

Strategien, Perspektiven und Szenarien der Unterrichtsdiagnostik

Was ist überhaupt Diagnostik?

Im Alltagssprachgebrauch gibt es eine Reihe von Begriffen, die oft synonym verwendet werden. Dies kann zu Verständigungsproblemen führen. Daher werden im folgenden Kasten einige verbreitete Konzepte skizziert, die im Bereich der Diagnostik wichtig sind.

Kasten 1: Inhaltlich verwandte Begriffe im Bereich der Diagnostik (nach Helmke, 2009, S. 268)

- *Erhebung (assessment)* ist eine unspezifische Bezeichnung für die Erfassung von Daten jeglicher Art.
- *Messung (measurement)* ist ganz allgemein der Prozess der Zuordnung von Zahlen zu Ausprägungen eines Merkmals und damit Grundlage jeder quantifizierenden Aussage.
- *Einschätzung (rating)* ist ein quantitatives Urteil über einen Sachverhalt, mit dem die Ausprägung eines Merkmals eingeschätzt werden soll. Oft geschieht dies durch die Einordnung in eine von mehreren vorgegebenen Antwortkategorien, die unterschiedliche Ausprägungsgrade des Merkmals bezeichnen. Beispielsweise kann die Ausprägung des Merkmals „Schülerbeteiligung im Unterricht“ durch die Antwortkategorien „stark“, „mittel“ oder „niedrig“ beurteilt werden.
- Unter *Diagnose* wird eine Feststellung über den Zustand einer Person oder eines Sachverhalts verstanden, die sich an bestimmten vorgegebenen Merkmalen (z. B. von Personen oder der Unterrichtsqualität) orientiert, in aller Regel theorie- und hypothesengeleitet vorgenommen wird, meistens auf der Basis geeigneter Messinstrumente (Tests, Fragebögen, Ratings) erfolgt und spezifischen Gütekriterien unterworfen ist. Das Wort leitet sich aus dem griechischen *diagnostikos* (zum Unterscheiden geschickt) ab. *Diagnostik* ist eine professionelle, systematische, wissenschaftlich und methodisch fundierte Tätigkeit mit dem Ziel, Erkenntnisse über Merkmalsträger zu gewinnen oder Entscheidungen über nachfolgende Maßnahmen treffen zu können. Dabei wird ein Ist-Stand erfasst. In Abgrenzung davon geht es bei der *Prognose* um die Beschreibung eines in der Zukunft liegenden und bei der *Retrognose* um die Beschreibung eines in der Vergangenheit liegenden Sachverhaltes.
- *Monitoring* ist die systematische und regelmäßige Überwachung von Prozessen oder Systemen (z. B. des Bildungssystems: Bildungsmonitoring) mit dem Ziel, im Falle eines unerwünschten Verlaufs (z. B. Unter- oder Überschreitung von Schwellenwerten) steuernd einzugreifen.
- *Benchmarking* ist die Standortbestimmung durch Vergleiche mit einer Bezugsgruppe, einem Referenzwert (z. B. Durchschnitt) oder mit herausragenden Personen/Firmen/Gruppen.
- *Evaluation* (Bewertung) ist der Prozess (auch: das Ergebnis) der Beurteilung des Wertes eines Produktes, Prozesses, Programms oder einer Maßnahme unter Zugrundelegung eines Gütemaßstabes, um das Programm oder Produkt selbst zu verbessern. „Evaluation“ leitet sich ab aus dem lateinischen *valere* = wert sein, bei Kräften sein; hieraus abgeleitet: *valoir* (franz.) = kosten, wert sein, *value* (engl.) = Wert, Nutzen, *evaluation* (engl.) = Schätzung, Bewertung. Von Evaluation im Sinne der *Evaluationsforschung* spricht man dagegen nur dann, wenn der Evaluation anerkannte sozialwissenschaftliche Regeln und Methoden zugrunde liegen.

Wie lässt sich Unterrichtsdiagnostik verorten?

