

Norbert Maritzen

Evidenzbasierte Unterrichtsentwicklung – Über einige Vereinfachungen und Fehlvorstellungen

Summary: *Mit nationalen und internationalen Leistungsvergleichsuntersuchungen ist in den deutschsprachigen Ländern auch die Frage virulent geworden, was mit den Ergebnissen auf der Ebene des Unterrichts eigentlich anzufangen sei. Bildungspolitik und -administration erwarten jedenfalls von Lehrkräften und Schulleitungen, dass sie Konsequenzen aus solchen Ergebnissen für die pädagogische Praxis ziehen. Dass solche Erwartungen bisher eher enttäuscht wurden, hängt nicht zuletzt mit verengten Vorstellungen von Nutzungsvoraussetzungen und -prozessen zusammen. Dem versucht der Beitrag nachzugehen mit dem Ziel, ein differenziertes Verständnis von „evidenzbasierter Unterrichtsentwicklung“ zu skizzieren.*

Einleitung

Der sukzessive Auf- und Ausbau von Systemen des Bildungsmonitorings, der in den deutschsprachigen Ländern im internationalen Vergleich ein wenig verspätet einsetzte, hat für Lehrende an Schulen zur Folge, dass sie mit der Erwartung konfrontiert sind, extern zur Verfügung gestellte Daten auch systematisch zu nutzen. Der bildungspolitische Kontext und die Modalitäten der Implementierung von Verfahren und Instrumenten des Bildungsmonitorings mögen in den Ländern durchaus unterschiedlich sein. Gemeinsam ist ihnen, dass unter Rekurs auf meist vage Konzepte von „Daten“ oder „Evidenz“ die mehr oder weniger verbindliche Aufforderung an Lehrende und Schulleitungen ergeht, die eigene Praxis auf Basis von Daten zu reflektieren und weiterzuentwickeln, um so das pädagogische Geschäft gleichsam aus der empirischen Ahnungslosigkeit herauszuführen. Man wäre fast geneigt, dieser Entwicklung einen aufklärerischen Impetus zu unterstellen frei nach Kants „Ausgang des Menschen aus der selbstverschuldeten Unmündigkeit“ (Kant 2015), wenn nur die Evidenz-Forderung nicht eindeutig in einem Kontext stünde, der die pädagogische Profession in einen Zwang zur Rechenschaftslegung einbindet. Empirisch fundiert ist beides, Entwicklung des Lernens und Lehrens einerseits und Ausweitung von staatlichen Kontrollansprüchen andererseits, von vornherein eng miteinander verknüpft. Das ist nun keineswegs beklagenswert, sondern historisch gesehen in der Gleichzeitigkeit von Freiheit und Zwang geradezu die Signatur des modernen staatlichen Erziehungswesens. Schwieriger ist indes, dass mit der Erwartung, Schul- und Unterrichtsentwicklung empirisch zu fundieren, im angedeuteten Systemreformkontext zugleich die Eigengesetzlichkeit des pädagogischen Handlungskontextes weitgehend ausgeblendet wird und kurzschlüssige Ableitungsmythen gepflegt werden. Als ergäben sich pädagogisch-didaktische Interventionen direkt aus Daten. Der Beitrag versucht, dieser Problematik nachzugehen, indem zunächst einige verhaltensrelevante Herausforderungen skizziert werden, die die Nutzung datengestützter Informationen mit sich bringt. Anschließend fokussieren wir auf die Vielschichtigkeit der Datennutzungsbedingungen im schulischen Kontext. Im dritten Abschnitt

soll verdeutlicht werden, dass evidenzbasierte Unterrichtsentwicklung auf höchst voraussetzungsvolle Infrastrukturen der Datenbereitstellung angewiesen ist. Am Schluss wird dann resümierend auf einige gängige Vereinfachungen und Fehlvorstellungen verwiesen, die es zu vermeiden gilt, wenn Evidenzbasierung für die Weiterentwicklung unterrichtlicher Praxis fruchtbar gemacht werden soll.¹

Was passiert, wenn Evidenz auf unterschiedliche Akteure trifft?

Was passiert, wenn in bestimmten Verfahren gewonnene Informationen im Schulbereich an handelnde Akteure rückgemeldet werden in der Erwartung, dass diese dann etwas mit den Informationen anfangen, um Zustände zu verbessern? Das ist nie ganz vorhersehbar und beherrschbar. Die erforderlichen Prozesse der Informationsverarbeitung sind vielfältig und komplex. Informationen werden in bestehende kognitive Strukturen integriert. Jenseits kognitiver Prozesse spielen auch Emotionen und Haltungen eine zentrale Rolle: Zuschreibungsmuster, Angst vor Veränderung, Angst vor dem Scheitern, Abwehrreaktionen etc., ein ganzes Bündel widersprüchlicher Gefühle kann einer gezielten Datennutzung im Wege stehen (vgl. *Diedrich et al. 2020*). Das ist normal. Insbesondere die emotionalen Prozesse dürften mit dazu beitragen, dass Evidenz keinesfalls in der Schrittfolge von Rezeption, Reflektion, Aktion und Evaluation verarbeitet wird, wie beispielsweise das vielfach adaptierte Modell von *Helmke* (vgl. *Helmke 2012*) suggeriert. Datennutzung folgt keinen rein rationalen Informationsverarbeitungsprozessen. Wer dies in der Implementierung von Datenfeedbackstrukturen unterstellt, wird unweigerlich enttäuscht werden. Wenn wohlmeinende Evidenz-Propagandisten, insbesondere solche mit hierarchischer Positionsmacht, gern Ignoranz oder Reaktanz unwilliger Empirie-Banausen für die träge Wirksamkeit ihrer Evidenzprogrammatik verantwortlich machen, verkennen sie die Tatsache, dass die frohe Botschaft der Evidenzbasierung nicht so einfach verfangen kann. Sie verkennen auch, dass das, was im Umgang mit Datenfeedbacks auf allen Ebenen als Vermeidungsstrategien sichtbar zu werden scheint, bei näherer Betrachtung aus der jeweiligen Innensicht durchaus eine erhebliche Rationalität aufweist, der gründlicher nachzugehen sich lohnen würde.

