

Andreas Helmke / Friedrich-Wilhelm Schrader

## Merkmale der Unterrichtsqualität: Potenzial, Reichweite und Grenzen

### 1. Einführung

Die Jahrzehnte lange Vernachlässigung der Wirkungen von Schule und Unterricht, etwa in Gestalt mess- und evaluierbarer Schülerkompetenzen, hat seit der empirischen Wende der Bildungspolitik in Deutschland ein Ende gefunden. Seit der 1995 durchgeführten TIMS-Studie hat sich in Deutschland in wenig mehr als 10 Jahren Gewaltiges getan: Von der regelmäßigen Teilnahme aller Bundesländer an den internationalen Schulleistungsstudien PISA und IGLU, der Ergänzung von PISA durch die DESI-Studie (Deutsch Englisch Schülerleistungen International) bis hin zur Verabschiedung länderübergreifender Bildungsstandards, der Gründung des IQB und zu flächendeckenden Vergleichsarbeiten wie dem Programm VERA, das in der Grundschule (Ende 3. Klasse) beginnend nach und nach auf die 6. und 8. Klassenstufe ausgeweitet wurde. Die überfällige Orientierungen an nachweislichen *Wirkungen* („Output“) von Schule und Unterricht ist in Deutschland nicht mehr rückgängig zu machen; und auch in der Schweiz (Programm Harnos) sowie in Österreich ist man in der gleichen Richtung unterwegs.

Der KMK gebührt das Verdienst, neben den Bildungsstandards auch die *Professionalitätsstandards* thematisiert zu haben – ein eminent wichtiges Thema, mit dem sich gerade diese Zeitschrift wiederholt beschäftigt hat, z. B. im Schwerpunktheft „Kompetenzerwerb in der Lehrerbildung“ (Heft 1/2008), siehe Helmke, Goellner, Kleinbub, Schrader & Wagner (2008). Dieser Thematik gilt auch weiterhin das Interesse der Bildungspolitik. So hat das BMBF in seinem Rahmenprogramm zur Förderung der empirischen Bildungsforschung mit einer kürzlich ausgeschriebenen Förderlinie „*Entwicklung von Professionalität des pädagogischen Personals in Bildungseinrichtungen*“ Fragen der Kompetenz und Expertise von Lehrkräften ausdrücklich aufgegriffen<sup>1</sup>

Komplizierter verhält es sich mit der *Prozessqualität des Unterrichts*, die im Fokus dieses Artikels steht. Wie wir aus der Unterrichtsforschung wissen, hängt die Qualität des Unterrichts nicht nur von der *professionellen Kompetenz*, dem Wissen und Können von Lehrpersonen, sondern zu einem erheblichen Teil auch vom vorgefundenen *Kontext* ab, insbesondere von der Klassenzusammensetzung im Hinblick auf lernrelevante Merkmale. Dies erschwert die Benennung und die Messung von Merkmalen der Unterrichtsqualität. Gleichwohl wird seit langem darüber nachgedacht, welches Schlüsselmerkmale der Unterrichtsqualität sind. Einige dieser Zusammenstellungen sollen im nächsten Abschnitt berichtet werden. Im Mittelpunkt dieses Artikels steht die Frage, welches Potenzial, welchen Nutzen solche Merkmalskataloge der Unterrichtsqualität haben und wie ihre Reichweite und Grenzen zu bestimmen sind. Dieser Versuch erscheint uns angesichts zahlreicher Missverständnisse, Fehl- und Überinterpretationen in der pädagogischen und bildungspolitischen Diskussion überfällig zu sein.

<sup>1</sup> [http://www.bmbf.de/pub/foerderung\\_der\\_empirischen\\_bildungsforschung.pdf](http://www.bmbf.de/pub/foerderung_der_empirischen_bildungsforschung.pdf)

## 2. Merkmale der Unterrichtsqualität: eine Übersicht

Die Suche nach wesentlichen fachübergreifenden Wirkprinzipien oder Gütekriterien des Unterrichts und nach Merkmalen der Unterrichtsqualität hat eine lange Tradition. So entwickelte *Brunnhuber* bereits vor knapp 40 Jahren „Prinzipien effektiver Unterrichtsgestaltung“, die großenteils noch heute aktuell sind; darunter versteht er „allgemeine und wesentliche Grundsätze, die für das Unterrichten und Lernen gleichermaßen Geltung beanspruchen“ und „Grundsätze der Steuerung jener Bedingungsfaktoren für Lernleistungen, die es zu erkennen und anzuwenden gilt, um ein möglichst wirkungsvolles Zusammenwirken aller Faktoren zu erreichen“ (Brunnhuber, 1971, S. 14). Diese Prinzipien sind: Motivierung, Zielorientierung, Strukturierung, Aktivierung, Angemessenheit und Leistungssicherung.