Diagnostik im Kontext des Lehrens und Lernens kann sich auf sehr unterschiedliche Gegenstandsbereiche beziehen, wie Abbildung 1 veranschaulicht:

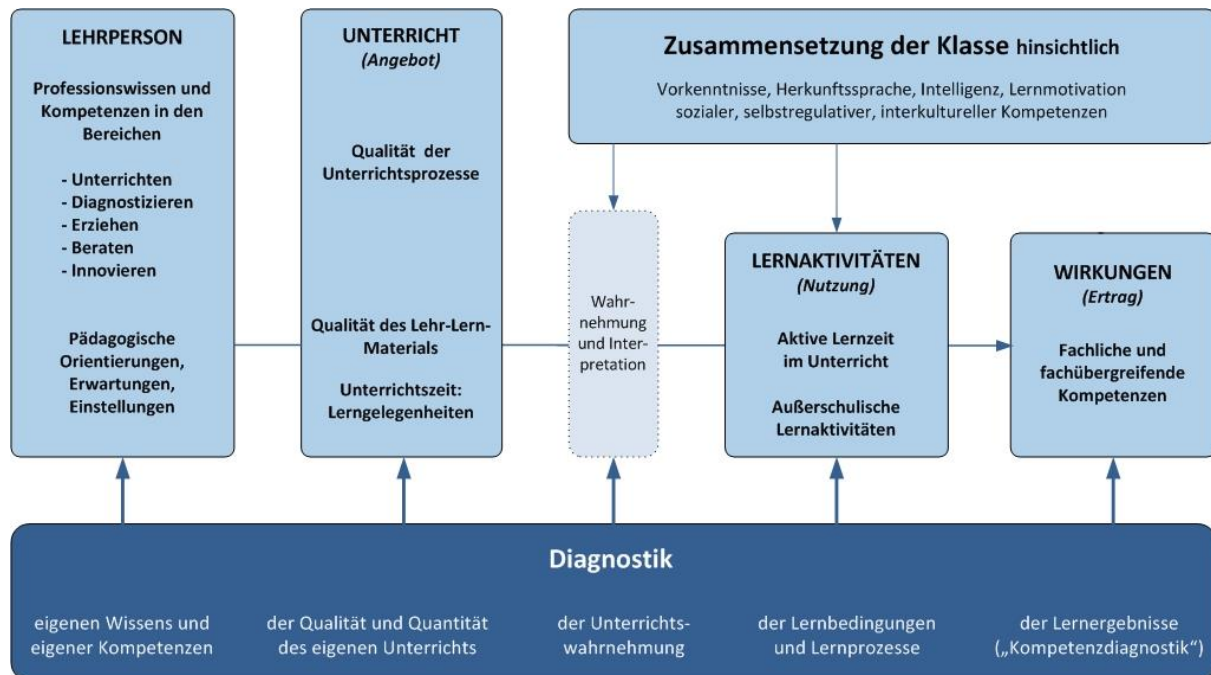


Abbildung 1: Gegenstandsbereiche der pädagogischen Diagnostik des Lehrens und Lernens

Während die *schülerbezogene Diagnostik* versucht, Lernvoraussetzungen, Lernprozesse und Lernerfolge beim Schüler zu ermitteln (siehe hierzu den UDiKom Studienbrief zur Individualdiagnostik), zielt die *Unterrichtsdiagnostik* darauf ab, den Unterricht, d. h. die Qualität der Unterrichtsprozesse zu diagnostizieren. Die Diagnose ist natürlich kein Selbstzweck, sondern dient dem Ziel, den Unterricht zu verbessern. Wir gehen dabei davon aus, dass der Unterricht das "Kerngeschäft" von Schule darstellt und daher im Zentrum der Professionalisierung von Lehrpersonen stehen sollte.

Koordinaten der Unterrichtsdiagnostik

Für die Erfassung der Unterrichtsqualität gibt es verschiedene *Methoden*, von denen im Folgenden einige zentrale kurz dargestellt werden sollen.