Tab. 1. Was passiert, wenn „Evidenz“ auf unterschiedliche Akteure trifft?

Akteur	Was häufig im Zentrum steht	Was im Zentrum stehen sollte
Einzelperson (Lehrkraft, Schulleitung)	neue Evidenz in vorhandene Sinnstrukturen so integrieren, dass Routinen und Handlungsfähigkeit erhalten werden können	eigene Annahmen und Haltungen hinterfragen, eigene Veränderungsnotwendigkeiten erkennen, Schlüsse für das eigene Handeln ziehen
Gruppen (Lehrer/innen kollegien)	sicherstellen, dass niemand beschämt oder bloßgestellt wird, Unterschiede zudecken	Schlussfolgerungen für Professionalisierung und Organisationsentwicklung ziehen, Unterschiede zwischen Kolleginnen und Kollegen sichtbar machen
Aufsichtspersonen	die eigene Überlebensfähigkeit sichern, mit operativer Überlast umgehen, Ambiguitäten ausblenden	ein Verständnis für die organisationalen Veränderungsnotwendigkeiten entwickeln und Konsequenzen für die Schulentwicklung ziehen
Personen im Unterstützungs-/ Fortbildungssystem	Dissonanzen vermeiden oder verringern, eigene Evidenzen gegen neue Evidenz setzen	Relevanzentscheidungen treffen, Evidenz mit bereits vorhandenem Wissen verknüpfen, Einschätzungen mit der Schule rückkoppeln

Evidenz trifft also zwangsläufig auf unterschiedliche Akteure mit unterschiedlichen normativen Erwartungen, Zielvorstellungen und Interessen. Sie kann nicht eo ipso gewährleisten, dass involvierte Akteure dieselben Maßnahmen der Schul- und Unterrichtsentwicklung favorisieren oder ergreifen. Evidenz stellt noch keine systemische Passung an den Schnittstellen der unterschiedlichen Akteure her. Evidenz selbst ist immer auch eine Angelegenheit von Aushandlung, Reflexion, professionellem Diskurs. Ihre Wirksamkeit im Sinne der Qualitätsentwicklung baut auf spezifische Voraussetzungen, die von den Akteuren erst in zumeist mühevollen Aushandlungsprozessen herzustellen sind. Eine nicht unwesentliche Gelingensbedingung sind in diesem Kontext Aspekte, die man etwas altmodisch unter dem Begriff „Haltung“ subsummieren könnte. Dazu gehören ganz praktische Regeln wie z. B. immer Daten aus unterschiedlichen Quellen, zu unterschiedlichen Dimensionen möglichst in längsschnittlicher Perspektive und nicht Einzelergebnisse zur Grundlage von Entscheidungen über Individuen/Institutionen zu machen oder auch die Gewichtung unterschiedlicher Datengrundlagen für Entscheidungen immer explizit zu machen. Anspruchsvoller ist schon, dass Lehrpersonen und Führungskräfte sich für ihre spezifische Aufgabe bei der Datennutzung schulen und die eigenen Aufgaben, das eigene Interventionsrepertoire, die (Neben-)Folgen des eigenen Agierens kritisch reflektieren. Zentral ist die mühevoll Einübung eines forschend-fragenden Habitus im professionellen Agieren. Er bedeutet nicht weniger als einen Abschied vom lieb gewordenen Bescheidwissen.

Was passiert, wenn eine Lehrkraft und ein Datum sich begegnen?

Im Folgenden sollen nun die konkreten Bedingungen skizziert werden, die sich aus der erweiterten Verfügbarkeit von datenbasierten lern-, lerner- und institutionenbezogenen Informationen für kooperative Handlungsstrategien von Lehrpersonen ergeben. Zugrunde gelegt wird ein mehrfach geschichtetes Modell des Lehrerhandelns, in dem sukzessive unterschiedliche Voraussetzungen, Bedingungen und Prozesse der „Datennutzung“ entfaltet werden. Mit diesem Versuch soll auch Annahmen widersprochen werden, die in simplifizierender Weise davon ausgehen, datengestützte Informationen z. B. zu Lernergebnissen sprächen für sich selbst (Daten sprechen nicht!) oder es bedürfe nur der Erläuterung von Kennwerten, Diagrammen und Tabellen durch „Rückmeldemoderatoren“, um hinreichende Voraussetzungen für die Datennutzung in der Schule zu schaffen. Solchen Fehlkonzepten erliegen nicht nur ministerielle Instanzen, die unter dem Druck der großen Zahl flächendeckende und finanzierbare Nutzungskonzepte implementieren müssen, sondern auch Fortbildungsinstitutionen oder Qualitätsagenturen, wenn sie sich mit Angeboten an einzelne Lehrpersonen wenden.