Die international bekannteste Zusammenstellung von Merkmalen der Unterrichtsqualität ist vermutlich diejenige von Brophy (2000):

Tab. 1: Schlüsselvariablen der Unterrichtsqualität nach Brophy

Faktoren von Brophy	(Freie) Übersetzung
A Supportive Classroom Climate	Unterstützendes Klassenklima
Opportunity to Learn, Content covered	Lerngelegenheiten, Unterrichtszeit
Curricular Alignment	Orientierung am Lehrplan
Establishing Learning Orientations	Herstellung einer Lern- und Aufgabenorientierung
Coherent Content	Inhaltliche Kohärenz, „roter Faden“
Practice and Application Activities	Anwendung, Übung, Sicherung
Thoughtful Discourse	Aktivierende Gespräche
Scaffolding Students' Task Engagement	Unterstützung der Lerntätigkeit
Strategy Teaching	Lehren von Lernstrategien
Co-operative Learning	Kooperatives Lernen
Goal-oriented Assessment	Kriteriumsorientierte Beurteilung
Achievement Expectations	Hohe Leistungserwartungen

Eine im deutschen Sprachraum sehr verbreitete Darstellung stammt von Meyer (2004):

Tab. 2: Merkmale guten Unterrichts nach Hilbert Meyer

- Klare Strukturierung des Unterrichts (Prozess-, Ziel- und Inhaltsklarheit; Rollenklarheit, Absprache von Regeln, Ritualen und Freiräumen)
- Hoher Anteil echter Lernzeit (durch gutes Zeitmanagement, Pünktlichkeit, Auslagerung von Organisationskram, Rhythmisierung des Tagesablaufs)
- Lernförderliches Klima (durch gegenseitigen Respekt, verlässlich eingehaltene Regeln, Verantwortungsübernahme, Gerechtigkeit und Fürsorge)

- Inhaltliche Klarheit (durch Verständlichkeit der Aufgabenstellung, Monitoring des Lernverlaufs, Plausibilität des thematischen Gangs, Klarheit und Verbindlichkeit der Ergebnissicherung)
- Sinnstiftendes Kommunizieren (durch Planungsbeteiligung, Gesprächskultur, Schülerkonferenzen, Lerntagebücher und Schülerfeedback)
- Methodenvielfalt (Reichtum an Inszenierungstechniken; Vielfalt der Handlungsmuster; Variabilität der Verlaufsformen und Ausbalancierung der methodischen Großformen)
- Individuelles Fördern (durch Freiräume, Geduld und Zeit; durch innere Differenzierung und Integration; durch individuelle Lernstandsanalysen und abgestimmte Förderpläne; besondere Förderung von Schülern aus Risikogruppen)
- Intelligentes Üben (durch Bewusstmachen von Lernstrategien, passgenaue Übungsaufträge, gezielte Hilfestellungen und „überfreundliche“ Rahmenbedingungen)
- Transparente Leistungserwartungen (durch ein an den Richtlinien oder Bildungsstandards orientiertes, dem Leistungsvermögen der Schülerinnen und Schüler entsprechendes Lernangebot und zügige förderorientierte Rückmeldungen zum Lernfortschritt)
- Vorbereitete Umgebung (durch gute Ordnung, funktionale Einrichtung und brauchbares Lernwerkzeug)

Im angloamerikanischen Sprachraum ist Slavins (1997) QuAIT-Modell sehr verbreitet. QuAIT ist ein Akronym und bezieht sich auf die Konzepte: Quality, Appropriateness, Incentives, Time.

Tab. 3: QuAIT-Modell von Slavin (1997)

- *Unterrichtsqualität (Quality of instruction)*: Das Ausmaß, in dem Informationen oder Fertigkeiten so vermittelt werden, dass die Schüler sie leicht lernen können. Die Unterrichtsqualität ist hauptsächlich ein Ergebnis der Qualität des Lehrplans sowie der konkreten Inszenierung des Unterrichts.
- *Angemessene Lernniveaus (Appropriate levels of instruction)*: Das Ausmaß, in dem der Lehrer darauf achtet, dass die Schüler darauf vorbereitet sind, eine neue Unterrichtseinheit zu lernen (das heißt, dass sie die notwendigen Fertigkeiten und das notwendige (Vor-)Wissen haben, sie zu lernen), ohne sie aber bereits gelernt haben. In anderen Worten, das Lernniveau ist angemessen, wenn eine Unterrichtseinheit weder zu schwer noch zu leicht für die Schüler ist.
- *Anreiz (Incentive)*: Das Ausmaß, in dem der Lehrer darauf achtet, dass die Schüler motiviert sind, an ihren Aufgaben zu arbeiten und sich mit dem vorgegebenen Stoff auseinander zu setzen.
- *Zeit (Time)*: Das Ausmaß an Zeit, das den Schülern gewährt wird, um den Unterrichtsstoff zu lernen.

