S. 151–162). Springer.
- Pietsch, M. (2011). *Nutzen und Nutzung der Schulinspektion. Befunde der Hamburger Schulleitungsbefragung*. Behörde für Schule und Berufsbildung Hamburg; Institut für Bildungsmonitoring. https://e-pub.sub.uni-hamburg.de/epub/volltexte/2021/128812/pdf/pdf_zufriedenheitsstudie_2011.pdf
- Popp, K., Melzer, C. & Methner, A. (2023). *Förderpläne entwickeln und umsetzen* (4. Aufl.). Ernst Reinhardt.
- Prabhakar, G. V., Bhargavi, G. S., Reddy, P. R. & Shehab, M. (2023). Influence of performance feedback and academic performance on parent-family involvement and parent satisfaction in US schools. *Journal of Community Psychology*, 52, 27–38. <https://doi.org/10.1002/jcop.23083>
- Praetorius, A.-K. & Charalambous, C. Y. (Hrsg.). (2023). *Theorizing teaching. Current status and open issues*. Springer. <https://doi.org/10.1007/978-3-031-25613-4>
- Praetorius, A.-K., Jentsch, A., Luoto, J., Keller, S. D. & Fauth, B. (2025). Context in research on teaching quality: bringing a complex idea into the spotlight. *School Effectiveness and School Improvement*, 36(2), 263–296. <https://doi.org/10.1080/09243453.2025.2482577>
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM – Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Prediger, S., Dröse, J., Stahnke, R. & Ademmer, C. (2023). Teacher expertise for fostering at-risk students' understanding of basic concepts: Conceptual model and evidence for growth. *Journal of Mathematics Teacher Education*, 26(4), 481–508. <https://doi.org/10.1007/s10857-022-09538-3>
- Prediger, S., Fischer, C., Selter, C. & Schöber, C. (2019). Combining material- and community-based implementation strategies for scaling up: The case of supporting low-achieving middle school students. *Educational Studies in Mathematics*, 102(3), 361–378. <https://doi.org/10.1007/s10649-018-9835-2>
- Prediger, S. & Karst, K. (i. E.). *Datengestützte Unterrichtsentwicklung erfordert fachdidaktische Fokussierung und systematische Unterstützung*.
- Prediger, S., Rösike, K.-A. & Wischgoll, A. (2025). Beyond basic skills: An effective foundation intervention for low-achieving fifth graders' understanding of basic concepts. *Studies in Educational Evaluation*, 85, Artikel 101452. <https://doi.org/10.1016/j.stueduc.2025.101452>
- Prediger, S., Selter, C., Götze, D., Halleemann, S., Holzäpfel, L., Kreuziger, A., Pant, H. A. & Rösken-Winter, B. (2024). *QuaMath – Unterrichts- und Fortbildungsqualität in Mathematik entwickeln: Konzept des Zehnjahres-Programms von DZLM und KMK* (GDM-Mitteilungen Nr. 116). <https://ojs.didaktik-mathematik.de/index.php/mgdm/article/view/1226>



- Prenzel, M., Artelt, C., Baumert, J., Blum, W. & Hammann, M. (Hrsg.). (2008). *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich*. Waxmann.
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (Hrsg.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Waxmann.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. & Schiefele, U. (Hrsg.). (2005). *PISA 2003. Der zweite Vergleich der Länder in Deutschland – Was wissen und können Jugendliche?* Waxmann.
- Prenzel, M., Carstensen, C. H., Senkbeil, M., Ostermeier, C. & Seidel, T. (2005). Wie schneiden SINUS-Schulen bei PISA ab? *Zeitschrift für Erziehungswissenschaft*, 8(4), 540–562.
<https://doi.org/10.1007/s11618-005-0158-6>
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Waxmann. <https://elibrary.utb.de/doi/book/10.31244/9783830980018>
- Preuße, D., Pohl, J. & Gärtner, H. (2019). *Wahrgenommene Auswirkungen der Schulinspektion aus Sicht von Schulleitungen und Schulaufsicht in Berlin*. Institut für Schulqualität der Länder Berlin und Brandenburg e. V. (ISQ). https://www.isq.berlin/wordpress/wp-content/uploads/2019/09/Ergebnisbericht_Berlin_2019-07.pdf
- Prilop, C. N., Weber, K. E. & Kleinknecht, M. (2019). Entwicklung eines video- und textbasierten Instruments zur Messung kollegialer Feedbackkompetenz von Lehrkräften. In T. Ehmke, P. Kuhl & M. Pietsch (Hrsg.), *Lehrer. Bildung. Gestalten. Beiträge zur empirischen Forschung in der Lehrerbildung* (S. 153–163). Beltz Juventa.
- Prokupek, L., Hummel, T. G., Blaurock, S., Cohen, F. & Anders, Y. (2025). Potential of a digital parenting app to support parents of toddlers: Relations between the intensity of app use, language-related parental self-efficacy and the home literacy environment. *British Educational Research Journal*, 52(1), 643–666. <https://doi.org/10.1002/berj.70022>
- Pukrop, J. (2019). *Rückmeldungen aus Schulleistungstests an Lehrkräfte durch interaktive Informationsvisualisierungen* [Dissertation]. Universität Bremen. <https://media.suub.uni-bremen.de/entities/publication/24da4eb7-80c0-4afa-8419-2924d54fa773>
- Pushor, D. (2012). Tracing my research on parent engagement: Working to interrupt the story of school as protectorate. *Action in Teacher Education*, 34(5-6), 464–479.
<https://doi.org/10.1080/01626620.2012.729474>
- QUA-LiS NRW (o. J.). *Schulabsentismus*. <https://www.qua-lis.nrw.de/schulentwicklung-nrw/schulentwicklung/uebergreifende-themen/erziehung-und-praevention-g>
- Rakoczy, K., Buff, A. & Lipowsky, F. (2005). *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie. "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis". Teil 1 Befragungsinstrumente* (Materialien zur Bildungsforschung Nr. 13). Gesellschaft zur Förderung Pädagogischer Forschung. <https://doi.org/10.25656/01:3106>
- RatSWD (2022). *Aufbau eines Bildungsverlaufsregisters: Datenschutzkonform und forschungsfreundlich. Positionspapier des RatSWD*. Rat für Sozial- und Wirtschaftsdaten. <https://www.konsortswd.de/wp-content/uploads/Positionspapier-RatSWD-Aufbau-eines-Bildungsverlaufsregisters.pdf>
- Rauschenbach, T. & Meiner-Teubner, C. (2024). Von der Jugendwohlfahrt bis zum Ganzttag: 100 Jahre Kinder- und Jugendhilfe. Der stille Aufstieg der Sozial- und Erziehungsberufe – eine Rekonstruktion. In G. Graßhoff & M. Sauerwein (Hrsg.), *Das Personal am Ganzttag. Bildung, Betreuung und Erziehung in Zeiten des Fachkräftemangels – Herausforderungen für Kinder- und Jugendhilfe und Schule* (S. 14–41). Beltz Juventa.



- Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., Auquier, P., Power, M., Abel, T., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J. & European Kidscreen Group (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics & Outcomes Research*, 5(3), 353–364. <https://doi.org/10.1586/14737167.5.3.353>
- Reeves, A. R. (2011). *Where great teaching begins. Planning for student thinking and learning*. ASCD.
- Reichert-Garschhammer, E., Knoll, S., Helm, J., Harbecke, L., Möncke, U., Holand, G. & Lorenz, S. (2025). *KitaApps – Apps und Softwarelösungen für mittelbare pädagogische Aufgaben in der Kita IFP-Expertise* (3., überarb. Aufl.). Staatsinstitut für Frühpädagogik und Medienkompetenz. https://cdn.table.media/assets/wp-content/uploads/2025/05/28212544/KitaApps_IFP-Expertise_Auflage_3_Mai_2025.pdf
- Reiß, F., Behn, S., Erhart, M., Strelow, L., Kaman, A., Ottová-Jordan, V., Bilz, L., Moor, I. & Ravens-Sieberer, U. (2024). Subjektive Gesundheit und psychosomatische Beschwerden von Kindern und Jugendlichen in Deutschland: Ergebnisse der HBSC-Studie 2009/10 – 2022. *Journal of Health Monitoring*, 9(1), 7–24. <https://doi.org/10.25646/11867>
- Reiss, K., Klieme, E., Köller, O. & Stanat, P. (Hrsg.). (2017). *PISA Plus 2012 – 2013. Kompetenzentwicklung im Verlauf eines Schuljahres*. Zeitschrift für Erziehungswissenschaft, 33. Sonderheft. Springer VS.
- Relevant Assessment and Pedagogies for Inclusive Digital Education (2023). *Guidelines for ethical use of data in education*. <https://rapide-project.eu/web/sites/default/files/2023-03/APPENDIX%203%20-%20Guidelines%20on%20ethical%20use%20of%20data%20in%20education.pdf>
- Renkl, A. (2022). Meta-analyses as a privileged information source for informing teachers' practice? *Zeitschrift für Pädagogische Psychologie*, 36(4), 217–231. <https://doi.org/10.1024/1010-0652/a000345>
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. Hogrefe.
- Richter, I. (1994). Theorien der Schulautonomie. *Recht der Jugend und des Bildungswesens*, 1, 5–16.
- Ricking, H., Dunkake, I., Wiegand, J., Arnkens, C., Pielage, V., Fischer, A.-C., Sannen, L., Menting, A. & Frauen, T. (2024). *Jeder Schultag zählt. Praxishandbuch für die Schule zur Prävention und Intervention bei Absentismus* (3., aktual. Aufl.). Joachim Herz Stiftung. https://doi.org/10.53200/JHS_Jeder_Schultag_zaeht
- Rimm-Kaufman, S. E., Soland, J. & Kuhfeld, M. (2025). Social and emotional competency development from fourth to 12th grade: Relations to parental education and gender. *American Psychologist*, 80(3), 359–374. <https://doi.org/10.1037/amp0001357>
- Röhl, S. (2021). Effects of student feedback on teaching and classes: An overview and meta-analysis of intervention studies. In W. Rollett, H. J. E. Bijlsma & S. Röhl (Hrsg.), *Student feedback on teaching in schools. Using student perceptions for the development of teaching and teachers* (S. 139–156). Springer. https://doi.org/10.1007/978-3-030-75150-0_9
- Rollett, W., Bijlsma, H. J. E. & Röhl, S. (Hrsg.). (2021). *Student feedback on teaching in schools. Using student perceptions for the development of teaching and teachers*. Springer. <https://doi.org/10.1007/978-3-030-75150-0>
- Rönnau-Böse, M., Anders, Y., Fröhlich-Gildhoff, K., Blaurock, S., Burghardt, L., Hausladen, K., Limberger, J., Lorenzen, A., Pasquale, D., Pult, G., Oppermann, E. & Wolf, K. M. (2023). Evaluationsstudie zur Wirksamkeit des Gesetzes zur Weiterentwicklung der Qualität und zur Verbesserung der Teilhabe in Tageseinrichtungen und in der Kindertagespflege (KiTa-Qualitäts- und Teilhabeverbesserungsgesetz – KiQuTG). Abschlussbericht. In Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend (Hrsg.), *Zweiter Bericht der Bundesregierung zur Evaluation des Gesetzes zur*



- Weiterentwicklung der Qualität und zur Verbesserung der Teilhabe in Tageseinrichtungen und in der Kindertagespflege (KiQuTG)*. <https://www.bmbfsfj.bund.de/resource/blob/228532/e277c994e4e58be8c31031f045eb58ef/zweiter-bericht-der-bundesregierung-zur-evaluation-des-kiquotg-data.pdf>
- Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of Education Review*, 23(2), 143–151. [https://doi.org/10.1016/S0272-7757\(03\)00081-5](https://doi.org/10.1016/S0272-7757(03)00081-5)
- Roßbach, H.-G., Tietze, W., Kluczniok, K. & Nattfort, R. (Hrsg.). (2018). *Kindergarten-Skala-Erweiterung (KES-E)*. Deutsche Fassung der ECERS-E. Verlag das Netz.
- Sabio, J. D., Balagtas, M. U. & David, A. P. (2015). “Backwash Effects” of testing on learning mathematics. *The Normal Lights*, 9(2), 156–179. <https://doi.org/10.56278/tnl.v9i2.130>
- Sacher, W. (2012). Erziehungs- und Bildungspartnerschaften in der Schule: zum Forschungsstand. In W. Stange, R. Krüger, A. Henschel & C. Schmitt (Hrsg.), *Erziehungs- und Bildungspartnerschaften. Grundlagen und Strukturen von Elternarbeit* (S. 232–243). Springer VS. https://doi.org/10.1007/978-3-531-94279-7_24
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144. <https://doi.org/10.1007/BF00117714>
- Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 129–152). Springer. https://doi.org/10.1007/978-3-658-06604-8_5
- Sälzer, C., Ricking, H. & Feldhaus, M. (2024). Addressing school absenteeism through monitoring: A review of evidence-based educational policies and practices. *Education Sciences*, 14, Artikel 1365. <https://doi.org/10.3390/educsci14121365>
- Sánchez Puerta, M. L., Valerio, A. & Gutiérrez Bernal, M. I. (2016). *Taking stock of programs to develop socio-emotional skills. A systematic review of program evidence*. World Bank Group. <https://doi.org/10.1596/978-1-4648-0872-2>
- Scheer, D. & Melzer, C. (2020). Trendanalyse der KMK-Statistiken zur sonderpädagogischen Förderung 1994 bis 2019. *Zeitschrift für Heilpädagogik*, 71. <https://doi.org/10.25656/01:24724>
- Scheerens, J. (2008). *Review of research on school and instructional effectiveness*. University of Twente.
- Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, 61(3), 257–273. <https://doi.org/10.1080/00131881.2019.1625716>
- Schildkamp, K., Lai, M. K. & Earl, L. M. (Hrsg.). (2013). *Data-based decision making in education. Challenges and opportunities*. Springer. <https://doi.org/10.1007/978-94-007-4816-3>
- Schildkamp, K., Poortman, C. L., Ebbeler, J. & Pieters, J. M. (2019). How school leaders can build effective data teams: Five building blocks for a new wave of data-informed decision making. *Journal of Educational Change*, 20(3), 283–325. <https://doi.org/10.1007/s10833-019-09345-3>
- Schildkamp, K., Poortman, C., Luyten, H. & Ebbeler, J. (2017). Factors promoting and hindering data-based decision making in schools. *School Effectiveness and School Improvement*, 28(2), 242–258. <https://doi.org/10.1080/09243453.2016.1256901>