Die Grundlagen, die nötig sind, um diesen theoretischen Rahmen zu entfalten, sind von der deutschsprachigen Forschung bisher kaum hinreichend entfaltet, geschweige denn empirisch systematisch getestet worden. Sichtet man neuere Forschungsbefunde aus dem deutschsprachigen Raum zum schulischen Umgang mit Inspektionsdaten, mit Daten aus Lernstandserhebungen und aus zentralen Abschlussprüfungen, kommt man zu dem ernüchternden Schluss: Daten werden auf Schulebene offensichtlich wenig und unsystematisch für die Schulentwicklung, Unterrichtsentwicklung oder Fortbildungsplanung genutzt. Es gibt wenig Befunde zur Nutzung verschiedener Datenquellen im Zusammenhang und zu den schulexternen Kontextbedingungen einer wirksamen Nutzung. Die Befunde zu Effekten sind insgesamt widersprüchlich; Letzteres liegt nicht selten auch an unzureichenden Forschungsdesigns (Altrichter et al. 2016, Maritzen 2017).

Günstiger ist für unser Thema die wissenschaftliche Ausgangslage, wenn man die niederländische oder US-amerikanische Forschungslandschaft sichtet. Am systematischsten sind *Schildkamp* und Mitarbeiter/innen (Universität Twente, Enschede, Niederlande) und *Mandinach* und Mitarbeiter/innen (WestEd, San Francisco, USA) seit vielen Jahren in zahlreichen empirischen Studien den schulischen Voraussetzungen, Bedingungen und Wirkungen der Datennutzung von Lehrpersonen nachgegangen (vgl. exemplarisch die jüngeren Arbeiten: z. B. *Gummer* et al. 2015, *Mandinach* et al. 2016, *Mandinach* et al. 2020, *Schildkamp* et al. 2018, *Schildkamp* et al. 2019). Interessant für unseren Kontext ist, dass sie in ihren zum Teil experimentell angelegten Forschungen insbesondere auf nicht-lineare Prozessverläufe der Datennutzung, auf Unterstützungsbedarfe und auf Kooperationsstrukturen bei der Datennutzung in „data teams“, die man als Spezialfall von professionellen Lerngemeinschaften verstehen kann, eingehen.

Im Folgenden wird auf diese Arbeiten rekurriert, um ein Modell der Datennutzung durch Lehrpersonen zu entwerfen, das sowohl das Prozesshafte wie die komplexe individuelle und institutionelle Bedingungsstruktur abzubilden versucht. Was geschieht also, wenn sich ein „Datum“ und eine Lehrperson begegnen? Man darf sich den Vorgang nicht als rein rezeptiven oder simplen Ableitungsprozess vorstellen, sondern als mehrschrittigen, rekursiven Transformationsprozess, der von der Datenrezeption über die aktive Konstruktion von Information und von Wissen zum pädagogischen Entscheidungshandeln führt.

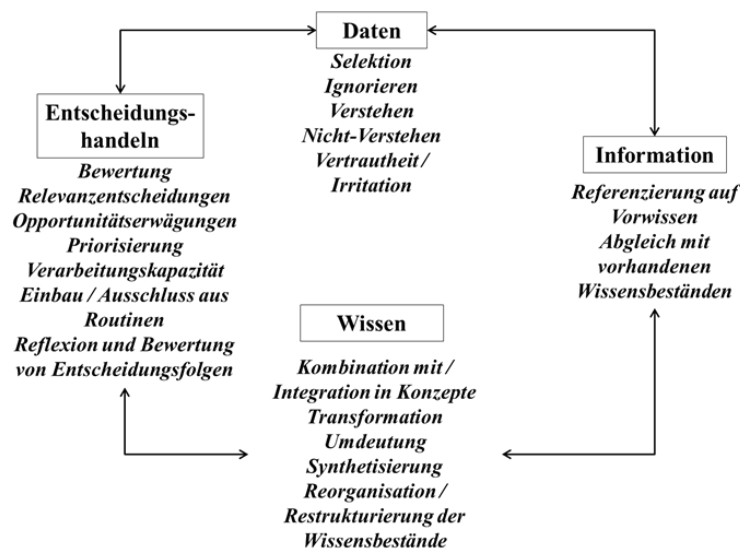


Abbildung 1: Komponenten des Zusammenhangs von Daten, Information, Wissen und Entscheidungshandeln auf individueller Ebene

Im ersten Schritt selektiert, ignoriert bewusst/unbewusst, versteht/missversteht eine Lehrkraft angebotene Daten, hat vertraute Déjà-vu-Erlebnisse oder fühlt sich irritiert. Schon die „Erstbegegnung“ mit einem Datum hat kognitive, aber auch affektive Komponenten, ein Umstand, der – obwohl handlungssteuernd – in einem Rationalitätsparadigma der Datennutzung nicht hinreichend in Rechnung gestellt wird. Kein Datum „im Rohzustand“ erklärt

sich selbst. Es wird vielmehr erst dann für eine Lehrperson zu einer bedeutungsvollen Information, wenn es anschlussfähig an ihr Vorwissen ist (z. B. zu einem fachlichen Inhalt, zu einem Schüler, zu situativen Gegebenheiten oder zu anderen Daten). Auf die erste Wahrnehmung eines Datums folgt immer ein zum Teil unbewusster Abgleich mit vorhandenen Wissensbeständen; erst dieser Referenzrahmen schafft Information. Manche Wirkungslosigkeit von Datenrückmeldungen erklärt sich vermutlich nicht durch intentionale Ignoranz oder Inkompetenz von Lehrpersonen, sondern dadurch, dass mit Datenrückmeldungen der Referenzhorizont von Lehrpersonen, der immer auch eine Relevanzzuschreibung markiert, gar nicht erreicht wird.