Auf der Basis dieses Modells hat Ditton (2000) eine detailliertere Gliederung bedeutender Faktoren des Unterrichts vorgelegt:

Tab. 4: Faktoren des Unterrichts nach Ditton (2000)

<p><b>Qualität</b></p> <ul style="list-style-type: none"> <li>● Struktur und Strukturiertheit des Unterrichts</li> <li>● Klarheit, Verständlichkeit, Prägnanz</li> <li>● Variabilität der Unterrichtsformen</li> <li>● Angemessenheit des Tempos</li> <li>● Angemessenheit des Medieneinsatzes</li> <li>● Übungsintensität</li> <li>● Behandelter Stoffumfang</li> <li>● Leistungserwartungen und Anspruchsniveau</li> </ul> <p><b>Angemessenheit</b></p> <ul style="list-style-type: none"> <li>● Angemessenheit des Schwierigkeitsgrades</li> <li>● Adaptivität</li> <li>● Diagnostische Sensibilität/Problemsensibilität</li> <li>● Individuelle Unterstützung und Beratung</li> <li>● Differenzierung und Individualisierung</li> <li>● Förderungsorientierung</li> </ul> <p><b>Motivierung</b></p> <ul style="list-style-type: none"> <li>● Bedeutungsvolle Lehrinhalte und Lernziele</li> <li>● Bekannte Erwartungen und Ziele</li> <li>● Vermeidung von Leistungsangst</li> <li>● Interesse und Neugier wecken</li> <li>● Bekräftigung und Verstärkung</li> <li>● Positives Sozialklima in der Klasse</li> </ul> <p><b>Unterrichtszeit</b></p> <ul style="list-style-type: none"> <li>● Verfügbare Zeit</li> <li>● Lerngelegenheiten</li> <li>● Genutzte Lernzeit</li> <li>● Inhaltsorientierung, Lehrstoffbezogenheit</li> <li>● Klassenmanagement, Klassenführung</li> </ul>
---

In den Qualitätsagenturen (oder vergleichbaren Institutionen) der Bundesländer bzw. Kantone sind eine Reihe von interessanten Klassifikationen entstanden; sie fungieren zugleich als Basis für die Entwicklung entsprechender kategorienbasierter Bögen zur Beurteilung („Rating“) von Merkmalen der Unterrichtsqualität. Exemplarisch für die Vielzahl vorliegender Instrumente wird auf die zusammenfassende Darstellung der Konstrukte verwiesen, die dem bayerischen Unterrichtsbeobachtungsbogen<sup>2</sup> zugrunde liegen.

<sup>2</sup> <http://www.isb.bayern.de/isb/>; der Erstautor ist von Anfang an an der Ausbildung der bayerischen Evaluationsteams (durch die Qualitätsagentur am Staatsinstitut für Schulqualität und Bildungsforschung und die Akademie für Lehrerfortbildung und Personalführung) beteiligt gewesen.

Tab. 5: Kurzbeschreibung der Qualitätsdimensionen des bayerischen Unterrichtsbeurteilungsbogens