- Schildkamp, K., Vanhoof, J., van Petegem, P. & Visscher, A. J. (2012). The use of school self-evaluation results in the Netherlands and Flanders. *British Educational Research Journal*, 38(1), 125–152. <https://doi.org/10.1080/01411926.2010.528556>
- Schildkamp, K. & Visscher, A. J. (2009). Factors influencing the utilisation of a school self-evaluation instrument. *Studies in Educational Evaluation*, 35(4), 150–159. <https://doi.org/10.1016/j.stueduc.2009.12.001>
- Schildkamp, K., Visscher, A. J. & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement*, 20(1), 69–88. <https://doi.org/10.1080/09243450802605506>
- Schleswig-Holsteinischer Landtag (2016). *Drucksache 18/3719. Bericht der Landesregierung zur schulischen Qualitätsentwicklung in Schleswig-Holstein*. <https://www.landtag.ltsh.de/info-thek/wahl18/drucks/3700/drucksache-18-3719.pdf>
- Schnell, R. (2022). *Verknüpfung von Bildungsdaten in einem Bildungsregister mittels Record-Linkage auf Basis von Personenmerkmalen. Expertise für das Bundesministerium für Bildung und Forschung*. Research Methodology Group Universität Duisburg-Essen. https://duepublico2.uni-due.de/servlets/MCR-FileNodeServlet/duepublico_derivate_00076069/wp-grlc-2022-03.pdf
- Schnitzler, C., Henschel, S., Stehr, C., Ferencik-Lehmkuhl, D., Jost, J. & Stanat, P. (2024). Lesekompetenz in der Sekundarstufe I datengestützt fördern. Untersuchung der Wirksamkeit einer Lehrkräftefortbildung zur datengestützten Unterrichtsentwicklung mit den VERA-Ergebnissen. In Trägerkonsortium BiSS-Transfer (Hrsg.), *BiSS-Transfer: Potenziale sichern. Arbeitsergebnisse, Länderperspektiven und Impulse* (S. 45–50): Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache. <https://doi.org/10.3278/9783763978014>
- Schoenfeld, A. H. (2018). Video analyses for research and professional development: The Teaching for Robust Understanding (TRU) framework. *ZDM – Mathematics Education*, 50(3), 491–506. <https://doi.org/10.1007/s11858-017-0908-y>
- Scholes, V. (2016). The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, 64(5), 939–955. <https://doi.org/10.1007/s11423-016-9458-1>
- Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *The Future of Children*, 27(1), 137–155. <https://doi.org/10.1353/foc.2017.0007>
- Schonert-Reichl, K. A. (2019). Advancements in the landscape of social and emotional learning and emerging topics on the horizon. *Educational Psychologist*, 54(3), 222–232. <https://doi.org/10.1080/00461520.2019.1633925>
- Schult, J. (2022). VERA fair-bessern: Ein Vergleich von Strategien zur Berücksichtigung von Kontextbedingungen bei schulischen Vergleichsarbeiten. *Journal for Educational Research Online*, 14(2), 3–24. <https://doi.org/10.31244/jero.2022.02.01>
- Schulte, K., Fickermann, D. & Lücken, M. (2016). Das Hamburger Prozessmodell datengestützter Schulentwicklung. *Die Deutsche Schule*, 108(2), 176–190. <https://doi.org/10.25656/01:25956>
- Schürer, S. & Lintorf, K. (2023). Erziehungs- und Bildungspartnerschaften mit Eltern von Kindern mit sonderpädagogischem Förderbedarf. Ein einlösbarer Anspruch am Grundschulübergang? In M. Haider, R. Böhme, S. Gebauer, C. Gößinger, M. Munser-Kiefer & A. Rank (Hrsg.), *Nachhaltige Bildung in der Grundschule* (S. 158–164). Klinkhardt.



- Schütte, K., Hasselhorn, M. & Köller, O. (2024). Bildungsungleichheit besser verstehen und vermindern. *Frühe Bildung*, 13(2), 65–70. <https://doi.org/10.1026/2191-9186/a000662>
- Schwippert, K. (2005). Vergleichende Lernstandsuntersuchungen, Bildungsstandards und die Steuerung von schulischen Bildungsprozessen. *Berufs- und Wirtschaftspädagogik online*(8), 1–14.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). SAGE.
- Seidel, T., Prenzel, M., Duit, R. & Lehrke, M. (2003). *Technischer Bericht zur Videostudie "Lehr-Lern-Prozesse im Physikunterricht"*. Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN). https://www.researchgate.net/publication/264436084_A_Review_of_Early_Numeracy_Interventions_for_Children_at_Risk_in_Mathematics
- Selter, C. (2017). Förderorientierte Diagnose und diagnosegeleitete Förderung. In A. Fritz, S. Schmidt & G. Ricken (Hrsg.), *Handbuch Rechenschwäche. Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie* (3., vollst. überarb. Aufl., S. 375–395). Beltz.
- Sendzik, N., Helbig, M., Demski, D., Bellenberg, G., Eiden, S., Edelstein, B. & Hugo, J. (Hrsg.). (2025). *Ungleich fördern – gerecht steuern. Begründungen, Umsetzungen und Wirkungserwartungen einer bedarfsorientierten Ressourcensteuerung im Bildungssystem*. Die Deutsche Schule, 20. Beiheft. Waxmann. <https://doi.org/10.31244/9783830999812>
- Sergis, S. & Sampson, D. G. (2017). Teaching and learning analytics to support teacher inquiry: A systematic literature review. In A. Peña-Ayala (Hrsg.), *Learning analytics: Fundamentals, applications, and trends. A view of the current state of the art to enhance e-learning* (S. 25–63). Springer. https://doi.org/10.1007/978-3-319-52977-6_2
- Setzer, S., Rohde, J., Güttgemanns, V. & Rothe, P. (2024). Die Cell-Key-Methode in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. Teil 1: Vorstellung des neuen Geheimhaltungsverfahrens. *WISTA – Wirtschaft und Statistik*(3), 31–45. https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2024/03/cell-key-methode-teil1-032024.pdf?__blob=publicationFile
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Siraj, I., Kingston, D. & Melhuish, E. (2015). *Assessing quality in early childhood education and care. Sustained Shared Thinking and Emotional Well-being (SSTEW) Scale for 2-5 year-old provision*. Trentham Books Ltd.
- Slavin, R. E., Cheung, A. C. K., Holmes, G., Madden, N. A. & Chamberlain, A. (2013). Effects of a data-driven district reform model on state assessment outcomes. *American Educational Research Journal*, 50(2), 371–396. <https://doi.org/10.3102/0002831212466909>
- Sliwka, A., Klopsch, B. & Deinhardt, L. (2023). Digital, nachhaltig, gerecht. Eine strategische Kernroutine zur adaptiven Förderung in der Schule. *MedienPädagogik: Zeitschrift für Theorie Und Praxis Der Medienbildung*, 52, 169–190. <https://doi.org/10.21240/mpaed/52/2023.02.09.X>
- Solarski, L. A. (2021). *Clarifying coaching: A mixed methods analysis of a math content-focused model and its impact on teachers' practice, content knowledge, and dispositions* [Dissertation]. Loyola University Chicago. https://ecommons.luc.edu/luc_diss/3851/
- Sousa, E. B. G. de, Alexandre, B., Ferreira Mello, R., Pontual Falcão, T., Vesin, B. & Gašević, D. (2021). Applications of learning analytics in high schools: A systematic literature review. *Frontiers in Artificial Intelligence*, 4, Artikel 737891. <https://doi.org/10.3389/frai.2021.737891>



- Souvignier, E., Förster, N. & Salaschek, M. (2014). quop: Ein Ansatz internetbasierter Lernverlaufsdiagnostik mit Testkonzepten für Lesen und Mathematik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (S. 239–256). Hogrefe.
- Souvignier, E. & Hasselhorn, M. (2018). Formatives Assessment. *Zeitschrift für Erziehungswissenschaft*, 21(4), 693–696. <https://doi.org/10.1007/s11618-018-0839-6>
- Speck, K., Peponis, M., Walter, J., Ruprecht, T. M., Bistritzky, H. & Ott-Jacobs, S. (2020). Kooperation von Kinder- und Jugendpsychiatrie/-psychotherapie, Schule und Kinder- und Jugendhilfe. Innovationsprojekt „Drei für Eins“ für psychisch belastete Kinder und Jugendliche in Hamburg. *unsere jugend*, 72(11-12), 492–497. <https://doi.org/10.2378/uj2020.art79d>
- Speck, K., Schäfer, A., Schmidt, N., Stauvermann, L., Walter, J., Ott-Jacobs, S., Ruprecht, T. M., Weber, R., Gringel, J., Peponis, M., Bistritzky, H. & Friedsam, P. (2023). Kooperation von Kinder- und Jugendhilfe, Schule und Kinder- und Jugendpsychiatrie, -psychosomatik und -psychotherapie: Frühzeitige Diagnostik, Behandlung und Förderung der gesunden Entwicklung und Bildungsteilhabe von Kindern und Jugendlichen im Innovationsprojekt DreiFürEins. In G. Groen, M. Peters, M. Schwabe & J. Weber (Hrsg.), *Krisenhafte Verläufe in den Erziehungshilfen. Kooperation, Risikopartnerschaften, Verantwortungsgemeinschaften* (S. 219–226). Beltz Juventa.
- Stacey, K., Steinle, V., Price, B. & Gvozdenko, E. (2018). Specific mathematics assessments that reveal thinking: An online tool to build teachers' diagnostic competence and support teaching. In T. Leuders, K. Philipp & J. Leuders (Hrsg.), *Diagnostic competence of mathematics teachers. Unpacking a complex construct in teacher education and teacher practice* (S. 241–261). Springer. https://doi.org/10.1007/978-3-319-66327-2_13
- Stadt Mönchengladbach (2023). *Handlungsansätze bei Schulabsentismus. Ein Leitfaden für Schulen in Mönchengladbach*. https://www.moenchengladbach.de/fileadmin/user_upload/FB40/Schulpsychologischer_Dienst/Leitfaden_Schulabsentismus_web.pdf
- Stanat, P., Klieme, E. & Pant, H. A. (Hrsg.). (2025). *Bildungsstandards in Deutschland*. Zeitschrift für Pädagogik, 71. Beiheft. Beltz.
- Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S. & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann. <https://doi.org/10.31244/9783830996064>
- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (2022a). *Basale Kompetenzen vermitteln – Bildungschancen sichern. Perspektiven für die Grundschule. Gutachten der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <https://doi.org/10.25656/01:25542>
- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (2022b). *Digitalisierung im Bildungssystem: Handlungsempfehlungen von der Kita bis zur Hochschule. Gutachten der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <https://doi.org/10.25656/01:25273>
- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (2023). *Lehrkräftegewinnung und Lehrkräftebildung für einen hochwertigen Unterricht. Gutachten der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <https://doi.org/10.25656/01:28059>
- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (2024). *Demokratiebildung als Auftrag der Schule – Bedeutung des historischen und politischen Fachunterrichts sowie Aufgabe aller Fächer und der Schulentwicklung. Stellungnahme der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <https://doi.org/10.25656/01:30061>



- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (2025). *Kompetenzen für den erfolgreichen Übergang von der Sekundarstufe I in die berufliche Ausbildung sichern. Gutachten der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <https://doi.org/10.25656/01:32815>
- Standing International Conference of Inspectorates (2013). *The Bratislava memorandum on inspection and innovation*.
- Stange, W. (2012). Erziehungs- und Bildungspartnerschaften – Grundlagen, Strukturen, Begründungen. In W. Stange, R. Krüger, A. Henschel & C. Schmitt (Hrsg.), *Erziehungs- und Bildungspartnerschaften. Grundlagen und Strukturen von Elternarbeit* (S. 12–39). Springer VS. https://doi.org/10.1007/978-3-531-94279-7_1
- Stange, W., Krüger, R., Henschel, A. & Schmitt, C. (Hrsg.). (2012). *Erziehungs- und Bildungspartnerschaften. Grundlagen und Strukturen von Elternarbeit*. Springer VS. <https://doi.org/10.1007/978-3-531-94279-7>
- Stauber, A., Stahl, J., Bruneforth, M. & Illetschko, M. (2024). Die individuelle Kompetenzmessung PLUS: Pädagogische Diagnostik im Spannungsfeld von Individualförderung und Systemmonitoring. In C. Schreiner, G. Schauer & C. Kraler (Hrsg.), *Pädagogische Diagnostik und Lehrer:innenbildung: Bildungswissenschaftliche und fachdidaktische Perspektiven* (S. 157–172). Klinkhardt.
- Strunk, K. O., Marsh, J. A., Bush-Mecenas, S. C. & Duque, M. R. (2016). The best laid plans: An examination of school plan quality and implementation in a school improvement initiative. *Educational Administration Quarterly*, 52(2), 259–309. <https://doi.org/10.1177/0013161X15616864>
- Sundermann, B. & Selter, C. (2006). *Beurteilen und Fördern im Mathematikunterricht*. Cornelsen Scriptor.
- Tan, C. Y., Gao, L. & Shi, M. (2022). Second-order meta-analysis synthesizing the evidence on associations between school leadership and different school outcomes. *Educational Management Administration & Leadership*, 50(3), 469–490. <https://doi.org/10.1177/1741143220935456>
- Tarkian, J., Lankes, E.-M. & Thiel, F. (2019). Externe Evaluation – Konzeption und Implementation in den 16 Ländern. In F. Thiel, J. Tarkian, E.-M. Lankes, N. Maritzen, T. Riecke-Baulecke & A. Kroupa (Hrsg.), *Datenbasierte Qualitätssicherung und -entwicklung in Schulen. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland* (S. 105–183). Springer VS. https://doi.org/10.1007/978-3-658-23240-5_5
- Tarkian, J., Maritzen, N., Eckert, M. & Thiel, F. (2019). Vergleichsarbeiten (VERA) – Konzeption und Implementation in den 16 Ländern. In F. Thiel, J. Tarkian, E.-M. Lankes, N. Maritzen, T. Riecke-Baulecke & A. Kroupa (Hrsg.), *Datenbasierte Qualitätssicherung und -entwicklung in Schulen. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland* (S. 41–103). Springer VS. https://doi.org/10.1007/978-3-658-23240-5_4
- Tarkian, J., Riecke-Baulecke, T. & Thiel, F. (2019). Interne Evaluation – Konzeption und Implementation in den 16 Ländern. In F. Thiel, J. Tarkian, E.-M. Lankes, N. Maritzen, T. Riecke-Baulecke & A. Kroupa (Hrsg.), *Datenbasierte Qualitätssicherung und -entwicklung in Schulen. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland* (S. 185–229). Springer VS. https://doi.org/10.1007/978-3-658-23240-5_6
- Tetzlaff, L., Simonsmeier, B., Peters, T. & Brod, G. (2025). A cornerstone of adaptivity – A meta-analysis of the expertise reversal effect. *Learning and Instruction*, 98, Artikel 102142. <https://doi.org/10.1016/j.learninstruc.2025.102142>



- Thiel, F., Ophardt, D. & Piwowar, V. (2013). *Abschlussbericht des Projekts Kompetenzen des Klassenmanagements (KODEK). Entwicklung und Evaluation eines Fortbildungsprogramms für Lehrkräfte zum Klassenmanagement*. Freie Universität Berlin. <https://www.ewi-psy.fu-berlin.de/erziehungswissenschaft/arbeitsbereiche/schulentwicklungsforschung/downloads/Abschlussbericht-KODEK1.pdf>
- Thiel, F., Schewe, C. M., Muslic, B., Lankes, E.-M., Maritzen, N. & Riecke-Baulecke, T. (Hrsg.). (2022). *Personalentwicklung in Schulen als Führungsaufgabe. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland*. Springer VS. <https://doi.org/10.1007/978-3-658-36925-5>
- Thiel, F. & Tarkian, J. (2019). Rahmenkonzepte zur Definition von Schulqualität in den 16 Ländern. In F. Thiel, J. Tarkian, E.-M. Lankes, N. Maritzen, T. Riecke-Baulecke & A. Kroupa (Hrsg.), *Datenbasierte Qualitätssicherung und -entwicklung in Schulen. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland* (S. 15–39). Springer VS. https://doi.org/10.1007/978-3-658-23240-5_3
- Thiel, F., Tarkian, J., Lankes, E.-M., Maritzen, N., Riecke-Baulecke, T. & Kroupa, A. (Hrsg.). (2019). *Datenbasierte Qualitätssicherung und -entwicklung in Schulen. Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland*. Springer VS. <https://doi.org/10.1007/978-3-658-23240-5>
- Thurm, D. & Graewert, L. A. (2022). *Digitale Mathematik-Lernplattformen in Deutschland. MINTUS – Beiträge zur mathematisch-naturwissenschaftlichen Bildung*. Springer Spektrum. <https://doi.org/10.1007/978-3-658-37520-1>
- Tietze, W., Becker-Stoll, F., Bensel, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B., Keller, H. & Leyendecker, B. (Hrsg.). (2013). *Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK)*. Verlag das Netz.
- Tietze, W., Roßbach, H.-G. & Schuster, K.-M. (Hrsg.). (2017). *Kindergarten-Skala (KES-RZ): Revidierte Fassung mit Zusatzmerkmalen. Deutsche Fassung der Early Childhood Environment Rating Scale von Thelma Harms, Richard, M. Clifford und Debby Cryer*. Verlag das Netz.
- Timmermans, A. C., de Wolf, I. F., Bosker, R. J. & Doolaard, S. (2015). Risk-based educational accountability in Dutch primary education. *Educational Assessment, Evaluation and Accountability*, 27(4), 323–346. <https://doi.org/10.1007/s11092-015-9212-y>
- Timmermans, A. C., Naayer, H. M., Meelissen, M. R. M., Gubbels, J., Scheerens, J. & van der Werf, G. (2022). Cross-national achievement surveys and educational monitoring in the Netherlands. In L. Volante, S. V. Schnepf & D. A. Klinger (Hrsg.), *Cross-national achievement surveys for monitoring educational outcomes: Policies, practices and political reforms within the European Union* (S. 91–114). Publications Office of the European Union. <https://doi.org/10.2760/406165>
- Timperley, H. (2013). *Learning to practise: A paper for discussion*. https://thehub.sia.govt.nz/assets/documents/42549_Learning-To-Practise_o.pdf
- Timperley, H. (2014). *Promoting and supporting improvement in schools through external review*. <https://www.evidence.ero.govt.nz/documents/promoting-and-supporting-improvement-in-schools-through-external-review>
- Timperley, H., Parr, J. M. & Bertanees, C. (2009). Promoting professional inquiry for improved outcomes for students in New Zealand. *Professional Development in Education*, 35(2), 227–245. <https://doi.org/10.1080/13674580802550094>
- Torres, R. (2021). *Does test-based school accountability have an impact on student achievement and equity in education? A panel approach using PISA* (OECD Education Working Papers Nr. 250). <https://doi.org/10.1787/0798600f-en>



- Trautwein, U., Sliwka, A. & Dehmel, A. (2022). *Grundlagen für einen wirksamen Unterricht. Wirksamer Unterricht Band 1*. Institut für Bildungsanalysen Baden-Württemberg (IBBW). [https://ibbw-bw.de/site/pbs-bw-rebrush2024/get/documents_E552984303/KULTUS.Dachmandant/KULTUS/Dienststellen/ibbw/Empirische%20Bildungsforschung/Programme-und-Projekte/Wirksamer Unterricht/IBBW_WU1_Trautwein et al%20%282022%29_GrundlagenWirksamerUnterricht.pdf](https://ibbw-bw.de/site/pbs-bw-rebrush2024/get/documents_E552984303/KULTUS.Dachmandant/KULTUS/Dienststellen/ibbw/Empirische%20Bildungsforschung/Programme-und-Projekte/Wirksamer%20Unterricht/IBBW_WU1_Trautwein_et_al%20%282022%29_GrundlagenWirksamerUnterricht.pdf)
- Trilling, B. & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Jossey-Bass. <https://psycnet.apa.org/record/2009-18745-000>
- Tripses, J., Hunt, J., Kim, J. & Watkins, S. (2015). Leading into the future: Perceptions of school board presidents on the essential knowledge and skills for superintendent preparation programs. *Education Leadership Review*, 16(2), 36–54. <https://www.icpel.org/ed-leadership-review.html>
- Turney, K. & Kao, G. (2009). Barriers to school involvement: Are immigrant parents disadvantaged? *The Journal of Educational Research*, 102(4), 257–271. <https://doi.org/10.3200/JOER.102.4.257-271>
- Ulber, D. (2011). Zielvereinbarungen einer Schule-Evaluation der Zielvereinbarungsgespräche zwischen Schulleitung und Schulaufsicht am Beispiel Berlin. *Schulverwaltung: Fachzeitschrift für Schulentwicklung und Schulmanagement. Hessen, Rheinland-Pfalz*, 6, 182–185.
- Ulferts, H., Wolf, K. M. & Anders, Y. (2019). Impact of process quality in early childhood education and care on academic outcomes: Longitudinal meta-analysis. *Child Development*, 90(5), 1474–1489. <https://doi.org/10.1111/cdev.13296>
- van der Bij, T., Geijsel, F. P. & ten Dam, G. T. M. (2016). Improving the quality of education through self-evaluation in Dutch secondary schools. *Studies in Educational Evaluation*, 49, 42–50. <https://doi.org/10.1016/j.stueduc.2016.04.001>
- van der Linden, S., van der Meij, J. & McKenney, S. (2021). Teacher video coaching, from design features to student impacts: A systematic literature review. *Review of Educational Research*, 92(1), 114–165. <https://doi.org/10.3102/00346543211046984>
- van der Scheer, E. A. & Visscher, A. J. (2016). Effects of an intensive data-based decision making intervention on teacher efficacy. *Teaching and Teacher Education*, 60, 34–43. <https://doi.org/10.1016/j.tate.2016.07.025>
- van Geel, M., Keuning, T., Visscher, A. J. & Fox, J.-P. (2016). Assessing the effects of a school-wide data-based decision-making intervention on student achievement growth in primary schools. *American Educational Research Journal*, 53(2), 360–394. <https://doi.org/10.3102/0002831216637346>
- van Geel, M., Keuning, T., Visscher, A. J. & Fox, J.-P. (2017). Changes in educators' data literacy during a data-based decision making intervention. *Teaching and Teacher Education*, 64, 187–198. <https://doi.org/10.1016/j.tate.2017.02.015>
- van Gronigen, B. A. & Meyers, C. V. (2021). Exploring the association between short-cycle school improvement planning and student achievement in underperforming schools. *Journal of School Leadership*, 32(2), 339–361. <https://doi.org/10.1177/10526846211018207>
- Verger, A., Parcerisa, L., Pagès, M. & Camphuijsen, M. (2024). School autonomy with accountability as a cross-national policy model: diverse adoptions, practices and impacts. *Nordic Journal of Studies in Educational Policy*, 10(3), 167–179. <https://doi.org/10.1080/20020317.2024.2430706>
- Verkroost, J. (2024). Indicators in risk-based educational supervision. *Journal of Educational Sciences Research*, 14(4), 1–12. <https://doi.org/10.22521/JESR.2024.14.4.37>