Wissen entsteht nun aus einer Information, wenn neue Informationen kombiniert werden mit vorhandenen Wissensbeständen, wenn es zur Integration in vorhandene Wissenskonzepte kommt. Dabei können auf der einen Seite datengestützte Informationen auch umgedeutet oder an das Vorgewusste angepasst werden. Aber auch eine partielle/ komplette Reorganisation der Wissensbestände kann die Folge sein. Mit den skizzierten Prozessen ist gleichwohl noch nicht verbürgt, dass neues Wissen auch für pädagogisches Entscheidungshandeln genutzt wird. Auf individueller Ebene setzt Letzteres immer eine Reihe von Relevanzentscheidungen, Bewertungen, Opportunitätserwägungen, Risiko- und Folgenabschätzungen und Bewertung von Entscheidungsfolgen voraus. Auch die Anschlussfähigkeit an routinierte Handlungsautomatismen oder die situative Verarbeitungskapazität der Lehrkraft – Lehrkräfte entscheiden oft unter Zeitrestriktionen und Handlungsdruck – spielen eine Rolle.

Auf Seiten der Lehrkraft ist Datennutzung ferner gesteuert von einem vielschichtigen Kompetenzbündel und mehr oder weniger stabilen Handlungsmustern, die in diesem Zusammenhang nicht alle erläutert werden können.

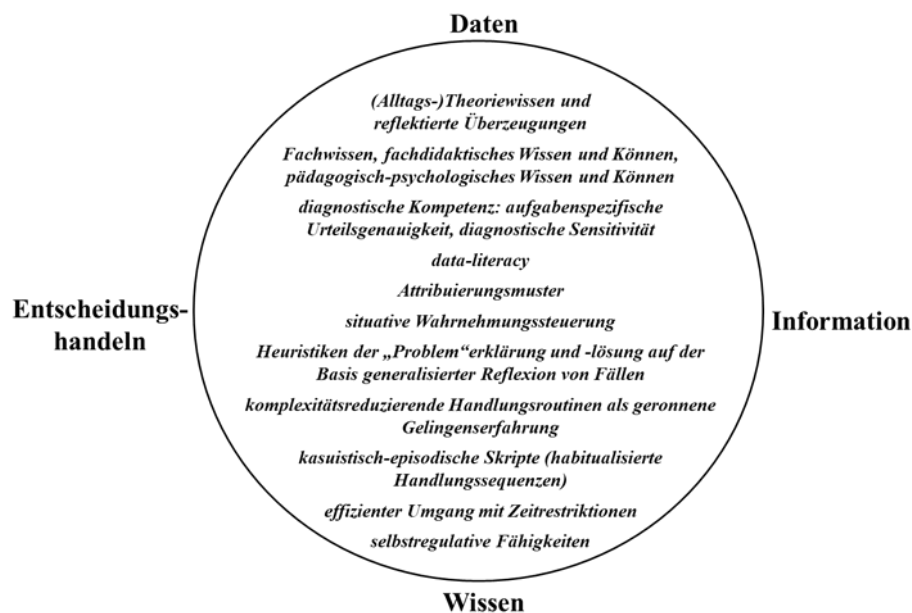


Abbildung 2: Kompetenzbündel und Handlungsmuster auf individueller Ebene

Hingewiesen sei nur auf den Teilaspekt „data-literacy“, nämlich die Fähigkeit, datengestützte Informationen in handhabbares, nutzbares Wissen zu transformieren. Erforderlich ist, Daten mit Fachwissen, didaktisch-methodischem Wissen und einem Wissen über Lernprozesse bei Schülerinnen und Schülern in Passung zu bringen (Gummer et al. 2015, Mandinach et al. 2016) und ein solches Wissen in kooperative Prozesse des „sensemaking“ einzubringen (Vanlommel et al. 2018). Dabei müssen Lehrkräfte sich der Tatsache bewusst sein, dass ihre Entscheidungen häufig high-stakes-Charakter haben, d. h. mit unmittelbaren Konsequenzen für Schüler/innen verbunden sind (Vanlommel et al. 2017).

Damit die beschriebene Wissensgenerierung gelingt und überhaupt in pädagogische Entscheidungsprozesse eingebracht werden kann, muss man sich diesen als einen mehrfach geschichteten Zusammenhang vorstellen.

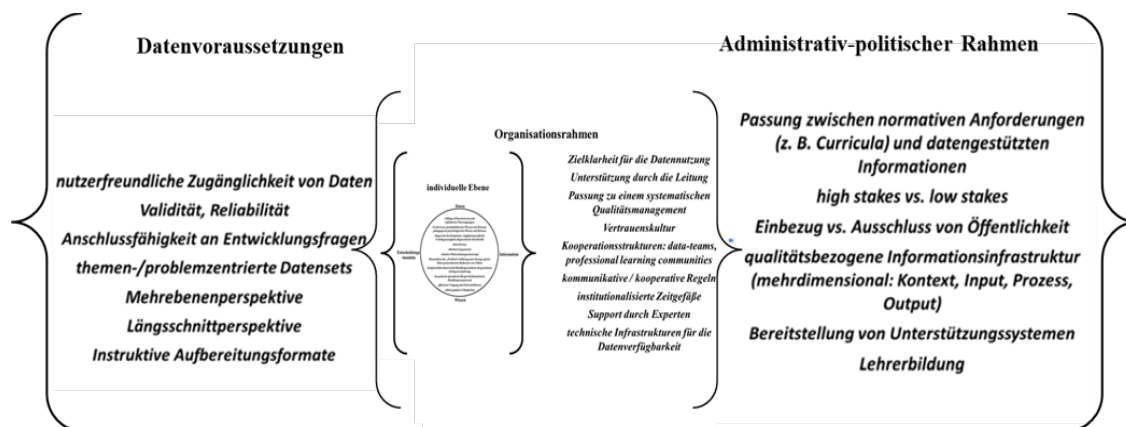


Abbildung 3: Datennutzung als mehrfach geschichteter Prozess (enthält verkleinert Abbildung 2)