Freie vs. kategorienbasierte Beurteilung

Die Beurteilung von Unterricht kann entweder ohne jede Vorgabe erfolgen oder an bestimmte Vorgaben (z. B. Items, Ratings, Fragen) gebunden sein. Je geringer die Vorgaben sind, desto stärker können die Besonderheiten des einzelnen Falles berücksichtigt werden. Für Zwecke der Beratung ist dies sicher günstig. Allerdings besteht dann die Gefahr, dass bei der Auswahl und Gewichtung der Kriterien stark subjektive Gesichtspunkte ins Spiel kommen. Als state-of-the-art gilt deshalb inzwischen sowohl für die interne wie für die externe Evaluation: Grundlagen sollten kategorienbasierte Werkzeuge sein. Dieses Fundament kann dann je nach Erkenntnisinteresse durch andere (eher qualitative) Verfahren ergänzt werden.

Hoch-inferente vs. niedrig-inferente Beurteilung

Unterricht kann direkt anhand globaler Qualitätsmerkmale beurteilt werden (z. B. Klarheit oder Strukturiertheit des Unterrichts). Ein solches Vorgehen hat aber verschiedene Nachteile. So kann sich das Merkmalsverständnis verschiedener Urteiler unterscheiden oder verschiedene Urteiler verwenden unterschiedliche Beobachtungsbelege (Indikatoren) für das Merkmal. Man bezeichnet solche allgemeinen Urteile als hoch-inferent, weil ein hohes Maß an Schlussfolgerungen (Inferenzen) nötig ist, um von konkreten Beobachtungen auf das Vorliegen bestimmter Merkmale zu schließen. Es ist deshalb üblich, ein solches, globales Merkmal durch konkretere und verhaltensnähere und damit auch besser beobachtbare Merkmale zu charakterisieren und dann durch mehrere Items (= Aussagen oder Fragen) zu erfassen. Es handelt sich hierbei um niedrig-inferente Urteile, weil nur ein geringes Maß an Schlussfolgerungen nötig ist, um das Vorliegen eines Sachverhaltes zu beurteilen.

Bei einem wissenschaftlich erprobten Messinstrument muss anhand von empirischen Daten belegt werden, dass Items, die einem bestimmten globalen Merkmal zugeordnet sind, auch tatsächlich alle dieses Merkmal repräsentieren. Dies ist dann der Fall, wenn die mit verschiedenen Items erfassten Sachverhalte mit einiger Wahrscheinlichkeit auch zusammen vorkommen. Dazu ein Beispiel für den Bereich Klarheit und Strukturiertheit: Lehrer, die auf wichtige Punkte im Unterricht hinweisen, fassen häufig auch am Ende die zentralen Ergebnisse zusammen. Das bedeutet nicht, dass beide Verhaltensweisen immer zusammen auftreten müssen. Man kann jedoch ihr gemeinsames Vorkommen mit einer bestimmten Wahrscheinlichkeit erwarten.

Referenzzeitraum

Urteile zum Unterricht können sich auf unterschiedliche Zeiträume beziehen, z. B. auf eine konkrete Unterrichtsstunde oder das gesamte Schuljahr. Viele Werkzeuge zur Unterrichtsdiagnostik, die einen Abgleich von Lehrer-Selbsteinschätzungen und Schülereinschätzungen vorsehen, verwenden Urteile, die sich auf einen längeren Zeitraum beziehen (z. B. das Instrument „Schüler als Experten für Unterricht“ (SEfU) aus Thüringen und Sachsen (<http://www.kompetenztest.de/sefu/sefu1.html>) sowie NRW (<http://nrw.sefu-online.de>), das Werkzeug des IS Berlin-Brandenburg (<http://sep.isq-bb.de/>) sowie des IQ Hessen (<http://www.iq.hessen.de/>)). Der Vorteil solcher Verfahren besteht darin, dass das Urteil auf einer breiteren Erfahrungsbasis basiert als auf nur einer einzelnen Stunde.