- Visscher, A. J. (2002). A framework for studying school performance feedback systems. In A. J. Visscher & R. Coe (Hrsg.), *School improvement through performance feedback* (S. 41–72). Routledge.
- Visscher, A. J. (2021). On the value of data-based decision making in education: The evidence from six intervention studies. *Studies in Educational Evaluation*, 69, Artikel 100899. <https://doi.org/10.1016/j.stueduc.2020.100899>
- Visscher, A. J. & Coe, R. (2003). School performance feedback systems: Conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement*, 14(3), 321–349. <https://doi.org/10.1076/sesi.14.3.321.15842>
- Vodafone Stiftung Deutschland (2013). *Qualitätsmerkmale schulischer Elternarbeit: Ein Kompass für die partnerschaftliche Zusammenarbeit von Schule und Elternhaus*. https://www.vodafone-stiftung.de/wp-content/uploads/2020/02/Qualit%C3%A4tsmerkmale-schulischer-Elternarbeit_vodafone-stiftung_druckversion.pdf
- Vogel, S., Blum, W., Achmetli, K. & Krawitz, J. (2016). Qualifizierung von Lehrkräften zum konstruktiven Umgang mit zentralen Lernstandserhebungen – Ergebnisse aus dem Projekt VELM-8. *Journal für Mathematik-Didaktik*, 37(2), 319–348. <https://doi.org/10.1007/s13138-016-0092-6>
- vom Hofe, R., Pekrun, R., Kleine, M. & Götz, T. (2002). Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Konstruktion des Regensburger Mathematikleistungstests für 5. bis 10. Klassen, 83–100. <https://doi.org/10.25656/01:3940>
- von Eichhorn, G. & Freunberger, R. (2024). *Vergleich und Validierung unterschiedlicher Lernfortschrittsmodelle* (IQS-Report 1/2024). Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen. <https://doi.org/10.17888/iqsreport-2024-1>
- Wagner, W., Göllner, R., Helmke, A., Trautwein, U. & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, 28, 1–11. <https://doi.org/10.1016/j.learninstruc.2013.03.003>
- Wartenberg, G., Aldrup, K., Grund, S. & Klusmann, U. (2023). Satisfied and high performing? A meta-analysis and systematic review of the correlates of teachers' job satisfaction. *Educational Psychology Review*, 35(4), Artikel 114. <https://doi.org/10.1007/s10648-023-09831-4>
- Waterfield, D. A., Coleman, O. F., Welker, N. P., Kennedy, M. J., McDonald, S. D. & Cook, B. G. (2026). IEPs in the age of AI: Examining IEP goals written with and without ChatGPT. *Journal of Special Education Technology*, 41(1), 57–71. <https://doi.org/10.1177/01626434251324592>
- Wayman, J., Midgley, S. & Stringfield, S. (2017). Leadership for data-based decision making: Collaborative educator teams. In A. B. Danzig, K. M. Borman, B. A. Jones & W. F. Wright (Hrsg.), *Learner-Centered Leadership* (S. 189–206). Routledge.
- Weitkämper, F. & Rutter, S. (2023). Sozioanalyse und Habitussensibilität als Schlüssel für inklusions- und sozialraumorientierte Schulentwicklungsprozesse. In S. Schuppener, N. Leonhardt & R. Kruschel (Hrsg.), *Inklusive Schule im Sozialraum. Entwicklungsprozesse durch Kooperation und Interprofessionalität in herausfordernden Lagen* (S. 19–35). Springer VS.
- West, L. & Staub, F. C. (2003). *Content-focused coaching: Transforming mathematics lessons*. Heinemann.
- West, M. R., Pier, L., Fricke, H., Hough, H., Loeb, S., Meyer, R. H. & Rice, A. B. (2020). Trends in student social-emotional learning: Evidence from the first large-scale panel student survey. *Educational Evaluation and Policy Analysis*, 42(2), 279–303. <https://doi.org/10.3102/0162373720912236>



- Weyrauch, I., Wick, F., Zimmer-Müller, M., Hosenfeld, I. & Strasser, J. (2025). „Klar, arbeiten wir mit den VERA-Ergebnissen weiter“ – Impulse für die Weiterentwicklung von Schule und Unterricht durch (die Arbeit mit Ergebnissen aus) Vergleichsarbeiten in Grund- und Sekundarschulen [Poster]. BMBF-Bildungsforschungstagung (BiFo) 2025, Bonn. https://www.empirische-bildungsforschung-bmbfsfj.de/img/WeSU_Poster_Bifo2025.pdf
- White, M., Göllner, R. & Kleickmann, T. (2025). Reconsidering the measurement of teaching quality using a lens model: Measurement dilemmas and trade-offs. *School Effectiveness and School Improvement*, 36(2), 192–210. <https://doi.org/10.1080/09243453.2025.2482583>
- Wild, E. (2021). Eltern als Erziehungs- und Bildungspartner von Schule? Warum und wie welche Eltern stärker in das Schulleben ihrer Kinder involviert werden sollten und könnten. In Sachverständigenkommission des Neunten Familienberichts (Hrsg.), *Eltern sein in Deutschland. Materialien zum Neunten Familienbericht* (433-536). Deutsches Jugendinstitut. <https://doi.org/10.36189/DJ1232021>
- Winne, P. H. & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmermann (Hrsg.), *Motivation and self-regulated learning: Theory, research, and applications* (S. 297–314). Lawrence Erlbaum Associates Publishers.
- Wisniewski, B., Zierer, K. & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, Artikel 3087. <https://doi.org/10.3389/fpsyg.2019.03087>
- Wittig, J. & Schneider, R. (2022). Kompetenzstufenbesetzungen im Fach Deutsch. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (S. 41–66). Waxmann.
- Wohlstetter, P., Datnow, A. & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement*, 19(3), 239–259. <https://doi.org/10.1080/09243450802246376>
- Wolf, D., Schneider, S. & Schütz, A. (2020). *Inventar und Situational Judgement Test zur Erfassung sozial-emotionaler Kompetenz(en) (SEC-I und SEC-SJT) : Diagnostische Verfahren zur multimethodalen Erfassung sozial-emotionaler Kompetenz(en) bei jungen Menschen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung*. <https://doi.org/10.20378/irb-48863>
- Woo, S., Lee, S. & Kim, K. (2015). Carrot and stick?: Impact of a low-stakes school accountability program on student achievement. *Economics Letters*, 137, 195–199. <https://doi.org/10.1016/j.econlet.2015.10.007>
- Wößmann, L. (2007a). International evidence on school competition, autonomy, and accountability: A review. *Peabody Journal of Education*, 82(2-3), 473–497. <https://doi.org/10.1080/01619560701313176>
- Wößmann, L. (2007b). Extern überprüfte Standards, Schulautonomie und Wettbewerb: Chancen für das deutsche Schulsystem. *Recht der Jugend und des Bildungswesens*, 55(1), 64–83. <https://www.nomos-elibrary.de/10.5771/0034-1312-2007-1-64.pdf>
- Wößmann, L., Lüdemann, E., Schütz, G. & West, M. R. (2009). *School accountability, autonomy and choice around the world*. Edward Elgar Publishing. <https://doi.org/10.4337/9781035363452>
- Wu, H. & Shen, J. (2022). The association between principal leadership and student achievement: A multivariate meta-meta-analysis. *Educational Research Review*, 35, 100423. <https://doi.org/10.1016/j.edurev.2021.100423>



- Wurster, S. (2022). Förderliche und hinderliche Bedingungen für die datengestützte Unterrichtsentwicklung mit Vergleichsarbeiten, zentralen Prüfungen und internen Evaluationen. *Zeitschrift für Pädagogik*, 68(1), 95–116. <https://doi.org/10.3262/ZP2201095>
- Wurster, S., Bez, S. & Merk, S. (2023). Does learning how to use data mean being motivated to use it? Effects of a data use intervention on data literacy and motivational beliefs of pre-service teachers. *Learning and Instruction*, 88, Artikel 101806. <https://doi.org/10.1016/j.learninstruc.2023.101806>
- Wurster, S. & Richter, D. (2016). Nutzung von Schülerleistungsdaten aus Vergleichsarbeiten und zentralen Abschlussprüfungen für Unterrichtsentwicklung in Brandenburger Fachkonferenzen. *Journal for Educational Research Online*, 8(3), 159–183. <https://doi.org/10.25656/01:12820>
- Wurster, S., Richter, D., Schliesing, A. & Pant, H. A. (2013). Nutzung unterschiedlicher Evaluationsdaten an Berliner und Brandenburger Schulen: Rezeption und Nutzung von Ergebnissen aus Schulinspektion, Vergleichsarbeiten und interner Evaluation im Vergleich. In I. van Ackeren, M. Heinrich & F. Thiel (Hrsg.), *Evidenzbasierte Steuerung im Bildungssystem? Befunde aus dem BMBF-SteBis-Verbund. Die Deutsche Schule*, 12. Beiheft (S. 19–50). Waxmann.
- Xuan, Q., Cheung, A. C. K. & Sun, D. (2022). The effectiveness of formative assessment for enhancing reading achievement in K-12 classrooms: A meta-analysis. *Frontiers in Psychology*, 13, Artikel 990196. <https://doi.org/10.3389/fpsyg.2022.990196>
- Yan, Y. (2019). Making accountability work in basic education: Reforms, challenges and the role of the government. *Policy Design and Practice*, 2(1), 90–102. <https://doi.org/10.1080/25741292.2019.1580131>
- Yao, Y., Amos, M., Snider, K. & Brown, T. (2024). The impact of formative assessment on K-12 learning: A meta-analysis. *Educational Research and Evaluation*, 29(7-8), 452–475. <https://doi.org/10.1080/13803611.2024.2363831>
- Yeager, D. S. & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47(4), 302–314. <https://doi.org/10.1080/00461520.2012.722805>
- York, B. N., Loeb, S. & Doss, C. (2019). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. *Journal of Human Resources*, 54(3), 537–566. <https://doi.org/10.3368/jhr.54.3.0517-8756R>
- Yotyodying, S. & Wild, E. (2016). Predictors of the quantity and different qualities of home-based parental involvement: Evidence from parents of children with learning disabilities. *Learning and Individual Differences*, 49, 74–84. <https://doi.org/10.1016/j.lindif.2016.05.003>
- Yotyodying, S. & Wild, E. (2019). Effective family-school communication for students with learning disabilities: Associations with parental involvement at home and in school. *Learning, Culture and Social Interaction*, 22, Artikel 100317. <https://doi.org/10.1016/j.lcsi.2019.100317>
- Yu, B. & Zhang, G. (2023). The effects of learning analytics on online self-regulated learning: A meta analysis. In W. B. A. Tso, S. K. K. NG, L. Law & T. S. Bai (Hrsg.), *Educational Communications and Technology Yearbook. The post-pandemic landscape of education and beyond: Innovation and transformation: Selected Papers from the HKAECT 2022 International Conference* (S. 148–161). Springer. https://doi.org/10.1007/978-981-19-9217-9_10
- Zimmer, R. (2019). *BaSiK: Begleitende alltagsintegrierte Sprachentwicklungsbeobachtung in Kindertageseinrichtungen* (3. Aufl.). Herder.



Anhänge

Anhang 1 Zum methodischen Vorgehen

Die SWK hat im Rahmen der Arbeit an dem vorliegenden Gutachten im Jahr 2025 sowohl eine Abfrage bei den Kultusministerien der Länder („SWK-Länderabfrage Schule“) als auch bei den in der Jugend- und Familienministerkonferenz organisierten Ministerien durchgeführt („SWK-Länderabfrage Frühe Bildung“). Die Länderabfragen sind sowohl im Text als auch im folgenden Anhang wiedergegeben. Alle Länder haben sich dankenswerterweise an der Abfrage beteiligt.

Zusätzlich hat die SWK auch öffentliche Daten zur Situation in den Ländern recherchiert (vgl. zum Vorgehen auch Thiel et al., 2019). Wenn zusammenfassende Darstellungen im Text erfolgen, sind die entsprechenden Quellen im Anhang aufgeführt.