Klassenführung	Effiziente Klassenführung ist eine zentrale Grundlage für Unterricht und Erziehung in der Schule, weil sie den nötigen <b>Orientierungsrahmen</b> für die Schüler schafft und ein hohes Maß an <b>aktiver Lernzeit</b> ermöglicht. Effiziente Klassenführung zielt durch <b>Planung</b> und <b>vorbeugende Maßnahmen</b> einerseits, durch <b>situationsgemäße Flexibilität</b> andererseits darauf ab, Disziplinprobleme von vornherein zu vermeiden und ihnen gegebenenfalls angemessen zu begegnen.
Unterrichtsklima	Eine <b>positive Grundeinstellung gegenüber Lernen und Leisten</b> sowie ein <b>vertrauensvolles Klima</b> zwischen Lehrkräften und Schülern und zwischen den Schülern sind Grundlagen für Lernbereitschaft und Lernvermögen.
Motivierung	Die für schulisches Lernen nötige <b>intrinsische Motivation</b> der Schüler kann durch entsprechende <b>Motivierungsstrategien</b> , vor allem durch Strategien zur Anregung von Neugierverhalten und Sachinteresse erhöht werden. Darüber hinaus tragen Methoden der <b>extrinsischen Motivierung</b> durch positive Verstärkung und überzeugendes Verhalten, durch das die <b>Lehrkraft als Modell</b> für die Schüler dient, zu einer Erhöhung der Lernmotivation bei.
Strukturiertheit	Neu zu erwerbendes Wissen muss mit bereits vorhandenem Wissen verknüpft werden, weshalb <b>Bezüge zu schulischem Vorwissen</b> , zu <b>Alltagswissen</b> und zu <b>alterstypischen Vorstellungen</b> der Schüler über den Gegenstandsbereich wichtig sind. Wenn Wissen vernetzt werden soll, müssen Unterrichtsinhalte zudem <b>sprachlich klar</b> und <b>gut strukturiert</b> , sowie in einem <b>angemessenen Unterrichtstempo</b> dargeboten werden.
Zielorientierung	Guter Unterricht orientiert sich sowohl an konkret <b>definierten Lernzielen für die einzelne Unterrichtsstunde</b> als auch an übergeordneten <b>Zielzusammenhängen</b> .
Individuelle Unterstützung/ Fördermaßnahmen	Guter Unterricht orientiert sich nicht am fiktiven Durchschnittsschüler, sondern nimmt Bezug auf die real vorhandenen Lern- und Leistungsunterschiede. Mit Hilfe einer entsprechenden <b>Anpassung der Lehrstrategien</b> wird versucht, bei allen Schülern ein individuelles Optimum an Lernfortschritten zu bewirken. Eine <b>fundierte Diagnose</b> von Lernpotenzialen einerseits und Leistungsgrenzen andererseits, die immer wieder überprüft und optimistisch interpretiert werden sollte, ist Voraussetzung für <b>unterschiedliche Lernziele und Aufgabenstellungen</b> .
Selbstständiges Lernen	Der Erwerb von anwendungsfähigem Wissen ist ein <b>aktiver und konstruktiver Prozess</b> . Selbst gesteuertes und selbst kontrolliertes Lernen sollte deshalb häufig sowohl <b>Mittel</b> als auch <b>Ziel</b> des Unterrichts sein.
Variabilität der Unterrichtsformen	Kein Lehrverfahren ist für alle und für alles gleichermaßen geeignet. <b>Sachgemäßer Methodenpluralismus</b> und <b>flexibles pädagogisches Handeln</b> kennzeichnen deshalb guten Unterricht. Lehrkräfte sollten also die spezifische Leistungsfähigkeit verschiedener Unterrichtsmethoden kennen und entsprechend eine <b>didaktisch begründete Wahl</b> der jeweils geeigneten vornehmen können.
Lernerfolgssicherung	In einem guten Unterricht gibt es ausreichend <b>Zeit zum Üben und Wiederholen</b> . Es wird überprüft, ob auch wirklich gelernt worden ist, was gelernt werden sollte. Die <b>Hausaufgaben</b> sind eine <b>sinnvolle Ergänzung</b> des Unterrichts. <b>Überforderungen</b> und <b>Unterforderungen</b> gilt es zu vermeiden.
Leistungserhebungen	Faire Leistungserhebungen sind gekennzeichnet durch <b>Klarheit</b> der Aufgabenstellungen und <b>Transparenz</b> der Leistungsanforderungen und Bewertungskriterien. Sie weisen einen <b>Bezug zum Unterricht</b> insofern auf, als einerseits die Anforderungen eindeutig aus diesem hergeleitet sind und die Ergebnisse andererseits als Rückmeldung für die künftige Unterrichtsgestaltung genutzt werden.

Weitere wichtige und einflussreiche Klassifikationen stammen von Walberg und Paik (2000) sowie von Borich (2007), dessen Bücher zum Thema „Effective teaching“ zu den auflagenstärksten Werken im angloamerikanischen Sprachraum zählen. Einen ausführlichen Überblick über gängige Klassifikationen verdanken wir Scheltwort (2006).

Jenseits der wissenschaftlichen Welt, aber (so steht zu befürchten) nicht weniger einflussreich, sind Listen und Schemata, mit denen Schüler ihre Lehrer anonym bewerten, etwa im Webportal spickmich.de. Dort heißt es beispielsweise:

Tab. 6: Auszug aus der Lehrer-Evaluations-Website von spickmich.de

*Guter Unterricht:* Hast du das Gefühl, im Unterricht was gelernt zu haben oder schwirren nachher 1000 Fragezeichen in deinem Kopf herum? Ist der Unterricht spannend und interessant gemacht?

*Cool & witzig:* Ist der Unterricht locker und interessant oder zum Einschlafen? Wer lernt nicht besser, wenn der Lehrer auch mal einen guten Witz reißt?

*Gut vorbereitet:* Auch Lehrer haben Hausaufgaben: Sie müssen sich für den Unterricht am nächsten Tag vorbereiten! Kommen sie also gut vorbereitet und mit kreativen Ideen zum Unterricht, oder müssen sie erstmal fragen, was in der letzten Stunde so gemacht wurde?

*Motiviert:* Motivation ist alles! Überrascht euch der Lehrer mit guten Ideen zu jeder Stunde? Sind die Arbeiten immer schnell korrigiert? Oder besteht der Unterricht nur aus dem Abschreiben uralter Folien?

*Fachlich kompetent:* Egal was man den Lehrer fragt, er hat immer die richtige Antwort parat. Er kennt Dinge, die bei „Wer wird Millionär“ erst ab 250 000 Euro gefragt werden? Oder ist „weiß ich nicht genau“ eine der häufigsten Antworten?

