

2 Evidenzbasiertes Handeln in der Schule – eine Eingrenzung

In Vorbereitung auf eine Arbeitsdefinition wird im folgenden Abschnitt zunächst evidenzbasiertes Wissen von anderen Wissensbeständen und Handlungsorientierungen abgegrenzt. Sodann wird der Versuch unternommen, für schulische Akteure potentiell handlungsleitende Informationsquellen zu systematisieren.

Indem es die Handlungskompetenz von Lehrkräften stärkt, erweist sich zunächst das *Erfahrungswissen* der schulischen Akteure als bedeutsam für ein professionelles pädagogisches Handeln (vgl. Gruber, Harteis & Rehrl 2006; vgl. auch van Ackeren et al. 2013: 56ff.). Mit dem Konzept einer stärkeren Evidenzbasierung schulischer Arbeit wird jedoch eine Objektivierung pädagogischer Praxis angestrebt, die u. U. eine Abkehr des individuellen Akteurs von einem Handeln auf Grundlage seines Erfahrungswissens erfordert. Ergebnismeldungen bzw. die Rezeption von Ergebnissen der Bildungsforschung könnten laut Kuper (2006: 8) eine „Brückenfunktion zwischen wissenschaftlichem und praktischem Wissen“ darstellen und somit das problematische Verhältnis dieser beiden Bereiche entlasten.

Pfeffer und Sutton (2006) grenzen evidenzbasiertes Handeln von einem Handeln auf Grundlage von *Ersatzwissen* (auch: *substitutorientiertes Wissen*) ab, wobei die Autoren unter Ersatzwissen Konzepte und Leitideen verstehen, die ohne systematische Analyse ihrer Wirksamkeit zum Einsatz kommen. Begründet werden ersatzwissenorientierte Entscheidungen und Handlungen bspw. durch Hinweise auf das eigene Bauchgefühl oder auf tradierte organisationale Praxis („Das haben wir schon immer so gemacht.“).¹⁶ Analog zu van Ackeren et al. (2013: 58) können *Evidenzorientierung* und *Ersatzwissenorientierung* als Variablen in einem 4-Felder-Schema gegenübergestellt werden (vgl. Tabelle 1). Neben dem Typus der evidenzbasierten Schule/des evidenzbasierten Akteurs bzw. des Gegenteils der

16 Von der Autorin der vorliegenden Arbeit ist keine Wertung der unterschiedlichen Wissensarten intendiert, die insbesondere der Begriff *Ersatzwissen* implizieren könnte. Zweifellos sind auch das Erfahrungs- und das Ersatzwissen für die pädagogische Praxis bedeutsam, indem diese den schulischen Akteuren Handlungssicherheit bieten und Unsicherheiten minimieren. Diese Formen des Wissens sollen nicht ersetzt, sondern durch Evidenzen fundiert und „angereichert“ werden (vgl. auch Datnow & Park 2015; Fullan & Hargreaves 2012).

ersatzwissenorientierten Schule/des ersatzwissenorientierten Akteurs (= evidenzabgewandt) existieren auch Schulen oder schulische Akteure, die ihre Entscheidungen weder ausgeprägt auf Evidenzen noch auf Ersatzwissen gründen; diese können als sklerotisch bezeichnet werden. Der vierte Typus der volatilen Schule bzw. des volatilen Akteurs zeigt sowohl ein hohes Maß an Evidenzorientierung als auch an Ersatzwissenorientierung und scheint somit bestrebt, „jede verfügbare Form von Information zu nutzen, um sie für Veränderungen in schulischen Prozessen und Vorgehensweisen einzusetzen“ (van Ackeren et al. 2013: 57).