In diesem bedarf es auch seitens der Infrastruktur der Datenbereitstellung einiger Voraussetzungen. Dazu gehören nicht nur technische Vorkehrungen, die eine nutzerfreundliche Zugänglichkeit der Informationen gewährleisten, und Formate der Datenaufbereitung, die für Lehrkräfte, auch für nicht datenaffine, instruktiv sind. Nicht trivial ist es auch, valide und reliable schulische Daten so zu modellieren, dass sie themen- oder problemzentrierte Anknüpfungspunkte für Entwicklungsfragen bieten, die in der Schule, dort z. B. bei einem Jahrgangsstufenteam oder in einer Fachgruppe, aktuell anstehen. Auch längsschnittliche Daten auf Individualebene, z. B. zum Übergang von Schülerinnen und Schülern an Schwellen ihrer Bildungsbiografie oder zu Lernentwicklungen über mehrere Jahre hinweg, sind für die Entwicklung von fachbezogenen oder übergreifenden pädagogischen Strategien von Lehrpersonen besonders aufschlussreich, aber alles andere als leicht und verlässlich bereitzustellen.

Auf der Organisationsebene der einzelnen Schule müssen auf Basis von Maßnahmen der Organisations- und Personalentwicklung ebenfalls wichtige Rahmenbedingungen geschaffen werden, damit Datennutzung überhaupt erfolgreich sein kann. Insbesondere die Schulleitung hat Sorge dafür zu tragen, dass im Kontext eines systematischen Qualitätsmanagements

- Ziele der Datennutzung klar formuliert sind;

- eine Vertrauenskultur etabliert ist, in der kommunikative Regeln zum professionellen Umgang mit Daten etabliert sind, die im Übrigen auch schutzwürdige Interessen von Einzelnen zu berücksichtigen haben;
- Kooperationsstrukturen fest verankert sind, in denen „data-teams“ und/oder professionelle Lerngemeinschaften auf Basis der Zusammenschau von Daten gemeinsam den Unterricht z. B. für bestimmte Schülergruppen weiterentwickeln;
- ggf. externe Beratung und Unterstützung akquiriert und finanziert wird;
- nicht zuletzt institutionelle Zeitgefäße bereitstehen, in denen diese Arbeit geleistet werden kann und
- die technische Infrastruktur für die Datenverfügbarkeit sorgt.

Über die einzelne Schule hinaus bedarf es eines mehrdimensionalen administrativ-politischen Rahmens, innerhalb dessen für eine Passung zwischen normativen Anforderungen (z. B. Curricula oder Qualitätsrahmen) und Datenerhebungsprogrammen gesorgt wird sowie die Grundfunktion von Maßnahmen des Bildungsmonitorings (high stakes vs. low stakes) geklärt ist. Datengestützte Qualitätsrückmeldungen sollten integraler Teil einer qualitätsbezogenen systemischen Infrastruktur sein, in der nach berechenbaren Regeln für unterschiedliche Adressaten und Verwendungszwecke verlässliche Informationen so aufbereitet werden, dass daraus ein funktional stimmiges und konzises, d. h. sich wechselseitig nicht störendes Zusammenspiel von Maßnahmen des Bildungsmonitorings wird.

Das auf Basis des internationalen Forschungsstandes skizzierte Modell macht deutlich: Datenrückmeldungen sind appellativ unspezifisch. Datennutzung ist kein Ableitungs- oder Transfer-, sondern ein iterativer Ko-Konstruktionsprozess. Datenrückmeldungen haben im besten Fall informatorischen Charakter, über dessen Relevanz und Integration in Wissensbestände und Handlungsfolgen der Nutzer entscheidet. Das reflektierte, fallspezifische Erfahrungswissen / die Performanz der Akteure hat im Prozessieren von Daten einen wesentlichen Stellenwert. Seitens der Datenvoraussetzungen gilt mit hoher Wahrscheinlichkeit: Nicht mehr Daten, sondern eine spezifischere, problemfokussiert aufbereitete, auf Erklärungszusammenhänge zugespitzte und in kooperativen Settings bearbeitbare „Datenzusammenschau“ ist erforderlich.

Wie kann eine Infrastruktur der evidenzbasierten Schul- und Unterrichtsentwicklung aussehen?

Evidenzbasierte Unterrichtsentwicklung ist ein voraussetzungsvolles Unterfangen. Und doch kann man an Systemen mit bereits seit langem ausgebauten Monitoringstrukturen zeigen, dass sich Kontexte schaffen lassen, in denen Evidenzbasierung der Unterrichts- und Schulentwicklung günstige Voraussetzungen hat. Im deutschsprachigen Bereich ist dies im Bundesland Hamburg bisher am weitestgehenden realisiert.

Hamburg hat bereits Mitte der 1990er Jahre begonnen, mit empirischen Verfahren Leistungsdaten zu Lernständen und Lernentwicklungen von Schülerinnen und Schülern zu erheben, sie Schulen und auch der Schulaufsicht zurückzumelden. Nach mehr als 20 Jahren Entwicklungsarbeit ist daraus eine komplexe Infrastruktur der Datenbereitstellung entstanden, die durch folgende allgemeine Merkmale gekennzeichnet ist:

- Sie ist mehrdimensional, d. h. datengestützte Informationen stehen nicht nur zu Testleistungen und Prüfungsergebnissen zur Verfügung, sondern auch zu schulischen Rahmenbedingungen (z. B. kleinräumig verfügbare Sozialdaten), zu Input-Größen (z. B.

Ressourcen, Personal etc.) und zu schulischen Prozessen (z. B. Verlaufsdaten der amtlichen Schulstatistik oder Daten zur Qualität des Unterrichts und des Schulmanagements).