Die aus unserer Sicht wesentlichen Qualitätsbereiche (siehe z. B. Helmke, 2009) sind:

Tab. 7: Merkmale der Unterrichtsqualität (Helmke, 2009)

1. Klassenführung
2. Klarheit und Strukturiertheit
3. Konsolidierung und Sicherung
4. Aktivierung
5. Motivierung
6. Lernförderliches Klima
7. Schülerorientierung
8. Kompetenzorientierung
9. Umgang mit Heterogenität
10. Angebotsvariation

Genau diese Struktur findet sich auch im rheinland-pfälzischen Qualitätsrahmen für Schulen<sup>3</sup>, an dessen Formulierung der Erstautor beteiligt war. Zu diesen Merkmalen und ihrer inhaltlichen Bedeutung wird auf das Buch „Unterrichtsqualität und Lehrerprofessionalität – Diagnose, Evaluation und Verbesserung des Unterrichts“ verwiesen: Es handelt sich um eine völlige Neubearbeitung des erstmals 2003 erschienenen und bis zur 5. Auflage 2006 in unveränderter Form publizierten Buches „Unterrichtsqualität – erfassen, bewerten, verbessern“, die Anfang 2009 im Verlag Klett-Kallmeyer erscheint.

<sup>3</sup> <http://aqs.rlp.de/externe-evaluation/orientierungsrahmen-schulqualitaet.html>

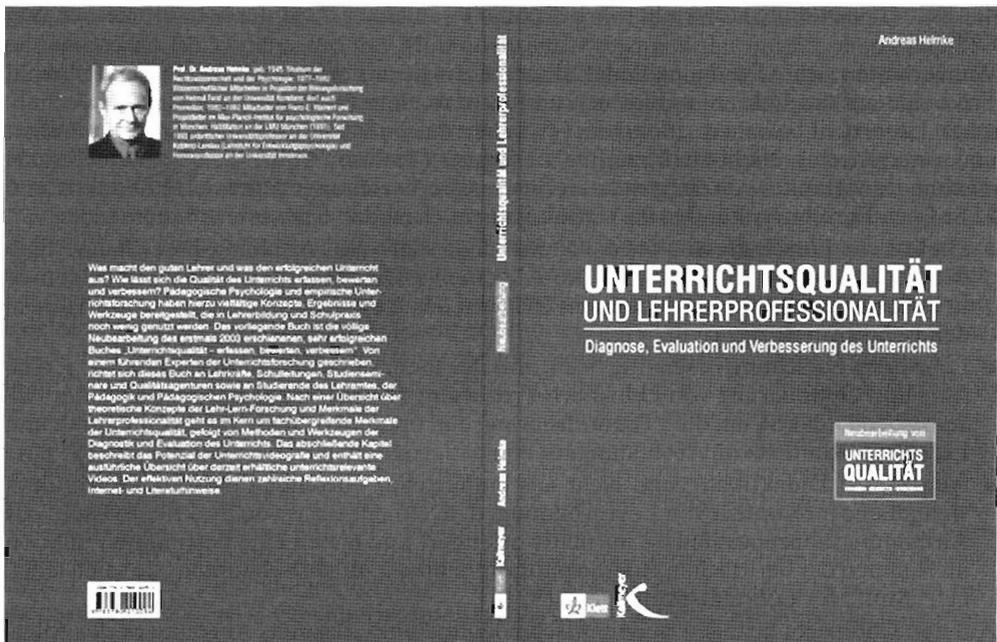


Abb. 1 „Unterrichtsqualität und Lehrprofessionalität“ (Helmke, 2009)

Allerdings – und damit kommen wir zum Kern dieses Artikels – sind solche Zusammenstellungen, wie wir aus zahlreichen Veranstaltungen sowohl der Lehrerausbildung als auch der Lehrerfortbildung erfahren haben, nicht unproblematisch. Allzu oft werden sie missverstanden, fehl- oder überinterpretiert. Darüber hinaus gibt es in der Pädagogik Stimmen, die jegliche Suche nach „Merkmalen“ der Unterrichtsqualität für obsolet und unsinnig halten, z. B. Mühlhausen (2008).

Welches Potenzial und welche Reichweite haben solche Auflistungen? Warum werden sie so häufig missverstanden? Wie ist es zu erklären, dass „Merkmale“, „Prinzipien“ oder „Variablen“ der Unterrichtsqualität in der internationalen empirischen Bildungs- und Unterrichtsforschung (Brophy 2000) eine überragende Rolle spielen, in Deutschland jedoch ein Schattendasein fristen? Ein Grund scheint zu sein, dass das in Deutschland vorherrschende Verständnis von Unterrichtsqualität und erfolgreichem Unterrichten stark von der *Allgemeinen Didaktik* geprägt ist, deren Denkweise sich erheblich von der der empirischen Unterrichtsforschung unterscheidet. Die damit verbundenen Fragen und die daraus resultierenden Konsequenzen für die Lehrerbildung bilden den Kern dieses Artikels.