Anhang 2 Weitere in den Ländern eingesetzte diagnostische bzw. formative Instrumente

Diagnostische und formative Instrumente, die in mehreren Ländern eingesetzt werden:

Name	Bildungsetappe	Nutzung in... Ländern
Potenzial und Perspektiven (2P)	Stufenübergreifend	8
Individuelle Lernstandsanalyse (ILeA)	Stufenübergreifend	6
Mathe sicher können (MSK)	Stufenübergreifend	5
quop	Stufenübergreifend	5
Bettermarks	Stufenübergreifend	4
Lernlinie	Stufenübergreifend	2
Lernstand	Stufenübergreifend	2
Meister Cody	Stufenübergreifend	2
Mit Mirola durch den Zauberwald	Primarstufe	2

Diagnostische und formative Instrumente, die jeweils nur in einem Land eingesetzt werden:

Name	Bildungsetappe
Bayrisches Screening des individuellen Sprachstands (BaSIS)	Frühe Bildung
Deutsch für den Schulstart	Frühe Bildung
Bayrisches Lesescreening (BYLES)	Primarstufe
Bielefelder Rechentest (BIRTE 2)	Primarstufe
Bielefelder Screening zur Früherkennung von Lese- und Rechtschreibschwierigkeiten (BISC)	Primarstufe
Deutscher Mathematiktest (DEMAT)	Primarstufe
Diagnostische Einschätzskalen zur Beurteilung des Entwicklungsstands und der Schulfähigkeit	Primarstufe
eKidz Spracherwerbs-App	Primarstufe
Elementarmathematisches Basisinterview	Primarstufe
Fähigkeitsindikatoren in der Primarstufe (FIPS)	Primarstufe
Hamburger Rechentest (HaReT)	Primarstufe
Hamburger Schreibprobe (HSP)	Primarstufe
Leichter Lesen und Schreiben Lernen	Primarstufe
Leistung unterstützen, Potenziale entdecken (LUPE)	Primarstufe
Lesen macht stark (BiSS)	Primarstufe
Mathe macht stark	Primarstufe
Multidimensionales Begabungs-Entwicklungs-Tool (mBET)	Primarstufe
Münsteraner Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (MÜSC)	Primarstufe



Name	Bildungsetappe
Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (ZAREKI-R)	Primarstufe
Osnabrücker Test zur Zahlenbegriffsentwicklung (OTZ)	Primarstufe
Rundgang durch Hörhausen	Primarstufe
Wortgewandt und zahlenstark	Primarstufe
Berufsorientierung aktiv (BOaktiv)	Sekundarstufe I
Kapiert.de	Sekundarstufe I
Kompetenzen ermitteln (KERMIT)	Sekundarstufe I
Kompetenzfeststellung (KomPo)	Sekundarstufe I
LEO.MV	Sekundarstufe I
Lernausgangslagenerhebung (LALE)	Sekundarstufe I
Onlinediagnosen der Verlage	Sekundarstufe I
SCHNABEL	Sekundarstufe I
Audit zur begabungsfördernden Einrichtung	Stufenübergreifend
BASIS-MATH G 4+5	Stufenübergreifend
Diagnostikmaterialien des diff-Projektes	Stufenübergreifend
Einschätzungsbögen für überfachliche Kompetenzen (ÜKO)	Stufenübergreifend
Hamburger Verfahren zur Analyse der Sprachentwicklung 4 – 8-jähriger (HA-VASE)	Stufenübergreifend
Kompetenzerfassung in Kindergarten und Schule (KEKS)	Stufenübergreifend
Lernen mit Rückenwind	Stufenübergreifend
Lernserver	Stufenübergreifend
Lesen – Individuelle Fertigkeiten testen (LIFT)	Stufenübergreifend
Leseraum Online (LeOn)	Stufenübergreifend
Leseverständnistest für Erst- bis Siebtklässler (ELFE II)	Stufenübergreifend
Mathematische Basiskompetenzen digital und verstehensorientiert fördern (divomath)	Stufenübergreifend
Niemanden Zurücklassen (NZL)	Stufenübergreifend
Potenzialanalyse zur begabungsfördernden Schule	Stufenübergreifend
Züricher Ressourcenmodell (ZRM)	Stufenübergreifend

Anmerkung: „Stufenübergreifend“ steht für Instrumente, die in mehr als einer Bildungsetappe eingesetzt werden.

Quelle: SWK-Länderabfrage Schule

Anhang 3 Übersicht über die in den Ländern eingesetzten Verfahren zur Evaluation von Unterricht im Rahmen interner Evaluation¹⁰²

	Verfahren/Instrument	Öffentlich zugängliche Hinweise zur Validierung	Einschätzung durch ¹⁰³	Art der Rückmeldung	Umsetzung (digital / analog)	weiterführende Informationen
BW	Unterrichtsfeedbackbogen Tiefenstrukturen, QUBE-F	Präpilotierungs-, Pilotierungs- und Validierungsstudie (vgl. weiterführende Informationen)	Selbsteinschätzung, Eltern, LK, SL, FB, Aus-/Fortbildende	Ratingbogen inkl. Notizen, anschließendes Beratungsgespräch empfohlen	analog, bzw. ausfüllbare PDF	https://www.ibbw-bw.de/unterrichtsfeedbackbogen
	BEF-Portal <ul style="list-style-type: none"> Feedback-Tool Schul-Tool 	Nutzung validierter Skalen zur Erfassung von Unterrichtsqualität aus SuS- und LK-Sicht (Jaekel, 2021)	Selbsteinschätzung, LK, SL, SuS	automatisierter Ergebnisbericht	digital	https://befragung-bw.de/
	Basismodell zur Unterrichtsbeobachtung und -bewertung an beruflichen Schulen	k. A.	SL, SchA, Aus-/Fortbildende	k. A.	analog, digital	https://bs-s.seminare-bw.de/Lde/Startseite/Fortbildung/
BY	<ul style="list-style-type: none"> Selbsteinschätzungsbögen Befragungstool BETSI (unterschiedliche inhaltliche Bausteine) Kollegiale Hospitation: Unterrichtsbeobachtungsbögen (UBB) zu verschiedenen Beobachtungsschwerpunkten 	k. A. Validierungsstudie (GEBF-Vortrag Linberg & Taut, 2022)	LK, SuS, z. T. Eltern, päd. Personal, Kooperationspartner			https://www.isb.bayern.de/schulqualitaet/qualitaetssicherung/evaluation-konzeption-und-methoden/interne-evaluation/ https://betsie.isbbayern.de/de_DE/start/index.html

¹⁰² Die Tabelle wurde vom Institut für Bildungsanalysen Baden-Württemberg (IBBW) zur Verfügung gestellt. Grundlage sind Recherchen öffentlich zugänglicher Informationen über die Internetseiten der Landesinstitute. Die Angaben wurden im Zuge der Qualitätssicherung in der Geschäftsstelle der SWK überprüft. Eine Vollständigkeit der aufgeführten Instrumente und Verfahren kann nicht garantiert werden. Alle Links in der Tabelle aufgerufen am 23.02.2026.

¹⁰³ LK: Lehrkräfte, SL: Schulleitung, FB: Fachberatung, SuS: Schüler:innen, SchA: Schulaufsicht.

	Verfahren/Instrument	Öffentlich zugängliche Hinweise zur Validierung	Einschätzung durch ²⁰³	Art der Rückmeldung	Umsetzung (digital / analog)	weiterführende Informationen
BE/BB	Umfangreiche Toolbox: <ul style="list-style-type: none"> • Feed-UB (Gärtner et al., 2021) • EMU (Helmke et al., 2011) • KODEK (Thiel et al., 2013) • SINUS-Skala zur Wahrnehmung der Unterrichtsbedingungen (Seidel et al., 2003) • wahrgenommener Nutzen von Partner und Gruppenarbeit (Rakoczy et al., 2005) • Selbstevaluationsportal SEP-Klassik 	Zu Feed-UB: Validierungs- und Interventionsstudien durchgeführt und publiziert (Gärtner et al., 2021; Gärtner et al., 2022; Kellermann et al., 2022)	Selbsteinschätzung, SuS, LK	k. A.	k. A.	https://www.isq.berlin/wordpress/werkzeuge/toolbox/verfahren-zur-beurteilung-von-unterrichtsqualitaet/
HB	Verweis auf IQES Online	k. A.	k. A.	k. A.	k. A.	https://www.iqesonline.net/
HH	Befragungsportale <ul style="list-style-type: none"> • BEP.HH • SEP • edkimo 	k. A.	LK, SL, SuS, Eltern, außerschulische Partner	Ergebnisberichte	digital	https://ifbq.hamburg.de/schulinterne-evaluation/unsere-befragungsportale/
HE	Befragungsportal Feedback <ul style="list-style-type: none"> • Schülerfeedback: Befragungsbau- steine zur allgemeinen und fachspezifischen Unterrichtsqualität, darunter „7C RLP“ (Länderkooperation) • Schulleitungsfeedback: Befragungsbau- steine zu verschiedenen Themen des SL-Handelns 	k. A.	SuS, LK	Automatisiert erstellter Ergebnisbericht	digital	https://feedback.befragungsportal-hessen.de/de_DE/start/index.html
MV	Unterrichtsbeobachtungen	k. A.	k. A.	k. A.		https://www.bildung-mv.de/schule/schulentwicklung/daten-und-evaluation/evaluation/
	Befragungen via SEP-Klasse (Tool): <ul style="list-style-type: none"> • Schülerfeedback: Fragebogenbau- steine zu fachunspezifischen und fachspezifischen Unterrichtsaspekten, darunter u. a. Fragebogen „7C RLP“ • LK-/ Kollegium-Feedback: Bausteine „Überprüfung der Unterrichtsqualität“ & „Unterstützung kollegialer Unterrichtsentwicklung“ 	k. A.	LK, SuS	Ergebnisbericht	digital	https://www.bildung-mv.de/schule/schulentwicklung/daten-und-evaluation/evaluation/ https://klasse.sep-mv.de/de_DE/start/index.html

	Verfahren/Instrument	Öffentlich zugängliche Hinweise zur Validierung	Einschätzung durch ¹⁰³	Art der Rückmeldung	Umsetzung (digital / analog)	weiterführende Informationen
NI	Mehrperspektivische Unterrichtsbeobachtung (mUB)	k. A.	Selbsteinschätzung, LK, SuS	Ergebnisdarstellung als Diagramm, Gegenüberstellung der Perspektiven	k. A.	https://bildungsportal-niedersachsen.de/evaluation/interne-evaluation/mehrperspektivische-unterrichtsbeobachtung
	Interne Evaluation an berufsbildenden Schulen <ul style="list-style-type: none"> Selbstbewertung in berufsbildenden Schulen (SebeiSch-BBS) Handlungsorientierte Unterrichtsbeobachtung (UB-BBS) schulweite Lehrkräftebefragung (LeBe-BBS) schulweite Befragung der Schülerinnen und Schüler (SchüBe-BBS) 	Siehe Fußnote ¹⁰⁴	LK, SuS		digital (Evaluationsportal des NLO)	https://bildungsportal-niedersachsen.de/evaluation/interne-evaluation/portal-interne-evaluation-bbs
	Niedersächsisches Feedbackportal <ul style="list-style-type: none"> SuS-Befragung (Schul-/Unterrichtsqualität, selbstgesteuertes Lernen, motivational-affektives Erleben, überfachliche Kompetenzen) Self-Assessment für Lehrkräfte (Planungs-, Analyse- und Fortbildungskompetenzen) 	Siehe Fußnote ¹⁰⁵	Selbsteinschätzung, LK, SuS	Automatisch generierter, digitaler Ergebnisbericht (als PDF und interaktive Darstellung)	digital	https://feedbackportal-ni.de/de_DE/start/index.html
NW	Fragebögen zu unterschiedlichen Dimensionen des Referenzrahmen Schulqualität	k. A.	LK, SuS, Eltern	k. A.	digital (über LimeSurvey)	https://static.qua-lis.de/schulentwicklung

¹⁰⁴ Zum Schuljahr 2025/2026 gab es eine Weiterentwicklung der Qualitätsmanagement-Instrumente: Die angepassten QM-Instrumente-BBS wurden in Zusammenarbeit mit mehreren Schulen unter wissenschaftlicher Begleitung pilotiert und evaluiert, um eine möglichst praxisnahe und zielführende Anwendung sicherzustellen. Quelle: <https://wordpress.nibis.de/portal-interne-evaluation-bbs/2025/07/30/weiterentwickelte-qm-instrumente-ab-schuljahr-2025-26/> (aufgerufen am 23.02.2026).

¹⁰⁵ Sämtliche Fragebögen basieren auf maßgeblichen Überblicksbeiträgen, Metaanalysen und Primärstudien der empirischen Lehr-Lern- und Unterrichtsforschung. Bei den jeweiligen Self-Assessments handelt es sich um peer-reviewte Instrumente, die am ehemaligen Arbeitsbereich Schulpädagogik und Allgemeine Didaktik der Universität Vechta (heute Universität Siegen) validiert wurden. Im Zuge der Weiterentwicklung des Feedbackportals werden die einzelnen Skalen bzw. Assessments fortlaufend durch das Niedersächsische Landesinstitut für schulische Qualitätsentwicklung (NLO) psychometrisch geprüft. Quelle: https://feedbackportal-ni.de/de_DE/start/faq.html#Wie-kommen-die-Fragen-bzw-Aufgaben-der-Fragebogen-bzw-Self-Assessments-zustande-Aus-welchen-wissenschaftlichen-Studien-sind-sie-entnommen (aufgerufen am 23.02.2026).