Tabelle 1: Entscheidungs- und Handlungsorientierungen von schulischen Akteuren bzw. Einzelschulen

Ersatzwissen- orientierung		Evidenzorientierung	
		<i>gering</i>	<i>hoch</i>
	<i>gering</i>	sklerotisch	evidenzbasiert
	<i>hoch</i>	evidenzabgewandt	volatil

Quelle: veränderte Darstellung in Anlehnung an van Ackeren et al. 2013: 58

Anderson, Leithwood und Strauss betonen, dass schulische Akteure auch in der Vergangenheit evidenzbasiert gehandelt hätten, jetzt sei jedoch eine andere Art von Evidenzen gefordert, deren Wirksamkeit und Bedeutung für Schulentwicklungsprozesse es erst noch empirisch zu überprüfen gelte:

„Rather, it is an admonition to make greater use of a different type of evidence than administrators and teachers have typically used. This would be formally and systematically collected evidence for which there are, preferably, estimates of reliability and validity of the kind developed and championed by the academic test and measurement community; and objective criteria and standards for evaluating performance progress and outcomes. Whether or not shifting the weight of emphasis from unobtrusive to obtrusive evidence (‘data,’ in contemporary terms) for decision making will improve schools (...) is an empirical question” (Anderson et al. 2010: 295).

Zusammenwirken unterschiedlicher Wissensarten

Bisher ist das Zusammenspiel der unterschiedlichen Wissensarten und Handlungsorientierungen nicht hinreichend geklärt. Gibt es eine dominierende Wissensform oder kann sich Wissen unterschiedlichen Ursprungs gegenseitig befördern oder behindern? Bromme (1992: 100) stellt bei Lehrkräften die Besonderheit einer professionsbezogenen Integration unterschiedlicher Wissensarten heraus. Die Wis-

sensbereiche würden nicht einfach addiert, sondern miteinander vernetzt, wobei diese auch auf die praktischen Erfahrungen bezogen würden. „Die Verschmelzung von Kenntnissen verschiedener Herkunft ist das Besondere des professionellen Wissens von Lehrern gegenüber dem kodifizierten Wissen der Fachdisziplinen, in denen sie ausgebildet sind“ (ebd.).

Als hilfreich für die Verbindung von praxisorientiertem und evidenzorientiertem Wissen erweist sich an dieser Stelle die Unterscheidung von Arnold und Schüller (1998: 61) zwischen materialem und reflexivem Wissen. *Materiales Wissen* bezeichnen die Autoren auch als Know-how oder Speicherwissen, denn hierunter ist Wissen zur Speicherung bspw. von Daten, Fakten und Theorien zu verstehen. *Reflexives Wissen* kann in (1) Methodenwissen (Wissen um Verfahrensweisen zur Informationsbeschaffung und -präsentation sowie Kommunikation von Informationen), (2) Reflexionswissen (Wissen zur Hinterfragung, Kritik, Begründung und Folgenabschätzung von Konzepten) und (3) Persönlichkeitswissen (Wissen zur Erkennung eigener Anteile und Deutungen in Interaktionen) differenziert werden. Bereitgestelltes Steuerungswissen ist vor dem Hintergrund dieser Folie zunächst als materiales Speicherwissen zu verstehen, das in reflexives Wissen überführt werden muss.

Wissen als soziale Konstruktion

Bei der Transformation in Persönlichkeitswissen zeigt sich, dass Wissen im Sinne der konstruktivistischen Wissenssoziologie (Berger & Luckmann 2007) als soziale Konstruktion verstanden werden kann. Auch Latour und Woolgar (1979) sowie Knorr Cetina (2002) zeigen in ihren Laborbeobachtungen, dass selbst Forschungsbefunde in den Naturwissenschaften – gemeinhin als besonders objektiv und valide eingeschätzt – nicht unumstößlich, sondern vielmehr das Ergebnis eines sozialen Konstruktionsprozesses sind. Damit rücken konstruktive Aspekte von Wissensgenerierung und Wissensverwendung in den Fokus, da „wissenschaftliches Wissen das Produkt von *Handlungen* ist und nicht allein von *Repräsentationen*, d. h. passiven Abbildungen“ (Weingart 2003: 77; Hervorhebungen im Original).

Auch im Bildungssystem lassen sich derartige soziale Konstruktionen nachzeichnen: So verweisen bspw. Mintrop und Sunderman (2012: 13) auf die nicht eindeutigen Befunde zu den Leistungseffekten bedingt durch die US-amerikanischen Rechenschaftssysteme, die teilweise stark von methodischen Entscheidungen und zugrunde gelegten Maßeinheiten abhängen. In der Folge lassen sich in den USA auch unterschiedliche Effekte von staatlichen Tests (= *high-stakes tests* mit der Funktion der Rechenschaftslegung) und NAEP-Ergebnissen

(= *low-stakes tests* mit der Funktion der Systemüberprüfung) feststellen (ebd.: 13f.).