### 3. Die Sichtweise der Allgemeinen Didaktik

Die Allgemeine Didaktik (vom griechischen *didaskein* = unterrichten, lehren“ bzw. *didaktike techne* = Lehrkunst) ist ein Teilgebiet der deutschen Erziehungswissenschaft und steht traditionell im Fokus der Lehrerausbildung. Die Vielfalt der heute anzutreffenden Varianten didaktischer Theorien, Schulen und Lager hier auch nur ansatzweise darstellen zu wollen, würde den Rahmen dieses Aufsatzes sprengen; aus Sicht der Erziehungswissenschaft siehe hierzu die Hinweise bei Arnold, Koch-Priewe & Lin-Klitzing (2007), aus Sicht

der Lehrerausbildung Huwendiek & Bovet (2005), und aus Sicht der Schulpraxis Schmoll (2008).

Wie verhalten sich Allgemeine Didaktik und Lehr-Lern-Forschung zueinander? Obwohl sich beide auf das Lehren und Lernen und insbesondere (aber nicht nur) auf den schulischen Unterricht beziehen, „dominiert zwischen diesen beiden Teildisziplinen Fremdheit und organisiertes Nicht-zur-Kennntnis nehmen“ (Terhart, 2002, S. 77). Von Terhart stammt die Metapher von der gestörten Geschwisterbeziehung: Allgemeine Didaktik und Lehr-Lern-Forschung stünden im Verhältnis vom „fremden Schwestern“ zueinander:

- „Lehr-Lern-Forschung ist – wie der Name schon sagt – ein Forschungsbereich innerhalb der Pädagogischen Psychologie. Die Allgemeine Didaktik ist demgegenüber gerade kein Forschungsbereich, sondern ein Element des Ausbildungsprozesses von angehenden Lehrern“ (S. 80).
- Lehr-Lern-Forschung beschäftigt sich „mit der theoriegeleiteten Beschreibung, Erklärung und Optimierung von Lehr-Lern-Prozessen“ (S. 80). Dagegen geht es in der Allgemeinen Didaktik um die „Theoretisierung und operative Gestaltung von Lehren und Lernen im Kontext von Ausbildung für den pädagogischen Beruf des Lehrers“ (S. 80).
- Allgemeine Didaktiken „sind nicht aus empirischer Forschung entstanden, sondern weitgehend im Kontext von Lehrerbildung, d. h. auf der Basis von Praktikerwissen. Als Ausbildungskonzepte sind sie normativ orientiert und haben eine bestimmte Funktion im Rahmen der beruflichen Erst-Sozialisation; böse Zungen behaupten, sie seien lediglich „Stoff“ für Lehramtsprüfungen“ (S. 81).

Auf den Unterricht bezogen ist der entscheidende Punkt, dass das Augenmerk der Allgemeinen Didaktik auf *allgemeine Prinzipien und Regeln der Gestaltung* gerichtet ist, die bei der Planung und Realisierung des Unterrichts beachtet werden müssen. Insofern ist die Allgemeine Didaktik „grundlegend für die Professionalität und Berufsidentität der Lehrer und Lehrerinnen. Als Unterrichtsexperten brauchen sie didaktisches Grundwissen und entsprechende Reflexions- und vor allem Handlungskompetenz: Sie sollen Unterricht professionell vorbereiten und planen, durchführen und gestalten, analysieren und auswerten können“ (Huwendiek & Bovet, 2000, S. 11). Die allgemeine Didaktik hat in der Vergangenheit eine Vielzahl von Regeln, Schemata und Anweisungen für den „guten“ Unterricht und Konzepte idealtypischer Stundenverläufe hervorgebracht, die in der Lehrerausbildung noch heute dominieren. Den meisten Ansätzen ist gemeinsam, dass sie nicht empirisch, sondern *normativ* sind, d. h. es geht vornehmlich um die Identifikation und Begründung der Ziele und Inhalte schulischen Lernens.

Für die Beantwortung der Frage nach wissenschaftlich fundierten, empirisch belegbaren Merkmalen der Unterrichtsqualität ist die Allgemeine Didaktik dagegen wenig ergiebig, weil sie weitgehend empiriefremd ist: Von wenigen Ausnahmen abgesehen, hat die Allgemeine Didaktik kaum *empirische Forschung* angeregt, und wo sie es getan hat, geht es eher um Erfahrungsberichte, Modellversuche, Einzelfallstudien und deskriptive Berichte und nicht um hypothesenprüfende Studien der Lernwirksamkeit des Unterrichts, die methodischen Standards genügen (wie kontrollierte Längsschnittstudien, Unterrichtsexperimente und Trainingsstudien) und deren Aussagen den Anspruch erheben, verallgemeinerbar und belastbar zu sein. Es ist daher auch kaum etwas darüber bekannt, ob, in welchem Umfang und unter welchen Bedingungen die von der Allgemeinen Didaktik entwickelten Unterrichtskonzepte und -ansätze tatsächlich auch zu erfolgreichem Lernen beitragen, welche einzelnen Elemente dafür maßgeblich und welche vielleicht verzichtbar sind.