	Verfahren/Instrument	Öffentlich zugängliche Hinweise zur Validierung	Einschätzung durch ²⁰³	Art der Rückmeldung	Umsetzung (digital / analog)	weiterführende Informationen
NW	Selbstreflexionsbögen zum Referenzrahmen Schulqualität	k. A.	Selbsteinschätzung, SL, SuS, Eltern			https://static.qua-lis.de/schulentwicklung/static_e/e/schulinternevaluation/fokus-schulentwicklung/selbstreflexion/index.html
	Hinweis auf weitere Instrumente zur Unterrichtsbeobachtung: <ul style="list-style-type: none"> • Edkimo • EMU • Kooperation mit ISQ Berlin und Brandenburg: Selbstevaluationsportal (SEP) 	k. A.	LK, SuS	k. A.	überwiegend digital	https://static.qua-lis.de/schulentwicklung/static_e/e/schulinternevaluation/fokus-unterricht/index.html
RP	EVA Klasse <ul style="list-style-type: none"> • Modul „Unterricht“: verschiedene Fragebögen zu überfachlichen und fachlichen Unterrichtsaspekten, u. a. Fragebogen „7C RLP“ • Modul „Schulleitungshandeln“ 	k. A.	Selbsteinschätzung, LK, SL, SuS	Automatisch generierter, interaktiver Ergebnisbericht; Gegenüberstellung von Selbst- und Fremdeinschätzung	digital	https://klasse.eva-rlp.de/de_DE/start/index.html
	EVA Schule: <ul style="list-style-type: none"> • Fragebögen zu verschiedenen Themenbereichen, u. a. Fragebogen „7C RLP“ 	„Alle eingestellten Fragebögen sind wissenschaftlich fundiert bzw. praxiserprobt“; darüber hinaus keine weiteren Angaben	LK, SuS, Eltern, päd. Personal, außerschulische Partner	Automatisch generierter Ergebnisbericht; Gegenüberstellung verschiedener Perspektiven	digital	https://schule.eva-rlp.de/de_DE/start/index.html
	InES online: voreingestellte Fragebögen zu diversen Themen sowie Möglichkeit, eigene Fragebögen zu erstellen	k. A.	Selbsteinschätzung, LK, SL, SuS		digital	https://bildung.rlp.de/ines/befragungsportale/ines-online
SN	SIEVAS (Service-Portal interne Evaluation an sächsischen Schulen) <ul style="list-style-type: none"> • Fragebögen zu verschiedenen Themenbereichen • Beobachtungsbögen • Interviewleitfäden • Dokumentenanalyse 	k. A.	LK, SuS, Eltern, Ausbildungspartner	Automatisierte Auswertung	digital	https://www.sievas.de/

	Verfahren/Instrument	Öffentlich zugängliche Hinweise zur Validierung	Einschätzung durch ¹⁰³	Art der Rückmeldung	Umsetzung (digital / analog)	weiterführende Informationen
SN	SIEVAS-Unterricht: allgemeine und fachspezifische Aspekte von Unterrichtsqualität	k. A.	LK, SuS	Automatisierte Auswertung, ggf. Gegenüberstellung von Fremd- und Selbsteinschätzung	digital	https://sievas-unterricht.de/de_DE/start/index.html
ST	Schülerfeedback zum eigenen Unterricht (Feedbackbogen)	k. A.	SuS	Bei digitaler Durchführung: Auswertung durch LISA	digital oder analog	https://www.bildung-lsa.de/informationsportal/schule/schulentwicklung/qualitaetssicherung_und_entwicklung.htm
SH	Unterrichtsfeedbackbogen Tiefenstrukturen → siehe BW					https://fachportal.lern-netz.de/sh/fachanforderungen/allgemeiner-teil.html
	Schülerfeedback.SH <ul style="list-style-type: none"> Allgemeine Unterrichtsqualität Sozial-emotionales Lernen Fachunterricht Reflexion des eigenen Lernens 	Wissenschaftliche Begleitung durch Prof. Dr. Claus Buhren, Ergebnisse ausstehend	SuS	Befragungsergebnisse einsehbar im Onlinetool, Auswertungen wie Mittelwerte etc. muss LK selbst vornehmen	digital	https://fachportal.lern-netz.de/sh/themen/schuelerfeedback.html
Ländergemeinsam	EMU (Evidenzbasierte Methoden der Unterrichtsdiagnostik, Helmke et al., 2011)		Selbsteinschätzung, LK, SuS	Beobachtungsbögen, Software zur Visualisierung	digital und analog	https://www.unterrichtsdiagnostik.info/



Anhang 4 Externe Schulevaluationen der Länder¹⁰⁶

Land	Status	Links
Baden-Württemberg	Auf Abruf ¹⁰⁷	§ 114 SchG Baden-Württemberg Datengestützte Qualitätsentwicklung an Schulen (aufgerufen am 23.02.2026)
Bayern	Ausgesetzt	Externe Evaluation Qualitätssicherung Schulqualität (aufgerufen am 23.02.2026)
Berlin	Ausgesetzt	Externe Evaluation: Schulinspektion – ISQ (aufgerufen am 23.02.2026)
Brandenburg	Ausgesetzt (temporär im Schuljahr 2025/26)	Schulvisitation (aufgerufen am 17.03.2026)
Bremen	Ausgesetzt	-
Hamburg	Regelbetrieb	Zwischenbericht zur Schulqualität in Hamburg. 3. Zyklus Schulinspektion von Oktober 2019 bis Juli 2024 (aufgerufen am 23.02.2026) Hamburg – Schulinspektionsberichte (aufgerufen am 23.02.2026) Schulinspektion - IfBO (aufgerufen am 23.02.2026)
Hessen	Regelbetrieb ¹⁰⁸	Lehrkräfteakademie Hessen – Externe Evaluation selbstständiger Schulen in Hessen (aufgerufen am 23.02.2026) §§ 92, 98 Hessisches Schulgesetz (aufgerufen am 23.02.2026)
Mecklenburg-Vorpommern	Auf Abruf	§ 6 Schulqualitätsverordnung M-V (aufgerufen am 23.02.2026) Bildungsserver MV – Evaluation (aufgerufen am 23.02.2026) Bildungsserver MV – Daten und Evaluation (aufgerufen am 23.02.2026)

¹⁰⁶ Grundlage sind Recherchen öffentlich zugänglicher Informationen, die teilweise mit einzelnen Ländern rückgekoppelt wurden. Eine Vollständigkeit kann nicht garantiert werden.

¹⁰⁷ Aktuell auf Erprobung.

¹⁰⁸ Für selbstständige Schulen.



Niedersachsen	Ausgesetzt (allgemeinbildende Schulen) Verpflichtend (berufsbildende Schulen)	Externe Evaluation – Berufliche Bildung BBS Niedersachsen: Bildungsportal Niedersachsen (aufgerufen am 23.02.2026)
Nordrhein-Westfalen	Regelbetrieb	Bildungsland NRW – Qualitätsanalyse (aufgerufen am 23.02.2026) Qualitätsanalyse-Verordnung (aufgerufen am 23.02.2026)
Rheinland-Pfalz	Ausgesetzt	Bildungsserver RLP – Was ist Evaluation (aufgerufen am 23.02.2026)
Saarland	Regelbetrieb	Bildungsserver Saarland - Qualitätssicherung an allgemeinbildenden Schulen (aufgerufen am 23.02.2026)
Sachsen	Ausgesetzt	§ 3a Sächsisches Schulgesetz (aufgerufen am 23.02.2026) Externe Evaluation an Sachsens Schulen (aufgerufen am 23.02.2026)
Sachsen-Anhalt	Auf Abruf	Bildungsserver Sachsen-Anhalt – Qualitätssicherung und -entwicklung (aufgerufen am 23.02.2026)
Schleswig-Holstein	Regelbetrieb ¹⁰⁹	Schulfeedback SH (aufgerufen am 23.02.2026)
Thüringen	Regelbetrieb	MBWK – Qualitätsentwicklungsprozess an Schulen (aufgerufen am 23.02.2026) Staatliches Schulamt Nordthüringen - Qualitätsentwicklung Thüringer Schulen (aufgerufen am 23.02.2026)

¹⁰⁹ Aktuell für Schulen im Startchancen-Programm.

Anhang 5 Ziel- und Leistungsvereinbarungen in der Schulaufsicht¹¹⁰

Land	Bezeichnung	Zeitraum	Gespräche	Genutzte Daten			Verknüpfung zu Beratungen
				Externe Evaluation	Interne Evaluation	Weitere Daten	
BW	Ziel- und Leistungsvereinbarungen	Laufzeit von 1 bis 3 Jahren.	Jährliche Statusgespräche zwischen Schule und Schulaufsicht	Ja	Ja	Schülerleistung Schulstatistik weitere bildungsbezogene Daten	Fachberater:innen Schulentwicklung und Fachberater:innen Unterricht beim ZSL
	<ol style="list-style-type: none"> § 32 SchG Baden-Württemberg Ziel- und Leistungsvereinbarungen - ZSL Referenzrahmen Schulqualität Baden-Württemberg 						
BY	Ziel ¹ - bzw. Handlungsvereinbarungen ²	Ziel- und Handlungsvereinbarungen werden nach der externen Evaluation vereinbart. Idealerweise alle 5 Jahre.	-	Ja	-	-	Schulentwicklungsmoderator:innen, QmbS-Berater:innen
	<ol style="list-style-type: none"> Bayerisches Gesetz über das Erziehungs- und Unterrichtswesen (BayEUG) 240208 Broschuere ISB OA 25 Seiten final.pdf 						
BE	Kriteriengestützte Zielvereinbarungen (Schulverträge)	Jährlich	Jährliche Bilanzgespräche	Ja	-	Schülerleistung Schulstatistik Eigene Daten der Schule z.B. durch eigene Erhebungen	Unterstützungsmöglichkeiten werden mit der Schulaufsicht besprochen
	<ol style="list-style-type: none"> § 9 SchulG Berlin SenBJF - Der Schulvertrag 						
BB	Zielvereinbarungen	2 Jahre	Jährlich stattfindenden Statusgespräche	Ja	-	-	-
	<ol style="list-style-type: none"> Gesetz über die Schulen im Land Brandenburg (BbgSchulG) MBJS - Schulvisitation Verwaltungsvorschrift über die Schulvisitation im Land Brandenburg 						
HB	Zielvereinbarungen	Es soll ein Zeitraum festgelegt werden	Zielvereinbarungsgespräche	-	-	weitere bildungsbezogene Daten	-
	<ol style="list-style-type: none"> Der Senator für Kinder und Bildung - Zielvereinbarungen 						

¹¹⁰ Grundlage sind Recherchen öffentlich zugänglicher Informationen, die teilweise mit einzelnen Ländern rückgekoppelt wurden. Eine Vollständigkeit kann nicht garantiert werden. Alle Links in der Tabelle aufgerufen am 23.02.2026.

Anhänge

Land	Bezeichnung	Zeitraum	Gespräche	Genutzte Daten			Verknüpfung zu Beratungen
				Externe Evaluation	Interne Evaluation	Weitere Daten	
HH	Ziel- und Leistungsvereinbarungen [allgemein] ¹ Basis-Zielvereinbarung [berufsbildende Schulen] ²	2 Jahre ¹	-	Ja	-	Schülerleistung	-
HE	Zielvereinbarungen	-	Jährliche Schulentwicklungsgespräche	Ja	Ja	-	-
MV	Zielvereinbarungen	-		Ja	Ja	-	-
NI	Zielvereinbarungen (werden derzeit an Startchancen-Schulen eingeführt, sowie an berufsbildenden Schulen)	-	Jährliche Zielvereinbarungsgespräche	Ja	-	Schülerleistung	-
NW	Zielvereinbarungen	Bspw. 4 Jahre		Ja	-	-	Beratung durch Schulaufsichtsbehörde oder andere Einrichtungen, insbesondere Fortbildungsträger
RP	Zielvereinbarungen	Es soll ein Zeitraum festgelegt werden			Ja	Schülerleistung	Berater:innen für Schulentwicklung, für Lehr-/Lernkultur (BBS, EQuL), Enior-Expert:innen, Interne Evaluation an Schulen.

Anhänge

Land	Bezeichnung	Zeitraum	Gespräche	Genutzte Daten			Verknüpfung zu Beratungen
				Externe Evaluation	Interne Evaluation	Weitere Daten	
SL	Zielvereinbarungen (im Projekt Qualitätsverbesserungsprozess - QVP)	-	-	-	-	-	-
	<ol style="list-style-type: none"> MBK Saarland - Qualitätsverbesserungsprozess ME Saar - Qualitätsverbesserungsprozess 						
SN	Zielvereinbarungen	-	-	-	Ja	-	Beratung und Unterstützung durch Schulaufsicht
	<ol style="list-style-type: none"> SMK - Schulentwicklung 						
ST	Zielvereinbarungen	-	-	Ja	-	-	-
	<ol style="list-style-type: none"> Bildungsserver Sachsen Anhalt - Qualitätssicherung und -entwicklung 						
SH	Zielvereinbarungen im Rahmen von Datenblattgesprächen	Es soll ein Zeitraum festgelegt werden	Jährlich verbindliche Datenblattgespräche	-	-	Schülerleistung Schulstatistik Überwachungsaudit (AZAV), integrierte Ausbildungsberichterstattung, bereitgestellte Daten von den Kammern und Übergangsquoten	-
	<ol style="list-style-type: none"> Schule der Zukunft: Datengestützte Qualitätsentwicklung an Schulen in Schleswig-Holstein 						
TH	Zielvereinbarungen	Bis zu 3 Jahren	-	Ja	-	-	Die Dienstleister erstellen nach dem Schulbesuch im Rahmen der externen Evaluation individuelle Angebote in Zusammenarbeit mit der Schule, leiten an und unterstützen/begleiten somit die Schule auf ihrem Weg der Weiterentwicklung
	<ol style="list-style-type: none"> § 40b ThürSchulG Eigenverantwortliche Schule und schulische Evaluation TMBWK - Datengestützte Qualitätsentwicklung Staatliches Schulamt Nordthüringen 						