Formen und Funktionen von Steuerungswissen

Aus diesem Grunde expliziert Fend (2011: 18) die Notwendigkeit einer differenzierten Betrachtung und Beschreibung unterschiedlicher Instrumente der Neuen Steuerung, in der die mit diesen Instrumenten verbundenen Funktionen berücksichtigt werden: „Je nach Testungsart (intern oder extern, standardisiert oder lehrerabhängig) und je nach den damit verbundenen Konsequenzen gehen die betroffenen Akteure unterschiedlich mit den Ergebnissen um. Wenn man die Wirksamkeit einer standardorientierten und testbasierten Steuerung untersuchen will, ist es somit ganz entscheidend, diese Intervention präzise zu beschreiben, um Hypothesen zu bilden, wie sie bei unterschiedlicher Ausgestaltung ‚wirkt‘, d. h. in Schüler/innenleistungen umgesetzt wird“. Der Autor unterscheidet interne und externe Instrumente der Ergebnismessung und berücksichtigt die mit diesen Instrumenten verbundenen rechtlichen Regelungen und Konsequenzen für Schülerinnen und Schüler sowie für Lehrkräfte (vgl. Tabelle 2).

Tabelle 2: Kombination von Instrumenten der Ergebnismessung und rechtlichen Regelungen ihrer Durchführung

RECHTLICHE REGELUNGEN DER DURCHFÜHRUNG UND FOLGEN				
		Low stakes	High stakes für Schüler/innen	High stakes für Lehrpersonen
INSTRUMENTE DER ERGEBNISMESSUNG	intern lehrerabhängig und curriculumnah	Formative Evaluation	Klassenarbeiten mit Noten und Berechtigungen	
	extern klassenübergreifend und curriculumnah	Vergleichsarbeiten Aufgabenstellungen durch Vertreter/innen der Lehrerschaft	Zentralabitur Aufgabenstellungen durch Vertreter/innen der Lehrerschaft	
	extern testbasiert und curriculumübergreifend	Standardtestungen	Standardisierte Reifeprüfung	Standardtestungen mit Klassenkennung

Quelle: Fend 2011: 18 (Nachbildung D.D.)

Es lässt sich schließen, dass für die unterschiedlichen Akteure im Bildungssystem Evidenzen und das mit ihnen verknüpfte Steuerungswissen verschiedene Formen annehmen können und die mit den Instrumenten und Verfahren verbundenen Effekte differenziert zu betrachten sind. Im nächsten Kapitel erfolgt eine Annäherung an eine Definition von Evidenzen in der Schule, indem eine Kategorisierung der potentiell zur Verfügung stehenden Informationsquellen vorgenommen wird.

2.1 Mögliche Evidenzen in der Schulpraxis

Trotz der Forderungen nach einer evidenzbasierten Steuerung und Entwicklung von Schule und Unterricht auf allen Ebenen des Bildungssystems bleibt der Begriff der *Evidenz* in der Literatur weitgehend diffus und wird oftmals nicht näher definiert und operationalisiert (z. B. Jornitz 2008; Bellmann & Müller 2011a). Verglichen mit dem Diskurs im deutschsprachigen Raum blicken v. a. die USA und Großbritannien auf eine längere Forschungstradition zur Frage der Evidenzbasierung schulischer Akteure zurück. Die Übertragbarkeit der in der anglo-amerikanischen Forschung beschriebenen Konzeptionalisierungen und Forschungsbefunde auf das deutsche Bildungssystem muss jedoch aufgrund der länderspezifischen Strukturen zumindest kritisch hinterfragt werden. Trotz der längeren Forschungstradition ist zudem auch in den englischsprachigen Ländern eine Uneinheitlichkeit der Definitionen und Konzepte festzustellen. Hier finden sich bspw. Begrifflichkeiten wie *evidence-based*, *evidence-informed*, *research-based*, *data-driven* oder *what works*, die trotz ähnlicher oder sogar identischer Bedeutung koexistieren (Davies 1999). Dabei werden auch die Begriffe *evidence*, *data*, *information* und *knowledge* teilweise synonym verwendet, teilweise wird jedoch auch von unterschiedlichen Bedeutungen bzw. Konnotationen ausgegangen (Honig & Coburn 2008: 583). Coburn und Turner (2011: 175) bezeichnen die Forschung zur Datennutzung aus diesem Grunde als „disorganized“. Mit Blick auf dieses Forschungsfeld lassen sich nach Ansicht von Coburn und Turner (2012: 100ff.) in der Literatur bisher vorwiegend drei Forschungsstränge identifizieren (vgl. auch die empirischen Befunde in Kapitel 4):