Ein wesentlicher Grund für die Kluft zwischen Lehr-Lern-Forschung und Allgemeiner Didaktik liegt nach unserer Einschätzung in der an *Variablen* orientierten Denk- und Forschungsweise der Lehr-Lern-Forschung, die sich von der ganzheitlichen und auf Abläufe

gerichteten Sichtweise der Allgemeinen Didaktik grundlegend unterscheidet. Für Lehrpersonen und für andere, die mit Methoden und Prinzipien der empirischen Sozialforschung und der Statistik wenig vertraut sind, ist die „Zerlegung“ des Unterrichts in unterschiedliche Aspekte (oder Komponenten, Segmente, Teilmerkmale) zunächst ein ungewohnter, ja befremdlicher Zugang (siehe hierzu Helmke, A., Helmke, T., Heyne, Nordheider & Schrader, 2006). Wenn es darum geht, den Unterricht anhand von Kriterien zu beurteilen, liegt in der *Schulpraxis* der Schwerpunkt meist darauf, welche *Methoden* eingesetzt werden, um *Unterrichtsabläufe* lernwirksam zu gestalten. In der *Empirischen Unterrichtsforschung* dagegen charakterisiert man Unterricht im Hinblick auf bestimmte *Qualitätsdimensionen* (z. B. Strukturiertheit, Verständlichkeit, Motivierung), die nachweislich (belegt durch vorangegangene Untersuchungen) eine Rolle für den Lernerfolg spielen. Es handelt sich also um eine „variablenorientierte“ Betrachtungsweise, die davon ausgeht, dass der Unterricht verschiedener Lehrpersonen hinsichtlich bestimmter Merkmale (Variablen), die von Lehrkraft zu Lehrkraft, aber auch bei der gleichen Lehrkraft in verschiedenen Situationen (Fächer, Klassen, Zeitpunkte) *variieren* kann (z. B. hohe Ausprägung der Variable 'Verständlichkeit' bei Lehrperson A, niedrige Ausprägung dieser Variable bei Lehrperson B). Die *empirische Unterrichtsforschung* untersucht, ob unterschiedliche Ausprägungen solcher Variablen mit Unterschieden im Lernerfolg (z. B. Leistungszuwachs, Verbesserung der Lernfreude) einhergehen – eine Perspektive, die auch in der „output“-Orientierung der Bildungsstandards und neuen Kernlehrpläne ihren Niederschlag findet. Ein analoges Vorgehen, nämlich Abläufe durch Änderung einzelner Elemente (wenn möglich sogar experimentell) zu variieren und auf diese Weise die möglichen Wirkfaktoren des Lernerfolgs genauer zu identifizieren, ist dem ganzheitlichen Ansatz der Allgemeinen Didaktik aber fremd.

Wellenreuther (2005) weist darauf hin, dass insbesondere *experimentelle Methoden* – die methodisch am strengsten sind und am ehesten Aussagen über Wirkmechanismen und Kausalität zulassen – seitens der überwiegend geisteswissenschaftlich geprägten deutschen Pädagogik, insbesondere von der Kritischen Theorie, abgelehnt wurden: Die Unterrichtsforschung wurde „mit einem Anspruch konfrontiert, dem (sie) nicht genügen konnte: Sie sollte den umfassenden Einfluss schulischen Erziehens und Unterrichtens auf die Lernfähigkeit sowie die die Entwicklung der Heranwachsenden untersuchen ... Damit wurden die Weichen für eine Forschungsorientierung gestellt, die immer den ganzen Unterrichts- und Erziehungsprozess im Blick behält und die damit nicht mehr pragmatisch isolierte Lehr-Lern-Prozesses im Unterricht untersuchen kann. Experimentelle Forschung wird unter solchen Vorzeichen leicht als technologische Forschung diffamiert. Da sich experimentelle Forschung auf die Prüfung der Wirkung einzelner Faktoren konzentrieren muss, kann sie das Ganze nicht hinreichend im Blick behalten. Dieser überhöhte Anspruch an empirische Unterrichtsforschung trug wesentlich dazu bei, dass in der deutschen Pädagogik experimentelle Forschung nicht richtig etabliert ist. Deutschland ist hier noch immer ein Entwicklungsland“ (S. 32).

Die zentrale Rolle der Allgemeinen Didaktik in der Lehrerbildung mag auch damit zusammenhängen, dass die Brücke zur Unterrichtsgestaltung und -entwicklung hier – zumindest auf den ersten Blick – leichter zu schlagen ist. Die Allgemeine Didaktik gibt dem Lehrenden direkte Hinweise, wie er seinen Unterricht aufbauen kann, d. h. welche Schritte er ausführen sollte, und liefert ihm auf diese Weise gewissermaßen ein Handlungsgerüst. Die empirische Unterrichtsforschung liefert dem praktizierenden Lehrer dagegen Hinweise auf lernwirksame Merkmale und Prinzipien, deren Umsetzung und sachgerechte Nutzung alles andere als trivial ist. Auf die damit verbundenen Probleme wird in Abschnitt 5 (Frage 20) noch detaillierter eingegangen.