Anhang 6 Fachberatungen der Länder¹¹¹

Land	Links
Baden-Württemberg	<p>Verwaltungsvorschrift Fachberaterinnen und Fachberater BaWü (aufgerufen am 23.02.2026)</p> <p>Webseite ZSL – Fachberaterinnen und Fachberater Unterrichtsentwicklung (aufgerufen am 23.02.2026)</p>
Bayern	<p>Dienstanweisung für die Fachberatung bei den Staatlichen Schulämtern (aufgerufen am 23.02.2026)</p> <p>Bayerisches Landesamt für Schule – Fachberatungen (aufgerufen am 23.02.2026)</p>
Berlin	<p>Berliner Landesinstitut für Qualifizierung und Qualitätsentwicklung an Schulen (BLiQ) - Berlin.de (aufgerufen am 23.02.2026)</p>
Brandenburg	<p>Beratungs- und Unterstützungssystem (BUSS) Bildungsserver Berlin - Brandenburg (aufgerufen am 23.02.2026)</p> <p>Schul- und Unterrichtsentwicklung libra - Landesinstitut Brandenburg (aufgerufen am 23.02.2026)</p>
Bremen	<p>Schulqualität - Der Senator für Kinder und Bildung (aufgerufen am 23.02.2026)</p>
Hamburg	<p>§ 85 Schulgesetz Hamburg (aufgerufen am 23.02.2026)</p> <p>ReBBZ – Bildung und Beratung (aufgerufen am 23.02.2026)</p>
Hessen	<p>Hessischer Bildungsserver - Angebote und Aufgaben der Fachberatung Kulturelle Bildung (aufgerufen am 23.02.2026)</p> <p>Schulamts Hessen – Fachberatung Schulbibliothek (aufgerufen am 23.02.2026)</p>
Mecklenburg-Vorpommern	<p>Überblick Beratungs- und Unterstützungssysteme für Schulen in M-V (Stand April 2025) (aufgerufen am 23.02.2026)</p>

¹¹¹ Grundlage sind Recherchen öffentlich zugänglicher Informationen, die teilweise mit einzelnen Ländern rückgekoppelt wurden. Eine Vollständigkeit kann nicht garantiert werden.



Niedersachsen	RdErl. d. MK v. 01.08.2024 – 32-81420 – VORIS 22410 (aufgerufen am 23.02.2026)
Nordrhein-Westfalen	<p>RdErl. 10-32 Nr. 51. Fachberatung in der Schulaufsicht (aufgerufen am 23.02.2026)</p> <p>RdErl. 10-32 Nr. 51.2 - Runderlass Fachberaterinnen und Fachberater für die Fachkoordination Deutsch und Mathematik (aufgerufen am 23.02.2026)</p>
Rheinland-Pfalz	<p>Verwaltungsvorschrift - Regelungen für die Arbeit der Regionalen Fachberaterinnen und Fachberater an Gymnasien, Kollegs, Abendgymnasien, Integrierten Gesamtschulen und berufsbildenden Schulen (aufgerufen am 23.02.2026)</p> <p>Bildungsserver RLP – Homepage Regionalen Fachberaterinnen und Fachberater (aufgerufen am 23.02.2026)</p>
Saarland	§ 53 Schulgesetz Saarland (aufgerufen am 23.02.2026)
Sachsen	<p>Fachberater/innen für Begabungs- und Begabtenförderung (aufgerufen am 23.02.2026)</p> <p>Verwaltungsvorschrift des Sächsischen Staatsministeriums für Kultus über Bereiche besonderer schulischer Aufgaben an öffentlichen Schulen und die Tätigkeit von Fachberatern (aufgerufen am 23.02.2026)</p>
Sachsen-Anhalt	Landesschulamt Sachsen – Schulfachliche Beratung (aufgerufen am 23.02.2026)
Schleswig-Holstein	<p>Website IQSH - Fächer, Fachrichtungen und Fachberatung (aufgerufen am 23.02.2026)</p> <p>Website - Kreisfachberatung für Kulturelle Bildung (aufgerufen am 23.02.2026)</p>
Thüringen	Thüringer Schulportal – Unterstützungssysteme (aufgerufen am 23.02.2026)



Anhang 7 Mögliche Nutzung von Lernentwicklungsgesprächen als Ersatz und/oder in Ergänzung zum Halbjahreszeugnis¹¹²

Nötige Voraussetzungen wie z. B. vorausgehende Beschlüsse der Schulkonferenz werden in der Tabelle nicht aufgeführt. Nicht berücksichtigt sind Übergangsberatungen und anlassbezogene Gespräche bei Lern- oder Disziplinproblemen sowie Elternsprechtage o. ä.

Folgende Typisierung wurde zur Erstellung der Abbildung 10 auf S. 93 verwendet:

- A keine Angaben
- B1 Ergänzung möglich in Primarstufe, keine Angaben zur Sek I
- B2 Ergänzung möglich in Primarstufe und in der Sek I
- C1 Ersatz möglich in Primarstufe, keine Angaben zur Sek I
- C2 Ersatz möglich in Primarstufe, Ergänzung möglich in der Sek I
- C3 Ersatz möglich in Primarstufe und in der Sek I

Die Zuordnung erfolgt auf Basis expliziter Angaben, sofern keine expliziten Regelungen vorhanden sind, wurden keine Angaben (k. A.) eingetragen. Zuordnungen, die nur mit Einschränkungen gelten, sind mit Sternchen gekennzeichnet.

Stand: Februar 2026

Land	Rechtsnorm	LEG möglich als Ersatz/in Ergänzung für Halb- jahreszeugnisse		Typ
		in der Primar- stufe	in der Sekun- darstufe I	
Baden- Württemberg	LeistungsbeurteilungsVO GS, §2 (4), §3 (2)	Ersatz in Klasse 1 bis 3		C1*
	NotenbildungsVO, §4 (1)		k. A.	
Bayern	GrundschulO, §15 (7)	Ersatz in Klasse 1 bis 3; gilt auch für das Jah- reszeugnis		C2*
	Bay SchulO, §3, Anlage 1 Maßnahme Nr. 42		Ergänzung möglich in Form ei- ner Schülerbera- tungsstunde	
Berlin	GrundschulVO, §3 (8), §19 (3)	Ersatz in Klasse 3 und 4; alle Stufen in Ge- meinschaftsschulen und Integrierten Se- kundarschulen		C3*

¹¹² Grundlage sind Recherchen in den öffentlich zugänglichen Rechtsportalen der Länder (s. Links). Eine Vollständigkeit kann nicht garantiert werden. Alle Links in der Tabelle aufgerufen am 23.02.2026.



Land	Rechtsnorm	LEG möglich als Ersatz/in Ergänzung für Halb- jahreszeugnisse		Typ
		in der Primar- stufe	in der Sekun- darstufe I	
		Ergänzung verpflichtendes lern- prozessbegeleitendes Gespräch in je- dem Schulhalbjahr		
	Schulgesetz Berlin, §58 (4)		Ersatz in Integrierten Se- kundarschulen und Gemeinschaftsschu- len in Klasse 3 bis 9	
Brandenburg	GrundschulVO, §11 (3)	Ersatz in Klasse 1 und 2		C1*
	Brandenburgisches Schulgesetz, §57, §58; Sek-I-VO §14		k. A.	
Bremen	ZeugnisVO, §18; GrundschulVO im Land Bremen, §11	Ergänzung nur zum Zeugnis am Schuljahresende; keine Halbjahres- zeugnisse in Klasse 1 bis 3, verpflichtend zweimal im Schuljahr gemeinsame Verein- barungen zu nächs- ten Lernschritten		B1
	ZeugnisVO, §19		k. A.	
Hamburg	Hamburgisches Schulgesetz, §§44 (2) (3) (4); APO-GrundStGy §7	Ergänzung	Ergänzung	B2
Hessen	VO zur Ausgestaltung der Bildungsgänge und Schulformen der Grundstufe (Primar- stufe) und der Mittelstufe (Sekundarstufe I) und der Abschlussprüfungen in der Mittel- stufe, §14; Verordnung zur Gestaltung des Schulverhältnisses § 60; Hessisches Schul- gesetz § 74	Ergänzung zum Zeugnis am Schuljahresende in Klasse 1	k. A.	B1*
Mecklenburg- Vorpommern	Schulgesetz M-V § 63; Verordnung zur ein- heitlichen Leistungsbewertung an den Schulen des Landes Mecklenburg-Vorpom- mern § 3	k. A.	k. A.	A
Niedersachsen	Zeugnisse in den allgemeinbildenden Schulen, Abschnitt 5	k. A.	k. A.	A
Nordrhein- Westfalen	Ausbildungsordnung Grundschule § 6 Ausbildungs- und Prüfungsordnung Sekun- darstufe I §7 Schulgesetz NRW § 49	k. A.	k. A.	A



Land	Rechtsnorm	LEG möglich als Ersatz/in Ergänzung für Halb- jahreszeugnisse		Typ
		in der Primar- stufe	in der Sekun- darstufe I	
Rheinland- Pfalz	SchulO für die öffentlichen Grundschulen, §§39	Ersatz in Klasse 2, hier wird noch kein Halbjah- reszeugnis ausge- stellt		B2*
	SchulO für die öffentlichen Realschulen plus, Integrierten Ge- samtschulen, Gymnasien, Kollegs und Abendgymnasien (Übergreifende SchulO), §59; ÜSchO, §8	Ergänzung in Klasse 3 und 4	Ergänzung in der Integrierten Gesamtschule und der Realschule plus	
Saarland	Zeugnis- und Versetzungsordnung - Schul- ordnung - für die Grundschulen im Saar- land (ZVO-GS), § 3	Ergänzung in Klasse 1 und 2		B1*
	Verordnung - Schulordnung - über den Bildungsgang und die Abschlüsse der Sekundarschule (Sek-VO), § 9, § 10; Zeugnis- und Versetzungsordnung - Schulordnung - für die Klassenstu- fen 5 bis 10 des neunjährigen Gymnasiums, § 3, § 4		k. A.	
Sachsen	Schulordnung Grundschulen (SOGS), § 22, § 23; Schulordnung Gemeinschaftsschulen (SOGES), § 33, § 34; Schulordnung Gymna- sien Abiturprüfung (SOGYA), § 32	k. A.	k. A.	A
Sachsen- Anhalt	Leistungsbewertung und Beurteilung an der Grundschule und im Primarbereich an Förderschulen, Kap. 2, Kap. 6;	Ergänzung		B1
	Leistungsbewertung und Beurteilung an allgemeinbildenden Schulen und Schulen des Zweiten Bildungsweges der Sekun- darstufen I und II, Kap. 6, Kap. 8		k. A.	
Schleswig- Holstein	Landesverordnung über Grundschulen (GrVO), §6	Ersatz in Klasse 1		C1*
	Zeugnisverordnung (ZVO), §3; Schulgesetz (SchulG), § 16		k. A.	
Thüringen	Thüringer SchulO, §§59a, 60	Ergänzung verpflichtend in Klasse 3 oder 4 ein- mal im Schuljahr; in weiteren Klassen- stufen möglich	Ergänzung verpflichtend in Klasse 5 oder 6 sowie 7 oder 8 einmal im Schuljahr; in weite- ren Klassenstufen möglich	B2



Abbildungsverzeichnis

Abbildung 1: Zentrale Lernstandserhebungen in den Ländern	16
Abbildung 2: VERA-Rückmeldungen an Lehrkräfte.....	20
Abbildung 3: Zentrale Lernstandserhebungen (vgl. Empfehlung 4).....	41
Abbildung 4: VERA-Rückmeldungen an Schulleitungen.....	49
Abbildung 5: Angebote zur Fortbildung und Qualifikation zur datengestützten Schulentwicklung	52
Abbildung 6: VERA-Rückmeldungen an die Schulaufsicht.....	67
Abbildung 7: Externe Evaluation, Schulinspektion	70
Abbildung 8: Datenrückmeldung an Eltern bei VERA 3 und 8 inkl. genutzter Formate	87
Abbildung 9: Umfang der Rückmeldungen an Eltern in den Ländern relativ zur genutzten Anzahl der verfügbaren Instrumente	88
Abbildung 10: Rechtlich geregelte Möglichkeiten zur Nutzung von Lernentwicklungsgesprächen als Alternative zum Halbjahreszeugnis.....	93
Abbildung 11: Zentrale Beschlüsse der KMK zum Bildungsmonitoring und erste Durchführung der Studien	107
Abbildung 12: Zeitliche Taktung der Maßnahmen zur Feststellung der Leistungsfähigkeit des Bildungswesens	110
Abbildung 13: Regelungen der Länder zu internen und externen Evaluationen der pädagogischen Qualität der Kindertageseinrichtungen	136
Abbildung 14: Sprachstandserhebung im Vorschulalter in den Ländern	138

Bildnachweis

Abbildung 7, 13: "Germany Laender Sachsen-Anhalt" ©BMacZero, lizenziert unter CC BY-SA 4.0. Einfärbungen der Länder hinzugefügt. Original: https://commons.wikimedia.org/wiki/File:Germany_Laender_Sachsen-Anhalt.svg. Die modifizierte Version steht ebenfalls unter CC BY-SA 4.0. Weitere Informationen zur Lizenz: <https://creativecommons.org/licenses/by-sa/4.0/>.