1. Studien zum Zusammenhang zwischen Initiativen zur Beförderung von Datennutzung und Outcome-Variablen: Diese Untersuchungen stellen zwar eine Verbindung zwischen Input- und Outcome- bzw. Outputvariablen her, bieten jedoch wenig Erkenntnisse über die Prozesse, die dem Outcome zugrunde liegen. Dies ist insofern problematisch, als dass aufgrund des Mangels an Wissen über die innerschulischen Mechanismen und Prozesse keine Rückschlüsse darauf gezogen werden können, wie die Interventionen zu verändern

seien, um ihre Wirkung zu erhöhen (vgl. auch die Ausführungen zur Zweck- bzw. Konditionalprogrammierung in Kapitel 1.5).

2. Studien, die vorrangig Initiativen und Interventionen zur Datennutzung fokussieren: Hier ist besonders die Ausgestaltung des Programms bzw. der Interventionen selbst Gegenstand der Forschung, wohingegen die Beeinflussung der individuellen und organisationalen Praktiken durch die Intervention weitgehend vernachlässigt wird. Aus diesem Grunde bleibt in derartigen Studien unklar, warum sich in einigen Schulen positive Effekte im Nachgang der Einführung entsprechender Interventionen oder Programme nachzeichnen lassen, in anderen Schulen jedoch nicht.
3. Eher normative als analytische Forschung: Derartige Forschung stellt zwar auf optimistische Weise die Veränderungsmächtigkeit von Daten und Datennutzung heraus („transformative power of data use“; Coburn & Turner 2012: 101f.), liefert hierfür aber wenig Belege.

Honig und Coburn (2008) untersuchten anhand einer Literaturrecherche sowie einer Analyse der Programme der Tagungen der American Educational Research Association (AERA) aus den Jahren 2004 und 2005, welche Formen von Evidenzen von Beschäftigten in der Schulverwaltung (*district central office administrators*) genutzt werden. Aufgrund der Uneinigkeit in der Begrifflichkeit bezogen die Autorinnen folgende Begriffe in ihre Suche ein: *decision making, data based, data driven, data management, knowledge utilization, policy making, research, research based, evidence, evidence based* und *working knowledge*. Die Autorinnen kommen auf Basis ihrer Literaturanalyse zu dem Schluss, dass Beschäftigte der Schuladministration ihre Entscheidungen auf unterschiedliche Formen von Evidenzen gründen und dabei auch Informationsquellen in einem sehr weiten Verständnis von Evidenz als wichtig für die Schulentwicklung ansehen. Somit beschränken sie sich nicht auf jene Evidenzquellen, deren Nutzung durch die Bildungspolitik propagiert wird:

„(...) they tend to use forms of evidence not specifically mentioned in policy designs—especially what is sometimes referred to as practitioner knowledge or local knowledge such as feedback or other input from teachers and information from parents and students. These forms of evidence seem important to central office support for school improvement efforts” (Honig & Coburn: 2008: 586).

Bei der Frage nach Datennutzung im Rahmen einer qualitativen Interviewstudie stellen Anderson, Leithwood und Strauss (2010: 312) hingegen fest, dass Schulleitungen weniger an informell gewonnene Informationen (bspw. durch kollegiale Hospitationen oder Unterrichtsvisitationen) denken, sondern sich vorwiegend auf

standardisiert erhobene Schülerleistungen oder Merkmale der Schulstatistik beziehen. In der Folge ist die Spannweite der für die Entscheidungsfindung herangezogenen Daten als eher begrenzt anzusehen. Die Autoren führen zudem publizierte Forschungsergebnisse als weitere mögliche Quellen von Evidenz an und unterscheiden zwischen dem weiter gefassten Begriff *evidence use* und *data use*, wobei sie unter letzterem formalisiert gewonnene Daten fassen, die Eigenschaften und Leistungen der Schülerinnen und Schüler fokussieren (ebd.: 294).