- Sowohl rechtliche als auch technische Vorkehrungen ermöglichen diverse Datenverknüpfungen im Quer- und Längsschnitt, z. B. indem individuelle Schülermerkmale der Schulstatistik mit Sozialdaten der Schülerwohnorte verbunden werden.
- Die Infrastruktur umfasst alle Systemebenen, d. h. Informationen auf Schülerebene stehen ebenso zur Verfügung wie solche auf den Ebenen Klasse, Schule, Region und Gesamtsystem.
- Die Verfügungsrechte sind im Wesentlichen geklärt, d. h. wer auf welche Daten in welchem Zuschnitt Zugriff hat, ist definiert. Die professionelle Rolle / Aufgabe und deren situationsadäquate Ausfüllung bestimmen den Zuschnitt (z. B. Aggregatebene) und den Umgang mit Daten.
- Über die Jahre hinweg deutlich optimiert wurden die Formate, in denen Daten für unterschiedliche Adressaten aufbereitet werden. Sie sind wesentlich instruktiver geworden.

Nach Systemebenen geordnet seien einige Beispiele dieser Infrastruktur genannt:

- Auf *Ebene der Einzelschule* gibt es jährlich flächendeckende Erhebungen der Lernstände und Lernentwicklungen (KERMIT) in den Klassen 2, 3, 5, 7, 8 und 9. Schulen erhalten des Weiteren Rückmeldung über die Ergebnisse aller zentralen Abschlussprüfungen (abschlussbezogene Datenblätter). Auch fakultative schulinterne Evaluationsvorhaben auf Basis von Schulleistungstests (KESS) werden unterstützt. Die Schulinspektion beurteilt in sechsjährigem Intervall mit standardisierten Verfahren die Unterrichts- und Schulqualität. Schulleitungen erhalten einmal jährlich ein Datenblatt mit einer systematischen Zusammenstellung wichtiger Indikatoren: Herkunft der Schülerschaft, Förderbedarfe, Personal und Ressourcen, Inspektionsergebnisse, Ganztagesteilnahme, Übergänge und zentrale Abschlüsse. Ein Sozialindex gibt Hinweise auf die „soziale Belastung“ der Schülerschaft einer Schule, was nicht nur bei der Rückmeldung der Testergebnisse, sondern auch bei der Ressourcenzuweisung berücksichtigt wird. Der Unterrichtsausfall bzw. die Unterrichtsvertretung werden in den Schulen regelmäßig erfasst. Zu allen genannten Erhebungsquellen erhält die jeweils zuständige Schulaufsicht ebenfalls aufbereitete Informationen. Angesichts der Informationsdichte und -fülle besteht eine große Herausforderung darin, eine inhaltlich fokussierte Zusammenschau von Daten auf Einzelschulebene vorzunehmen mit dem Ziel, die Nutzungsvoraussetzungen zu verbessern.
- Auf *regionaler Ebene* stehen neben regional aufbereiteten Daten der amtlichen Schulstatistik z. B. für Bezirkskonferenzen von Schulleitungen und Schulaufsicht ad hoc erstellte Auswertungen zu Schulleistungstests, zur regionalen Verteilung von Sprachstandsdiagnosen oder von Ergebnissen zentraler Prüfungen zur Verfügung. Ein besonderes Instrument ist der web-basierte Regionale Bildungsatlas, der eine kartografische interaktive Darstellung von kleinräumigen Bildungs- und Sozialdaten für die Bereiche Kita/Krippen, Schule und öffentliche Freizeitangebote bietet. Auch die Schülerverteilung in Schuleinzugsgebieten ist dort einsehbar (<http://www.hamburg.de/bsb/regionaler-bildungsatlas-hamburg/>).
- Auf *Ebene des Gesamtsystems* sind z. B. diverse wiederkehrende Monitorings vorgesehen: zu den zentralen Abschlussprüfungen, zum Programm „Fördern statt Wiederholen“, zur Sprachförderung, zur Viereinhalbjährigen-Vorstellung (u. a. Sprachstandsfeststellung anderthalb Jahre vor der Einschulung), zum Unterrichtsausfall u.a.m. Ferner

veröffentlicht Hamburg regelmäßig einen Jahresbericht der Schulinspektion und einen Bildungsbericht. Schließlich werden zu ausgewählten Fragestellungen (z. B. Ausgaben für Bildung, Leistungsgerechtigkeit, Probleme des Übergangs nach Klasse 10, schulformspezifische Problemlagen) Systemanalysen erstellt oder – wie im Falle der Einführung der inklusiven Schule – in Kooperation mit externen Wissenschaftlerinnen und Wissenschaftlern Evaluationen von Reformvorhaben und Schulversuchen durchgeführt.

Diese Struktur der informationellen Grundversorgung bliebe weitgehend wirkungslos, wenn sie nicht auf allen Ebenen eingebunden wäre in eine verbindliche Gelegenheitsstruktur, d. h. in Settings, in denen mit geklärten Rollen die bereitgestellten Informationen zur Kenntnis genommen, interpretiert und auf die jeweils eigenen Entwicklungs- oder Steuerungsaufgaben und Entscheidungen bezogen würden. Auch dies beginnt in der Schule, wo strukturierte Beratungsprozesse zwischen Lehrkräften sowie zwischen Lehrkräften und Schulleitungen verbindlich institutionalisiert sind. Dazu sind Steuerungs- und Planungsinstrumente z. B. für Fachgruppen, für schulische Steuergruppen und für Funktionsträger nötig. Zwischen Schulleitungen und Schulaufsichten finden i. d. R. zweimal jährlich datengestützte QE-Gespräche statt, die einer fest vorgegebenen thematischen Struktur folgen. In Ziel- und Leistungsvereinbarungen legen Schulleitungen und Schulaufsichten Entwicklungsschwerpunkte fest, deren Umsetzung im Rahmen der QE-Gespräche ebenfalls bilanziert wird.