Dem steht aber auch ein grundlegender Nachteil gegenüber: Urteile, die sich auf einen längeren Zeitraum beziehen, erfordern eine Aggregation (Durchschnittsbildung) über viele Stunden und Situationen mit anschließender Gewichtung. Wie die Forschung gezeigt hat, erfolgt die Urteilsbildung in der Realität ganz anders, vor allem dann, wenn die Zeit knapp und die Items lang sind: Es werden einfache Heuristiken (Faustregeln) verwendet (z. B. wird auf besonders markante oder noch im Gedächtnis befindliche Ereignisse zurückgegriffen statt sich an alle relevanten Ereignisse zu erinnern). In der Folge besteht die Gefahr von Stereotypen und Klischees, zudem können Urteilsverzerrungstendenzen (wie der Halo-Effekt) leichter durchschlagen.

Dies ist *ein* Grund, warum die Unterrichtsbeobachtungen bei uns auf eine konkrete Stunde bezogen sind. Ein *zweiter* Grund ist, dass bei unserem Ansatz der evidenzbasierten Unterrichtsentwicklung der Kooperation in Tandems ein großes Gewicht eingeräumt wird. Kollegiale Hospitation setzt aber praktisch immer voraus, dass sich die Beobachtungen auf kurze Zeiträume wie z. B. eine einzelne Unterrichtsstunde beziehen, für den Hospitationspartner realistischerweise auch zur Verfügung stehen. Ein *dritter*

Grund ist, dass Maßnahmen der Unterrichtsentwicklung immer nur durch eine Veränderung konkreter Verhaltensweisen in bestimmten Situationen (z. B. einer Unterrichtsstunde) erfolgen können. Denn bei einer stundenbezogener Beurteilung kann der Bezug zu konkreten Verhaltensweisen sehr viel leichter hergestellt werden als bei einer auf einen langen Zeitraum bezogenen Beurteilung.

Verschiedene Perspektiven

Unterricht kann aus verschiedenen Perspektiven (z. B. Schüler- oder Lehrerperspektive) betrachtet und bewertet werden. Wie die Unterrichtsforschung zeigt, hat jede dieser Perspektiven ihre Vor- und Nachteile, ihre Stärken und Schwächen und ihre spezifischen blinden Flecken. Die Einholung eines "fremden Blicks" auf den Unterricht, der Abgleich verschiedener Perspektiven macht den Kern des vorliegenden Programms aus. Wir sehen hierfür zwei Abgleichsmöglichkeiten für die Selbsteinschätzung des Lehrers vor: a) Abgleich mit Schülerfeedback, b) Abgleich mit kollegialer Fremdeinschätzung auf der Grundlage einer kollegialen Hospitation.

Die Zeitdimension

Maßnahmen der Unterrichtsdiagnostik gewinnen zusätzlich an Wert, wenn sie nicht nur eine Momentaufnahme bieten, sondern die Möglichkeit mehrfacher Erhebungen vorsehen. Wenn Schulen die Unterrichtsdiagnostik für eine Reflexion der Ergebnisse nutzen und geeignete Interventionsmaßnahmen in Gang setzen, dann sollte im nächsten Schritt auch geprüft werden, welchen Ertrag die eingeleiteten Maßnahmen erbracht haben. Erst dadurch lässt sich die Strategie der evidenzbasierten Unterrichtsentwicklung in vollem Umfang realisieren. Unser Diagnosewerkzeug sieht deshalb ausdrücklich eine wiederholte Erhebung vor und bietet entsprechende Rückmeldungen an. Näheres hierzu siehe unter [Erfassung von Veränderungen](#).

Werkzeuge zur Unterrichtsdiagnostik

Unterrichtsdiagnostik setzt voraus, dass geeignete Messinstrumente (z.B. Fragebögen) zur Verfügung stehen, sachgerecht eingesetzt werden, die gewonnenen Daten korrekt aufbereitet und ausgewertet werden sowie die Ergebnisse in anschaulicher und praktisch nutzbarer Form verfügbar gemacht werden können. Dies alles ist mit einem gewissen Aufwand und Know-How verbunden. Um Lehrkräfte und Schulen zu entlasten, sind daher in der Forschung EDV-Werkzeuge entwickelt worden, mit denen die Dateneingabe und Auswertung erleichtert werden. Für die technische Unterstützung gibt es derzeit zwei Ansätze:

Online-Nutzung. Bei diesem internet-gestützten Ansatz, wie er in Berlin/Brandenburg und Thüringen/Sachsen praktiziert wird, loggen sich die Schulen unter ihrer amtlichen Schulnummer in einem Portal der Selbstevaluation ein. Lehrer und Schüler können dann ihre Fragebögen am PC online ausfüllen. Eine entsprechende Software gibt kurze Zeit nach der Eingabe Rückmeldungen, z. B. Vergleiche von Selbst- und Schülereinschätzungen. Dieses Verfahren hat die Vorteile, dass durch die Nutzung moderner Technologie Zeit gespart wird und dass die Schüler ihre Daten selbst eingeben und die Lehrer somit entlastet werden. Es setzt aber voraus, dass entsprechende technische Einrichtungen (wie Laptops) und Räume (z. B. PC-Räume) vorhanden sind und ständig gepflegt und aktualisiert werden (z. B. durch einen Administrator).

Offline-Nutzung. Bei dem hier vorgestellten Diagnosewerkzeug ist lediglich für das Herunterladen der Fragebögen und der Software ein Zugang zum Internet nötig. Die eigentliche Nutzung erfolgt

dagegen offline: Die Fragebögen werden nach dem Herunterladen ausgedruckt, im Klassenverband ausgefüllt und eingesammelt. Die Daten werden anschließend manuell in die vom Programm zur Verfügung gestellte Maske eingetippt. Nach den Erfahrungen der Pilotierung benötigt man bei 30 Items für die Eingabe der Daten eines Schülers ca. eine Minute, für eine Klasse mit 25 Schülern also eine knappe halbe Stunde. Das Programm erzeugt dann auf Knopfdruck vielfältige Auswertungen (zu Beispielen für mögliche Ergebnisse und Hinweisen zu ihrer Interpretation siehe [Was bedeuten die Ergebnisse?](#)). Eine solche offline-Version hat u. a. den Vorteil, dass die Datenhoheit sichtbar bei der Lehrperson selbst bleibt.

Der mit der Dateneingabe durch die Schulen verbundene Mehraufwand ist auf den ersten Blick eine Schwäche dieses Verfahrens. Allerdings haben die im Rahmen unserer Untersuchungen gewonnenen Erfahrungen gezeigt, dass es neben der Dateneingabe durch die Lehrperson selbst eine Reihe funktionierender Alternativen gibt, z. B. Eintippen durch ein Schülertandem oder die Delegation an das Sekretariat.

Gütekriterien

Die aus der pädagogisch-psychologischen Diagnostik bekannten Gütekriterien für Messungen – Objektivität, Reliabilität und Validität (siehe den Studienbrief zur Individualdiagnostik) – betreffen auch die Unterrichtsbeobachtung und –beurteilung. Auch bei der Entwicklung der hier verwendeten Messinstrumente wurde deshalb darauf geachtet, dass diese Gütekriterien erfüllt sind. Die Sicherstellung einer hohen Messgüte ist im Rahmen von wissenschaftlichen Untersuchungen unabdingbar. Aber auch für die Praxis ist die Messgüte wichtig, vor allem dann, wenn aufgrund einer Messung Entscheidungen getroffen werden, die bedeutsame und schwer korrigierbare Konsequenzen für die Betroffenen haben. Im vorliegenden Fall sind die getroffenen Entscheidungen insofern folgenreich, als sie möglicherweise nachhaltige Auswirkungen auf die Unterrichtsgestaltung und das Lernen der Schüler haben.

Objektivität

Objektivität bedeutet hier, dass Durchführung und Auswertung von Unterrichtsbeurteilungen von der Person des Urteilers möglichst unabhängig sind. Objektivität hängt von der Präzision und Eindeutigkeit der Instruktionen ab, mit denen ein vergleichbares Vorgehen gewährleistet werden soll. Ist das Vorgehen bei der Datenerhebung, -analyse und -auswertung *standardisiert* und sind die zu beurteilenden Sachverhalte zudem klar und eindeutig beschrieben, dann ist das Instrument weniger anfällig für Verzerrungen durch verschiedene Beobachter. Eine hohe Objektivität ist Bedingung für eine hinreichende Reliabilität und Validität.