Knapp, Copland und Swinnerton (2007: 80; Hervorhebungen im Original) betonen, dass Daten (*data*) erst durch Interpretationsleistungen und Rekontextualisierungen der Akteure zu Evidenzen (*evidence*) werden:

„We acknowledge that ‘data’ are not the same as ‘evidence.’ Put another way, data by themselves are not evidence of anything until users of the data bring concepts, criteria, theories of action, and interpretive frames of reference to the task of making sense of the data“.

Die unklare Befundlage deutet darauf hin, dass nicht von einem eindeutigen Verständnis von evidenzbasiertem Handeln und Evidenzen ausgegangen werden kann. So können Evidenzquellen im engeren und im weiteren Sinne ausgewiesen werden. „Im ersteren Verständnis geht es um einen vergleichsweise hohen Grad an Wissenschaftlichkeit, was zugleich mit zumeist extern generierten Datensätzen und Wissensbeständen einhergeht“ (Demski et al. 2012: 139). „In einem weiteren Begriffsverständnis kann der Impuls zur Generierung von Informationen auch ein interner sein, bei dem die Akteure in der Regel über keine oder eine geringere methodische Expertise verfügen, gleichwohl ist der Anspruch gegeben, mehr oder weniger systematisch Informationen über Arbeitsprozesse und -wirkungen sowie ihre Rahmenbedingungen zu erhalten“ (ebd.).

Dietrich und Lambrecht (2012: 60ff.) greifen mit Blick auf externe Evaluationen die *Unterscheidung* von Rürup (2008) *zwischen wahrheitsorientierten und angemessenheitsorientierten Inspektionsverfahren* auf. Bei wahrheitsorientierten Inspektionen erfolgt die Legitimation durch eine Standardisierung des Verfahrens und durch die Überprüfung auf Grundlage eines externen kriterialen Bezugsrahmens. Auf Basis des vorrangigen Einsatzes von standardisierten, quantitativen Erhebungsinstrumenten werden valide und objektiv-nachprüfbare Befunde gewonnen. Angemessenheitsorientierte Inspektionsverfahren sollen hingegen Charakteristika und Kontextbedingungen der Einzelschule berücksichtigen, sodass der Einsatz standardisierter Instrumente und kriterialer Bezugsnormen hier nicht zielführend ist; „vielmehr soll die Schulinspektion eine zutreffende Deutung der Schule liefern und damit eher einer einzelfallbezogenen Bezugsnorm gerecht werden“ (Dietrich & Lambrecht 2012: 61). In derartigen Inspektionsverfahren ist die Professionalität der Evaluierenden die Grundlage für die Legitimation.

Diese Legitimation über die Professionalität der Akteure berge nach Dietrich und Lambrecht jedoch die Gefahr, dass die schulischen Akteure – in ihrer Rolle als pädagogische Professionelle – die im Nachgang der Inspektion rückgemeldeten Daten nicht als handlungsleitende Evidenzen anerkennen und insofern den intendierten Steuerungsmechanismus unterlaufen:

„Weil die evidenzbasierte Steuerung Einsicht voraussetzt, verlagert sich die Deutungsmacht darüber, welches Wissen als Evidenzen anerkannt wird, in Richtung der schulischen Akteure. Bereits das unterschiedliche Verständnis von Evidenz auf Ebene der Schulinspektionsprogrammatik verweist immanent darauf, dass es ganz unterschiedliche Quellen für Evidenz gibt. Für die schulischen Akteure eröffnet sich die Option, eigensinnig und vor dem Hintergrund ihrer schulischen Praxis über die Plausibilität und Relevanz der angebotenen Evaluationsergebnisse zu entscheiden: Sie können bewerten, inwiefern ihnen die Inspektionsergebnisse im alltagsweltlichen Sinne *evident*, also augenscheinlich plausibel („self-evidence“; vgl. Jornitz 2008, S. 206f.) erscheinen“ (Dietrich & Lambrecht 2012: 68; Hervorhebung im Original).