Ziel dieser Prozesse, die nur langfristig zu erreichen und immer wieder aufs Neue sicherzustellen sind, muss etwas sein, was man Systemkohärenz nennen kann, d. h. ein aufeinander abgestimmtes, konzertiertes Gesamtpaket von Maßnahmen, deren je spezifische Funktion, deren Adressaten, deren Verantwortungs- und Nutzungskontexte möglichst weitgehend geklärt und transparent sind.

Welche Kurzschlüsse sind zu vermeiden, wenn Evidenzbasierung ernst gemeint ist?

Abschließend soll auf einige Fehlvorstellungen eingegangen werden, die in den Reformvorstellungen für eine Umstellung des Mehrebenen-Steuerungsregimes auf Evidenzbasierung häufig implizit oder explizit aufscheinen. Grund dafür ist wahrscheinlich, dass die Entwicklungen in den deutschsprachigen Ländern seit einiger Zeit viel stärker von Bildungspolitik und -administration als von der Community der Lehrenden forciert werden.

- **Fehlvorstellung 1:** Konzepten der datengestützten Unterrichtsentwicklung fehlt eine handlungstheoretische Grundlage.

In der Tat ist richtig, dass die schlichte Verfügbarkeit z. B. von Schulleistungsdaten und die rechtlich geregelte Obligatorik, damit etwas anfangen zu sollen, theoretisch begründeter und empirisch geprüfter Nutzungsmodelle bedarf, die das praktische Handeln der Lehrkräfte anleiten und die Erfolgswahrscheinlichkeit der Datennutzung für pädagogisch-didaktische Entscheidungsfindungen erhöhen. Solche Modelle liegen aus Ländern mit deutlichem Erfahrungs- und Erkenntnisvorsprung vor. Sie müssen viel systematischer, gerade auch von empirisch arbeitenden Fachdidaktiken rezipiert und in die hierzulande noch sehr abstrakte Diskussion um Datennutzung eingebracht werden.

- Fehlvorstellung 2: Datennutzung in Schule und Unterricht dient vornehmlich der Rechenschaftslegung.

Im Kontext Evidenzbasierung ist dem funktionalen Spannungsverhältnis von Kontrolle und Rechenschaftslegung einerseits und Entwicklungsförderung andererseits nicht zu entkommen. Besonders wenn Datennutzung von Lehrkräften politisch induziert ist und das Vorzeichen der administrativen Aufforderung erhält, droht das Potenzial empiriegestützten professionellen Handelns verschüttet zu werden. Nicht selten verdrängen dann compliance-Fragen (Mache ich, was von mir verlangt wird?) pädagogische Entwicklungsfragen (Was muss ich noch wissen, um eine didaktische Entscheidung zu treffen?). Rechenschaftsdruck materialisiert sich häufig durch Fokussierung auf den Umgang mit Testdaten, statt eine breitere Sicht auf Datengrundlagen für pädagogische Entscheidungen zu ermöglichen. Deshalb ist es wichtig, Schulen neben Testdaten ein differenziertes Spektrum unterschiedlicher datengestützter Informationen als wichtiges Kontext- und Erklärungswissen zur Verfügung zu stellen. Eine Folge des Rechenschaftsdrucks ist auch, dass Lehrkräfte im Umgang mit Daten den Blick auf Probleme und Defizite von Schülerinnen und Schülern verengen oder Schulen Strategien der Datenmanipulation praktizieren. Es ist deshalb wichtig, Schulen und Lehrkräften ausreichend Freiraum und Kompetenzen zu geben, um mittel- bis langfristig Strukturen einer entwicklungsförderlichen Datennutzung aufbauen zu können.

- Fehlvorstellung 3: Daten heißt Testergebnisse.

Die Einführung von obligatorischen flächendeckenden Schulleistungstests hat eine Fokussierung auf Testdaten zur Folge gehabt, wenn in den letzten Jahren Programme zur Datennutzung für die Unterrichts- und Schulentwicklung aufgelegt wurden. Dabei wurde häufig nicht in Rechnung gestellt, dass solche Daten wg. mehrfacher Ausschnitthaftigkeit (nur ausgewählte Fächer, dort curriculare Verengungen, nur ausgewählte Jahrgangsstufen, nur einmalige Querschnittserhebungen u.a.m.) oft einen erheblichen „Erklärungsabstand“ zum unterrichtlichen Geschehen haben. Wenn Unterrichtsentwicklung evidenzbasiert vorangetrieben werden soll, müssen ergänzend datengestützte Zugänge viel systematischer genutzt werden, die viel enger in der täglichen Unterrichtspraxis und die Lehr-Lernprozesse eingebettet sind: z. B. Unterrichtsbeobachtungen, Schüler/innen-Feedback, formative Assessments, adaptives Testen. Lehrkräfte müssen lernen, Daten aus solchen Quellen systematisch zu triangulieren, an fachliche Lernziele für Schüler/innen rückzubinden und angemessen rückzumelden. Je enger Datennutzung an der unterrichtlichen Praxis ist, umso größer ist jedoch das Risiko für Lehrkräfte, das eigene Vorwissen oder die eigene Voreinstellung bestätigt sehen zu wollen. Solchem „confirmation bias“ zu begegnen, bedarf gründlicher Schulung.