Reliabilität

Von großer Bedeutung bei der Unterrichtsbeobachtung und -beurteilung ist die Reliabilität. Sie betrifft die Genauigkeit der Messung bzw. die Verlässlichkeit, mit der ein (oft auf mehreren Einzelmessungen basierendes) Messinstrument Unterschiede zwischen den gemessenen Sachverhalten feststellen kann. Von besonderer Bedeutung ist hier die Situation, bei der verschiedene Urteiler als "Messinstrumente" fungieren und Aussagen über die "Messgüte" des einzelnen Urteilers bzw. der gesamten Urteilergruppe gemacht werden sollen (Inter-Rater-Reliabilität).

Inter-Rater-Reliabilität. Urteile sind umso reliabler (zuverlässiger), je stärker verschiedene Beurteiler bezogen auf den gleichen Sachverhalt (in unserem Fall also bezogen auf den gleichen Stundenausschnitt) in ihrem Urteil über ein bestimmtes Unterrichtsmerkmal übereinstimmen (Objektivität) und je mehr die Urteile zwischen *verschiedenen* Sachverhalten variieren. Ein allgemeines Maß dafür ist der *Intraclass-Korrelationskoeffizient*, der sowohl die Unterschiede zwischen den beurteilten Sachverhalten als auch die Unterschiede zwischen den Urteilern berücksichtigt.

Die Höhe der Übereinstimmung hängt sowohl vom zugrunde liegenden Instrument als auch von der Urteilskompetenz der Beobachter ab: Je mehr Urteils- und Ermessensspielraum die Formulierung einer Kategorie erlaubt, desto wahrscheinlicher ist es, dass die Urteile voneinander abweichen. Vor allem bei sogenannten "hoch-inferenten" Merkmalen (wie Klarheit und Verständlichkeit), deren Beurteilung auf einem hohen Ermessensspielraum beruht, spielt das Merkmalsverständnis des Urteilers und damit seine Urteilskompetenz eine wichtige Rolle. Um diese Quelle der Variation zwischen den Urteilern zu minimieren und damit die Reliabilität zu steigern, sind ein Training der Beobachter und eine genaue Beschreibung und Kommentierung der Kategorien inkl. der Angabe typischer Beispiele erforderlich. Es wäre allerdings unrealistisch zu erwarten, dass als Ergebnis von Unterrichtsbeobachtungstraining die Inter-Rater-Streuung komplett verschwindet. Für die praktische Unterrichtsdiagnostik bedeutet das, dass man versuchen muss, Quellen der Nicht-Übereinstimmung im Gespräch zu klären.

Validität

Das wichtigste Gütekriterium ist die *Validität* (Gültigkeit). Die Gültigkeit einer Messung bezieht sich darauf, ob mit der Beobachtung überhaupt das infrage stehende Merkmal, also die Qualität des Unterrichts gemessen wird. Eine hohe Übereinstimmung zwischen den Urteilern und eine hohe Reliabilität bedeuten nicht unbedingt, dass auch tatsächlich das Unterrichtsmerkmal erfasst wird, das gemessen werden soll. Die Validität kann aus sehr unterschiedlichen Gründen eingeschränkt sein – auf einige zentrale Punkte soll im Folgenden kurz eingegangen werden.

Perspektivenspezifität

Was ein Unterrichtsmerkmal bedeutet und wie es sich manifestiert, kann aus verschiedenen Urteilerperspektiven unterschiedlich verstanden werden. Beispiel „Verständlichkeit“: Ist das, was Lehrer als Experten für „verständlich“ halten, auch für Schüler verständlich? Es liegt auf der Hand, dass man für ein valides Gesamturteil unterschiedliche Perspektiven berücksichtigen muss.

Verallgemeinerbarkeit

Mit nur einem einzigen Unterrichtsbesuch lässt sich noch keine Aussage über „den“ Unterricht der betreffenden Lehrperson ableiten, geschweige denn ein Urteil über die Kompetenz der Lehrperson – das wäre ebenso unangemessen wie der Versuch, die Intelligenz einer Person mit einer einzigen Testaufgabe zu erfassen.