Fleischman (2009: 73f.) unterscheidet im Zusammenhang der unterschiedlichen Relevanz von Forschungsbefunden für die schulische Praxis zwischen *evidence-driven research reviews* und *user-driven research reviews*. Beim *evidence-driven*-Ansatz stehen die empirische Belastbarkeit der – in der Regel durch Wissenschaftler oder Experten gewonnenen – Befunde und die interne Validität der Studien im Vordergrund. Der *user-driven*-Ansatz hingegen fokussiert eine möglichst große externe Validität, die eine Generalisierbarkeit der Ergebnisse zulässt. Ausgangspunkt bei diesem Ansatz ist eine für Praktiker relevante Fragestellung, die es durch belastbare Befunde zu beantworten gilt; hier können auch Akteure aus der Praxis zur Problemlösung herangezogen werden. Durch die unmittelbare Relevanz für die schulische Arbeit könne die Frustration der Praktiker vermieden werden, die im Rahmen des evidenzbasierten Ansatzes durch die mangelnde Verwertbarkeit der Ergebnisse entstehen könne (Fleischman 2009: 73f.).

Die direkte Verwertbarkeit von Ergebnissen kann durch Ansätze schulischer Selbstevaluationen gewährleistet werden, die mit dem Verweis auf die Professionalisierung von Lehrkräften seit den 1990er Jahren etabliert wurden. Durch eine größere Nähe zur Praxis der Einzelschule und die Beteiligung der Schulleitungen und Lehrkräfte könne – so die Annahme – die Akzeptanz der generierten Befunde sichergestellt werden, die bei externen Evaluationen nicht zwingend gegeben sei und sich in der Folge in einer mangelnden Nutzung von Ergebnissen aus Fremdevaluationen manifestiere. Verfahren einzelschulischer Selbstevaluation werden meist ergänzend zu externen Evaluationen eingesetzt, teilweise jedoch auch bewusst als Alternative zur Fremdevaluation im Sinne eines „Korrektiv[s] für eine defizitär empfundene Praxis der externen Evaluation“ (Hense & Mandl

2009: 133). Gemeinhin werde allerdings bei internen Evaluationsverfahren von einer geringeren Güte als bei Fremdevaluationen ausgegangen (ebd.: 133f.).

Für die derartige anwendungsorientierte Forschung stellt Einsiedler (2011: 46ff.) die Gefahr einer „methodischen Beliebigkeit“ (ebd.: 48) heraus. In seinen Ausführungen zur didaktischen Entwicklungsforschung bedient sich der Autor des Quadrantenmodells von Stokes (1997: 70ff.). In diesem Modell werden Forschungsbemühungen hinsichtlich ihres Erkenntnisgewinns und ihres Nutzens klassifiziert (Abbildung 5). In den beiden Kategorien spiegeln sich die unterschiedlichen Erwartungshorizonte von Wissenschaftlern und Praktikern wider: Während die empirische Bildungsforschung in der Regel einen Erkenntnisgewinn fokussiert, hat für die schulischen Akteure die praktische Verwertbarkeit Priorität. Auch bezüglich des Aggregationsniveaus der Daten berichtet Kuper (2006: 11) unterschiedliche Interessenlagen. Rückgemeldete Ergebnisse stünden demnach im Spannungsfeld von Transparenz und öffentlichem Interesse einerseits sowie der Interpretationshoheit der Bildungsadministration und pädagogischen Professionen andererseits.

A: Erkenntnis

Ergebnis	ja	Typ I reine Grundlagenforschung	Typ III nutzenorientierte Grundlagenforschung
	nein	Typ IV	Typ II reine Anwendungsforschung
		nein	ja

B: Nutzen

Abbildung 5: Forschungstypen im Quadrantenmodell nach Stokes

Quelle: Einsiedler 2011: 47 (Nachbildung D.D; leicht verändert); vgl. auch Stokes 1997: 73