Evidenzbasierte Unterrichtsentwicklung oder – schlichter formuliert – eine Daten systematisch nutzende Unterrichtsentwicklung ist – das sollte gezeigt werden – ein voraussetzungsvolles, aber realisierbares Unterfangen. Im Ergebnis eröffnet sie für alle Akteure deutlich günstigere Voraussetzungen für die Fokussierung auf Strategien der Unterstützung von Lehr-Lernprozessen. Erforderlich ist allerdings auch der Erwerb einer Reihe von Kompetenzen. Es bedarf

- eines Grundwissens über Grundlagen, Aussagekraft, Reichweite und Grenzen bestimmter Datenerhebungen;
- einiger Grundkompetenzen im Lesen und Verstehen numerischer und grafischer Darstellungen;
- der Fähigkeit, deskriptive, schlussfolgernde und wertende Aussagen zu trennen;

- der Fähigkeit, Daten so auszuwählen, zu analysieren und aufzubereiten, dass sie in bedeutungsvolle Informationen für eine bestimmte Fragestellung umgewandelt werden können;
- der Fähigkeit, datengestützte Informationen zu verbalisieren und mit „Stakeholdern“ regelgeleitet zu teilen;
- der Fähigkeit, Daten schlüssig in Begründungszusammenhänge für Entscheidungen einzufügen und dies explizit zu machen;
- der Fähigkeit, mit Daten ethisch angemessen, d. h. in professioneller Anerkennung von Schutzansprüchen Dritter umzugehen;
- der Fähigkeit, entwicklungsförderliche Settings für datenbasierten Austausch herzustellen und nicht zuletzt
- der Fähigkeit, die Geltungsansprüche eigener Schlussfolgerungen bescheiden zu formulieren.

ANMERKUNG

- ¹ Bezeichnungen wie „evidenzbasiert“, „empiriebasiert“ oder „datengestützt“ werden in diesem Beitrag weitgehend synonym benutzt, wohl wissend, dass sie eigentlich der Unterscheidung bedürfen und hinsichtlich des zugrundeliegenden Konzeptes von „Evidenz“ kritisch hinterfragt werden müssten. Diese Klärung kann im Rahmen des Beitrags nicht geleistet werden (vgl. exemplarisch *Bellmann et al.* 2011, *Heinrich* 2015)

LITERATUR

- Altrichter H., Moosbrugger H., Zuber J.: Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In: Altrichter H., Maag Merki, K. (Hrsg.). *Handbuch Neue Steuerung im Schulsystem. Educational Governance*, Band 7. Wiesbaden: Springer VS 2016.
- Bellmann, J., Müller, Th.: *Wissen, was wirkt. Kritik evidenzbasierter Pädagogik*. Wiesbaden: VS Verlag für Sozialwissenschaften 2011.
- Diedrich, M., Maritzen, N.: Schulaufsicht im Datendschungel. In: Huber, S. G., Arnz, S., Klieme, T.: *Schulaufsicht im Wandel. Rollen und Aufgaben neu denken*. Stuttgart: Raabe 2020, S. 138–164.
- Gummer, E. S., Mandinach, E. B.: Building a Conceptual Framework for Data Literacy. In: *Teachers College Record* 117 (April 2015), 22 Seiten.
- Heinrich, M.: Zur Ambivalenz der Idee evidenzbasierter Schulentwicklung. In: *Zeitschrift für Pädagogik* 61 (2015), Heft 6, S. 778–792.
- Helmke, A.: *Unterrichtsqualität und Lehrerprofessionalität*. Seelze: Klett Kallmeyer 2012.
- Kant, I.: Beantwortung der Frage: Was ist Aufklärung? (1784) In: Kant, I.: *Beantwortung der Frage: Was ist Aufklärung? Drei Essays*. Berlin: Europäischer Literaturverlag 2015, S. 7–16.
- Mandinach, E. B., Schildkamp, K.: The complexity of data-based decision making: An introduction to the special issue. In *Studies in Educational Evaluation* (Juli 2020) (online: <https://reader.elsevier.com/reader/sd/pii/S0191491X20301541?token=BA074359DEF1C8E3C32B136F448B772172C715FE1FFE24F49A3B5E53442785BCB6EF8B98354C208F08454C542FA416C6>).
- Mandinach, E. B., Gummer, E. S.: What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. In: *Teaching and Teacher Education* 60 (2016), S. 366–376.
- Maritzen, N. (2017): Lehrer/innenfortbildung im Kontext des Bildungsmonitorings: Möglichkeiten und Herausforderungen einer evidenzbasierten Professionalisierung. In: Kreis, I., Unterköfler-Klatzer, D. (Hrsg.). *Beispiele wirksamer Fortbildung in der Lehrer/innenfortbildung* Innsbruck: StudienVerlag, S. 27–48.
- Schildkamp, K., Handelzalts, A., Poortman, C. L., Leusink, H., Meerdink, M., Smit, M., Ebbeler, J., Hubers, M. D.: *The Data Team™ Procedure: A Systematic Approach to School Improvement*. Basel: Springer International Publishing 2018.

- Schildkamp, K.: Data-based decision-making for school improvement: Research insights and gaps. In: *Educational Research*, 61 (2019) H. 3, S. 257–273.
- Vanlommel, K., Van Gasse, R., Vanhoof, J., Van Petegem, P. (2017). Teachers' highstakes decision making. How teaching approaches influence rational and intuitive data collection. In: *Teaching and Teacher Education* 71 (2017), S. 108–119.
- Vanlommel, K., Schildkamp, K.: How Do Teachers Make Sense of Data in the Context of High-Stakes Decision Making? In: *American Educational Research Journal* 56 (2019), Heft 3, S. 792–821.

ZUM AUTOR

Norbert MARITZEN, bis zu seiner Pensionierung 2018 Direktor des Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ) in Hamburg. Vorsitzender des Wissenschaftlichen Beirats des BIFIE / jetzt IQS. Zahlreiche Publikationen zur schulischen Qualitätsentwicklung und -sicherung sowie zum Bildungsmonitoring und zu Steuerungsproblemen im Bildungswesen.