Gleichwohl können "Momentaufnahmen" wichtige Indizien für die Qualität des Unterrichts liefern. Diese Indizien sind dabei umso gewichtiger, je mehr Perspektiven genutzt werden. Dazu kommt, dass jeder Unterricht durch routinemäßige Abläufe und schwer beeinflussbare Wirkungen geprägt ist. Deshalb lassen sich wünschenswerte Verhältnisse (wie z. B. eine aktive Beteiligung der Schülerinnen und Schüler am Unterrichtsgespräch) in den seltensten Fällen kurzfristig herstellen. Dies gilt insbesondere für all diejeni-

gen Merkmale des Lehrerverhaltens und der Lehrer-Schüler-Interaktion, die der Lehrperson gar nicht bewusst sind, beispielsweise die Wartezeit auf Schülerantworten, den Sprechanteil der Lehrkraft, eingeschliffene Muster der Lehrer-Schüler-Interaktion, den Umgang mit Fehlern sowie für non- und paraverbale Aspekte der Kommunikation. Die Ergebnisse der DESI-Videostudie (Lehrer als Diagnostiker des eigenen Unterrichts) machen deutlich, wie sehr sich Lehrpersonen irren, wenn sie den Prozentanteil der Sprechzeit im Unterricht schätzen sollen, der auf sie selbst entfällt.

Kontextspezifität

Wir wissen aus der Unterrichtsforschung, dass Unterricht Teil eines Systems ist: Nicht nur die Wirksamkeit des Unterrichts (gemessen anhand der Erreichung von Bildungszielen oder dem Kompetenzerwerb der Schüler/innen), sondern auch die Qualität des Unterrichts selbst hängen von Rahmenbedingungen ab, die durch die Lehrkraft nicht gesteuert werden können. Zu diesen Rahmenbedingungen gehört insbesondere die Klassenzusammensetzung. Zuletzt wurde dies in der Studie DESI der KMK deutlich: In Klassen mit günstigen Eingangsvoraussetzungen ist es einfacher, anspruchsvollen Unterricht zu veranstalten. Ein vermeintlich ineffizienter Unterricht, der die verfügbare Zeit weniger für Stoffbehandlung, sondern z. B. für sozialpädagogische Maßnahmen, für private Interaktionen, für Entspannung u. a. nutzt, kann in einer „schwierigen“ Klasse womöglich die einzige Möglichkeit sein, um überhaupt akzeptable Bedingungen für die Vermittlung des Lehrstoffs zu schaffen.

Fehlendes Kontextwissen. Die Unterrichtsbeurteilung ist dann erschwert, zum Teil sogar ausgeschlossen, wenn der Unterrichtsbeobachter (wie etwa bei der externen Evaluation) keine oder nur unzureichende Informationen darüber hat, was in den Stunden davor unterrichtet wurde, welche Ziele die beobachtete Unterrichtsstunde verfolgt und wie das Vorwissen der Schülerinnen und Schüler beschaffen ist. Im Rahmen der hier vorgestellten Unterrichtsdiagnostik (als Werkzeug der internen Evaluation) wird dieses Problem dadurch entschärft, dass mehrere Perspektiven (einschließlich der eigenen) einbezogen werden.

Subjektivität

Unterrichtsbeobachtungen und erst recht -beurteilungen sind in dem Umfang invalide, in dem nicht die Qualität des jeweiligen Unterrichts, sondern die subjektiven Theorien guten Unterrichts von Beobachtern erfasst werden. Solche subjektiven Färbungen kommen immer dann ins Spiel, wenn die Beurteilung eines Sachverhaltes eigene Auslegungen, Wertungen und Interpretationen erforderlich macht. Den Einfluss solcher individueller subjektiver Theorien kann man nie ganz ausschließen. Man kann ihn jedoch verringern, zum einen durch unmissverständliche Item-Formulierungen, zum anderen dadurch, dass man sich der eigenen subjektiven Vorstellungen zu gutem Unterricht überhaupt erst einmal bewusst wird. Genau hierzu soll die Nutzung der Unterrichtsdiagnostik einen Beitrag leisten.