Zwar wird anhand dieser Ausführungen deutlich, dass nicht an eine bestehende Definition von Evidenzen im Bildungssystem angeknüpft werden kann, doch es deutet sich an, „dass in der Unbestimmtheit des Begriffes auch sein Potential liegt, nämlich Bezugspunkt für unterschiedliche Akteure im Bildungssystem zu sein, die auf ein gemeinhin anerkanntes Streben nach systematischer, verobjektivierter Steuerung rekurren“ (Demski et al. 2012: 132). Es zeigt sich, dass Informationen für unterschiedliche Akteure verschiedene Stellenwerte einnehmen können: Was für den einen eine Evidenz im Sinne harter, verlässlicher Fakten darstellen kann, bleibt bei einem anderen gänzlich unberücksichtigt aufgrund mangelnder Einsicht in die Relevanz für die eigene pädagogische Praxis.

Aufbauend auf diesen Überlegungen lassen sich potentielle Evidenzquellen für die schulische Arbeit zusammenfassend u. a. im Hinblick auf die folgenden Dimensionen unterscheiden:

- die Ebene im Schulsystem, die in den Blick genommen wird
- den Steuerungsanspruch, der mit der Evidenz verknüpft ist und der explizit oder eher implizit formuliert sein kann
- den Produzenten der Information, der aus der Schule selbst stammen oder eine externe Person oder Organisation sein kann
- den Grad der Wissenschaftlichkeit und Systematik der Evidenz (der sich in den sozialwissenschaftlichen Gütekriterien Objektivität, Reliabilität, Validität manifestiert)
- den Grad der Standardisierung (im Sinne einer Wahrheitsorientierung; kriteriale Bezugsnorm) bzw. die Berücksichtigung des individuellen Untersuchungsobjektes (im Sinne einer Angemessenheitsorientierung; individuelle Bezugsnorm)
- den Fokus des Interesses, der eher auf der Beschreibung und Entwicklung von Prozessen oder der Messung des Outputs bzw. Outcomes liegen kann
- die Funktion(en), die mit der Informationsquelle verknüpft ist (sind) (z. B. Quellen zur Leistungsüberprüfung/Rechenschaftslegung, zum Benchmarking (soziale Bezugsnorm), zur individuellen Weiterentwicklung)
- die Art der Rückmeldung, die vertraulich oder öffentlich erfolgen kann
- das Aggregationsniveau der Daten bzw. Informationen

2.2 Arbeitsdefinition: Evidenzquellen in der Schule

Um nicht bereits a priori Informationsquellen mit dem Potential des Transfers wissenschaftlicher Befunde in die schulische Praxis auszuschließen, wird im Kontext

der vorliegenden Arbeit ein *breiter Evidenzbegriff* verwendet. Dazu werden Evidenzen als „systematisch generierte, verobjektivierte und explizierte Informationen und Wissensbestände zur Wirksamkeit von Bildungsprozessen und ihren spezifischen Rahmenbedingungen“ (Demski et al. 2012: 132) verstanden und insofern vom Handlungs- bzw. Erfahrungswissen abgegrenzt. Es sei jedoch ausdrücklich darauf hingewiesen, dass damit an dieser Stelle keine Wertung der unterschiedlichen Wissensarten intendiert ist. Zweifellos ist auch das Handlungs- und Erfahrungswissen für die pädagogische Praxis bedeutsam, indem es den schulischen Akteuren Handlungssicherheit bietet und Unsicherheiten minimiert.

Welche Informationen können für schulische Akteure handlungsleitend sein? Tabelle 3 stellt potentielle Evidenzquellen in der Schule dar, wobei die in Kapitel 2.1 eingeführte Unterscheidung zwischen Evidenzen in einem engeren und in einem weiteren Sinne beibehalten wird. Zudem wird zwischen formalisierten und eher informellen Evidenzquellen differenziert. Des Weiteren wird unterschieden, ob der Fokus eher auf Prozessen oder dem erzielten Output liegt. Unter Evidenzen im engeren Sinne werden dementsprechend eher formalisierte Verfahren und Instrumente gefasst, denen im Modell der Neuen Steuerung ausdrücklich eine potentiell steuernde Funktion zugeschrieben wird. Hierzu zählen Rückmeldungen aus Schulinspektionsverfahren, Vergleichsarbeiten bzw. Lernstandserhebungen sowie zentrale Abschlussprüfungen. Mit dem Fokus der Systemebene werden auch Berichterstattungen zu Schulleistungsvergleichen und die Bildungsberichterstattung den Evidenzen im engeren Sinne zugerechnet. Unter Evidenzen in einem weiteren Sinne werden nachfolgend Informationen verstanden, die in geringerem Maße formalisiert sind und sich durch eine vergleichsweise größere Prozessorientierung auszeichnen. Häufig werden diese Informationsquellen innerhalb der Einzelschule generiert. Berücksichtigt werden das Schülerfeedback zum Unterricht, gemeinsame Unterrichtsentwicklungsmaßnahmen, aber auch Parallelarbeiten zum Vergleich von Klassen innerhalb der gleichen Jahrgangsstufe sowie schulinterne Befragungen von Lehrkräften, Schülerinnen und Schülern oder Eltern.