Mitglieder der SWK

Prof. Dr. Felicitas Thiel

(Co-Vorsitzende der SWK und Vorsitzende der Redaktionsgruppe für dieses Gutachten)

Professorin für Schulpädagogik und Schulentwicklungsforschung
Freie Universität Berlin

Prof. Dr. Olaf Köller

(Co-Vorsitzender der SWK und Redaktionsgruppe für dieses Gutachten)

Geschäftsführender wissenschaftlicher Direktor des IPN
IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Prof. Dr. Isabell van Ackeren-Mindl (Redaktionsgruppe für dieses Gutachten)

Professorin für Bildungssystem- und Schulentwicklungsforschung
Universität Duisburg-Essen

Prof. Dr. Yvonne Anders (Redaktionsgruppe für dieses Gutachten)

Abteilungsleiterin
Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF)

Prof. Dr. Ulrike Cress

Direktorin des IWM
Leibniz-Institut für Wissensmedien (IWM), Tübingen

Prof. Dr. Claudia Diehl

Professorin für Mikrosoziologie
Universität Konstanz

Prof. Dr. Thilo Kleickmann

Leitung der Abteilung Schulpädagogik
Christian-Albrechts-Universität zu Kiel

Prof. Dr. Michael Krelle (Redaktionsgruppe für dieses Gutachten)

Professor für Fachdidaktik Deutsch
Technische Universität Chemnitz

Prof. Dr. Birgit Lütje-Klose

Professorin für Schulische Inklusion und sonderpädagogische Professionalität
Universität Bielefeld

Prof. Dr. Susanne Prediger

Professorin für Mathematikdidaktik und fachbezogene Professionalisierungsforschung
Technische Universität Dortmund; IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik



Prof. Dr. Susan Seeber

Professorin für Wirtschaftspädagogik und Personalentwicklung
Georg-August-Universität Göttingen

Prof. Dr. Birgit Ziegler

Professorin für Berufspädagogik und Berufsbildungsforschung
Technische Universität Darmstadt

Ständige Mitglieder

Prof. Dr. Samuel Greiff

Vorstandsvorsitzender Zentrum für internationale Vergleichsstudien (ZIB)
Technische Universität München

Prof. Dr. Kai Maaz

Sprecher Autorengruppe Bildungsberichterstattung
DIPF – Leibniz-Institut für Bildungsforschung und Bildungsinformation

Prof. Dr. Christian Reintjes (Mitglied bis 12/2025)

Vorsitzender wissenschaftlicher Beirat der Steuerungsgruppe „Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich“
Universität Osnabrück

Prof. Dr. Kerstin Schneider (Mitglied seit 01/2026)

Vorsitzende wissenschaftlicher Beirat der Steuerungsgruppe „Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich“
Bergische Universität Wuppertal, RWI – Leibniz-Institut für Wirtschaftsforschung

Prof. Dr. Petra Stanat (Redaktionsgruppe für dieses Gutachten)

Wissenschaftlicher Vorstand Institut zur Qualitätsentwicklung im Bildungswesen (IQB)
Humboldt-Universität Berlin



Mitwirkende

Unter Mitarbeit von

Prof. Dr. Dr. h.c. Manfred Prenzel (ehem. Universität Wien und TU München), Prof. Dr. Dr. Christiane Spiel (ehem. Universität Wien)

Mitarbeiter:innen aus der Geschäftsstelle der SWK

Dr. Per Tiedtke, Dr. Christiane Schöneberger, Dr. Karin Guill, Dr. Judith Schulte, Dr. Julia Suckut, Jennifer Adomat (bis 08/25), Carlotta Brüning (ab 12/2025)

Dank

Besonderer Dank gilt folgenden Expert:innen für ihre wertvollen Hinweise und Einschätzungen:

Prof. Dr. Hermann Josef Abs (Universität Duisburg-Essen), Dr. Andrea Albers (IfBQ Hamburg), Prof. Dr. Timm Albers (Universität Paderborn), Dr. Elke Alsago (ver.di), Anja Bensinger-Stolze (Gewerkschaft Erziehung und Wissenschaft), Kathrin Bock-Famulla (Bertelsmann-Stiftung), Prof. Dr. Oliver Böhm-Kasper (Universität Bielefeld), Dr. Desirée Burba (Ministerium für Allgemeine und Berufliche Bildung, Wissenschaft, Forschung und Kultur Schleswig-Holstein), Jacob Chammon (Telekom Stiftung), Prof. Dr. Peter Cloos (Studiengangstag Pädagogik der Kindheit), Dr. Martina Diedrich (DIPF), Prof. Dr. Charlotte Dignath (Universität Frankfurt), Prof. Dr. Hendrik Drachsler (DIPF), Stefan Düll (Deutscher Lehrerverband), Dr. Angela Ehlers (Verband Sonderpädagogik), Gabriele von Eichhorn (in Vertretung des IQS Österreich), Prof. Dr. Birgit Eickelmann (Universität Paderborn), Dr. Tobias Ernst (Stiftung Kinder forschen), Niels Espenhorst (Arbeitsgemeinschaft für Kinder- und Jugendhilfe), Prof. Dr. Benjamin Fauth (PH Freiburg), Martina Fiddrich (Cornelsen Verlag), Prof. Dr. Christian Fischer (Universität Münster), Dr. Matthias Forell (Universität Bochum), Michael Fritz (IZB Initiative für Zukunftsbildung), Prof. Dr. Holger Gärtner (FU Berlin), Dr. Sabine Geist (Laborschule Bielefeld), Prof. Dr. Frank Goldhammer (DIPF/Universität Frankfurt a.M.), Prof. Dr. Richard Göllner (Universität Potsdam), Marion Gutzmann (Grundschulverband e.V.), Dr. Anja Hagen (EdTech-Verband e.V.), Prof. Dr. Andreas Hartinger (Universität Augsburg), Prof. Dr. Marcus Hasselhorn (DIPF), Hendrik Haverkamp (FelloFish), Patrick Hawlitschek (IOB), Dr. Ursula Held (Institut für Qualitätsentwicklung im Land Bremen), Prof. Dr. Silke Hertel (Universität Heidelberg), Dirk Heyartz (Bundeselternrat), Dr. Andreas Jantowski (Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien), Iris Kalvelage (Westermann Verlag), Prof. Dr. Karina Karst (Technische Universität Kaiserslautern-Landau), Prof. Dr. Ina Kaul (Deutsche Gesellschaft für Erziehungswissenschaft), Matthias Graf von Kielmansegg (Vodafone Stiftung), Prof. Dr. Dagmar Killus (Universität Hamburg), Amy Kirchhoff (Bundesschülerkonferenz), Prof. Dr. Dominique Klein (TU Dortmund), Mag. Dr. Robert Klinglmair (in Vertretung des IQS Österreich), Prof. Dr. Katharina Kluczniok (pädquis/Freie Universität Berlin), Prof. Dr. Uta Klusmann (HU Berlin), Dr. Susanne Kollmann (IQHB Bremen), Dr. Katharina Kopp (Forschungsverbund DJI/TU Dortmund), Prof. Dr. Friederike Korneck



(Gesellschaft für Fachdidaktik), MD Alexander Kraft (Ministerium für Bildung, Wissenschaft und Kultur Schleswig-Holstein), Prof. Dr. Annelies Kreis (PH Luzern), Kerstin Krins (Rothenburg-Grundschule Berlin), Prof. Dr. Susanne Kuger (Deutsches Jugendinstitut/Universität München), Arndt Kwiatkowski (bettermarks GmbH), Prof. Dr. Rebecca Lazarides (Gesellschaft für Empirische Bildungsforschung), Kerstin Lehner (Wübben Stiftung Bildung), Dr. Christina Matschke (IWM Tübingen), Prof. Dr. Nele McElvany (TU Dortmund), Dr. Christiane Meiner-Teubner (Forschungsverbund DJI/TU Dortmund), Prof. Dr. Conny Melzer (Universität Leipzig), Tino Melzer (Thüringer Landesbeauftragter für den Datenschutz und die Informationsfreiheit), Prof. Dr. Detmar Meurers (IWM Tübingen), Prof. Dr. Sonja Nonte (Universität Osnabrück), Christoph Pienkoss (Verband Bildungsmedien e.V.), Dr. Birgit Pikowsky (Pädagogisches Landesinstitut Rheinland-Pfalz), Dr. Britta Pohlmann (IfBQ Hamburg), Katharina Queisser (Bundeselternvertretung der Kinder in Kindertageseinrichtungen und Kindertagespflege), Dr. Ulrike Rangel (IBBW Baden-Württemberg), Prof. Dr. Dirk Richter (Universität Potsdam), Anne Rolvering (Deutsche Kinder- und Jugendstiftung), Prof. Dr. Michael Schemmann (Deutsche Gesellschaft für Erziehungswissenschaft), Oliver Schmidt (stellvertretend für Simone Fleischmann, dbb beamtenbund und tarifunion), Jeanette Schnell (Deutscher Gewerkschaftsbund), Prof. Dr. Knut Schwippert (Universität Hamburg), Gerlind Sell (Konferenz der Schulaufsicht in der Bundesrepublik Deutschland), Matthias Siebert (Berufsverband Deutscher Psychologinnen und Psychologen, Sektion Schulpsychologie), Anika Smits (Verband für Kitafachkräfte NRW e.V.), Prof. Dr. Elmar Souvignier (Universität Münster), Prof. Dr. Karsten Speck (Universität Oldenburg), Anette Stein (Bertelsmann-Stiftung), Elisa Steinfeldt (Fröbel e.V.), Prof. Dr. Torsten Steinhoff (Universität Siegen), Vincent Steinl (Robert Bosch-Stiftung), Prof. Dr. Petra Strehmel (Bundesarbeitsgemeinschaft Bildung und Erziehung in der Kindheit e.V.), Prof. Dr. Ulrich Trautwein (Universität Tübingen), Mathias Tuffentsammer (Kitalino GmbH), Dr. Sonja Ugen (Universität Luxemburg), Waltraud Weegmann (Deutscher Kitaverband), Prof. Dr. Elke Wild (Universität Bielefeld), Sven Winkler (Allgemeiner Schulleitungsverband Deutschlands), Dr. Claudia Wirts (Staatsinstitut für Frühpädagogik und Medienkompetenz Bayern), Dr. Rachel Wollschläger (Universität Luxemburg), Alexander Wroblewski (Seewiesenschule Esslingen), Dr. York Xylander (bettermarks GmbH), Dr. Dirk Zorn (Bertelsmann-Stiftung).

Vertreter:innen der folgenden Ministerien und Gremien haben als Beobachter:innen die Hearings begleitet:

Bildungsministerkonferenz (BMK), Jugend- und Familienministerkonferenz (JFMK), Kommission Qualitätsentwicklung und Bildungsmonitoring (KomQB), Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend (BMBFSFJ).



Impressum

Das vorliegende Gutachten wurde von den Mitgliedern der Ständigen Wissenschaftlichen Kommission im März 2026 verabschiedet.

Gutachten bearbeiten eine umfangreiche Fragestellung unter verschiedenen disziplinären Perspektiven und enthalten Empfehlungen für längerfristige Maßnahmen und strategische Entwicklungen. Sie werden in Arbeitsgruppen unter Einbezug externer wissenschaftlicher Expertise erarbeitet. Weitere Informationen zu den Formaten der SWK finden Sie [hier](#).

Während der redaktionellen Überarbeitung dieses Gutachtens wurden verschiedene KI-Modelle verwendet, um die Sprache und die Lesbarkeit zu verbessern. Nach der Verwendung dieser Werkzeuge haben die Autor:innen den Inhalt überarbeitet und übernehmen die volle Verantwortung für den Inhalt.

Herausgeber: Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (SWK)

Geschäftsstelle

Graurheindorfer Str. 157
53117 Bonn

Tel. 0228 501 700
info@swk.kmk.org
www.swk-bildung.org

Verantwortlich: Dr. Isabelle Sieh (Leitung der Geschäftsstelle)

Die Veröffentlichungen der SWK stehen auf den Seiten der SWK zum [Download](#) zur Verfügung.




Zitationsvorschlag:

Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (SWK) (2026): *Datengestützte Entwicklung und Steuerung in Schulen und frühkindlicher Bildung*. Gutachten der der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz. <https://doi.org/10.25656/01:35028>

© Geschäftsstelle der SWK, Bonn 2026

Lizenz: CC BY-SA 4.0. Weitere Informationen zur Lizenz: <https://creativecommons.org/licenses/by-sa/4.0/>





Ständige Wissenschaftliche Kommission (SWK) ist ein unabhängiges wissenschaftliches Beratungsgremium der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK).

www.swk-bildung.org