„Schließlich werden auch mögliche Evidenzquellen aufgegriffen, die nicht mit einer unmittelbaren Steuerungserwartung verknüpft sind, sondern vielmehr vorhandene Wissensbestände aufgreifen und aufbereitet zur Verfügung stellen. Die Beschaffung und Wahrnehmung solcher ‚vermittelter Evidenzquellen‘ liegt allerdings stärker als bei den anderen genannten Kategorien in der Verantwortung der einzelnen Lehrerin bzw. des einzelnen Lehrers“ (Demski et al. 2012: 139f.).

Ein eher impliziter Steuerungsimpuls liegt bei Aufgabensammlungen von zentraler Stelle vor, bei denen anzunehmen ist, dass sie für schulische Akteure eine Orientierung darstellen und in der Folge steuernde Wirkungen entfalten; Ähnliches

gilt für Artikel mit bildungswissenschaftlichem Bezug in Tages- oder Wochenzei-
tungen bzw. Magazinen.

„Insofern ermöglicht das auf Schule bezogene weitere Verständnis von Evidenzen eine grundlegende Exploration eines Forschungsfeldes, zu dem bislang kaum Erkenntnisse vorliegen: Welche möglichen Quellen von Evidenz gibt es außerhalb des eigenen Handlungs- und Erfahrungswissens für schulische Akteure, welche Evidenzen nehmen sie wahr, welche sehen sie als grundsätzlich hilfreich und nützlich für die Schul- bzw. Unterrichtsentwicklung an und welche finden dann tatsächlich Eingang in die Reflexion von pädagogischen Situationen bzw. beeinflussen gar die Einstellungen und das Handeln von Schulleitungen, Steuergruppen und einzelnen Lehrerinnen und Lehrern?“ (ebd.: 132).

Tabelle 3: Potentielle Informationsquellen für die schulische Arbeit

		Evidenzquellen im engeren Sinne eher ausgeprägte Wissen- schaftsorientierung, eher externer Impuls	Evidenzquellen im weiteren Sinne eher geringe explizite Wissenschaftsorientierung, eher interner Impuls
eher formalisiert, explizierter Steuerungsanspruch	Fokus Systemebene	Schulleistungsstudien, Bildungsberichterstattung	
	Fokus Schul-/ Unterrichts- ebene	<i>outputorientiert:</i> Schulinspektion, Vergleichsarbeiten, zentrale Abschluss- prüfungen	<i>eher prozessorientiert:</i> Schülerfeedback zum Un- terricht, gemeinsame Un- terrichtsentwicklungsmaßnah- men wie kollegiale Hospita- tionen, schulinterne Befra- gungen <i>eher outputorientiert:</i> schulstatistische Auswer- tungen, Parallelarbeiten, innerschulisch eingesetzte Kompetenztests
	eher informell, implizites Steuerungs- potential	z.B. Bereitstellung von Aufgabensammlungen ,vermittelte Evidenzen‘ über Artikel in schulfachbezoge- nen oder überfachlichen schulbezogenen Zeitschriften, Bildungsteile von Tages-/Wochenzeitungen oder Maga- zinen	

Quelle: Demski et al. 2012: 140 (Nachbildung D. D.; leicht verändert)

Evidenzbasierte Schulentwicklung

Empirische Analyse eines Steuerungsparadigmas

Demski, D.

2017, XV, 462 S. 40 Abb., Softcover

ISBN: 978-3-658-18077-5