#### 4. Die Kritik von Mühlhausen

Aus anderen Gründen wendet sich Mühlhausen (2008) gegen die Zusammenstellung von Merkmalen der Unterrichtsqualität. Er betont in seinem Werk den Aspekt der weitgehenden Unberechenbarkeit und Unplanbarkeit des Unterrichts und fordert eine „überraschungsoffene Grundhaltung“. Er wendet sich zunächst – aus einer ganz anderen Perspektive als derjenigen der Empirischen Unterrichtsforschung – gegen die Dominanz der Allgemeinen Didaktik in der Lehrerausbildung:

Wie gering die Prägestkraft der in der Ausbildung vermittelten didaktischen Konzepte für die tägliche Unterrichtspraxis ist, wird inzwischen von vielen Seiten beklagt. Die geringe Aufnahmebereitschaft für didaktische Empfehlungen überrascht aber nicht, wenn man den Ausbildungsgang als Ganzen betrachtet. Angehende Lehrer sind in den verschiedenen Ausbildungsabschnitten mit jeweils anderen Ausbildern konfrontiert, die i. d. R. jeweils andere (fach-)didaktische Konzepte bevorzugen. So merken die Lehrer erst nach und nach, wie kontrovers und z. T. unvereinbar die konkurrierenden Positionen sind. Von Ausbilderseite wird dagegen das jeweils favorisierte Konzept als Nonplusultra ausgegeben, Schwächen und Ungereimtheiten werden eher selten angesprochen. Aber nicht lange bleibt Auszubildenden verborgen, dass ein gerade als aktuell und fundamental vorgestelltes Konzept von anderen Ausbildern überhaupt nicht geschätzt wird (manchmal nicht mal bekannt ist). Nicht verwunderlich ist daher, dass Lehrer spätestens nach der zweiten Phase dieser Art von (fach-) didaktischer Theorie misstrauen und eklektisch ihr eigenes Konzept zusammenbasteln, bei dem dann kaum noch Spuren des in der Ausbildung Vermittelten wieder zu finden sind [...] Didaktische Literatur wird benötigt für die schulpädagogische Ausbildung und kann spätestens nach der Prüfung vergessen werden. Ironisch prägnant resümiert der Leiter der deutschen PISA-Studie Andreas Schleicher: „Deutschland ist Weltmeister im Entwickeln didaktischer Konzepte. Leider werden sie hierzulande nicht eingesetzt.“ Polemisch ausgedrückt haben diverse Didaktiken in etwa denselben Charme wie visionäre Automodelle auf Automobilmessen, die – hochglanzpoliert, aber fahruntüchtig -- im Rampenlicht präsentiert werden. Im Unterschied zu den Produzenten didaktischer Konzepte kann es sich die Automobilindustrie allerdings nicht leisten, auf den Bau alltagstauglicher Fahrzeuge und deren Erprobung bis zur Serienreife zu verzichten. (S. 20/21)

Kasten 1: Mühlhausens Kritik an der Allgemeinen Didaktik

Auf dieser Grundlage begründet er, weshalb aus seiner Sicht Auflistungen und Kataloge von Merkmalen der Unterrichtsqualität sinnlos, ja sogar schädlich seien:

Viele didaktische Theorien wecken bei Lehrer/innen die trügerische Hoffnung, Unterricht könne planmäßig verlaufen, wenn sie sich nur an die – je nach Theorie anders lautenden – Vorgaben halten. Auch die Unterrichtsforschung bietet bis heute nicht mehr als eine je nach Systematik anders sortierte Auflistung mehr oder weniger bekannter, teils plausibler, teils unscharfer, teils kaum eindeutig fassbarer Merkmale nebst vieler Unter- und Nebenmerkmale, die angeblich guten Unterricht gewährleisten. Mit solchen Merkmalen können, wenn sie eindeutig zu definieren wären, bestenfalls notwendige Bedingungen für jeweils ganz bestimmte Auffassungen von gutem Unterricht beschrieben werden. Sie sind keinesfalls allgemeingültig für **den Guten Unterricht**, weil es eine solche allgemeingültige Auffassung nicht gibt. Zudem vermitteln solche Merkmalskataloge ein bloß statisches Wissen, das als solches nichts darüber verrät, auf welches Merkmal es mit welcher Akzentuierung in einer konkreten Unterrichtssituation ankommt. Solche Kataloge sind hilfreich, wenn es darum geht, ein gut abprüfbares Prüfungsthema für das Fach Pädagogik zu finden. Warum damit aber ausgerechnet jetzt eine Qualitätsverbesserung des Unterrichts erreicht werden könnte, wenn das in den letzten 40 Jahren nicht gelungen ist, bleibt das Geheimnis der Verfechter solcher Merkmalskataloge. M. E. ist es sinnvoller, zu untersuchen, welche Merkmale für einen Lehrer mit einem bestimmten Unterrichtskonzept bedeutsam sind (und welche nicht) und wie er diese situationsangemessen realisieren kann. (S.38/9)

Kasten 2: Mühlhausens Kritik an Merkmalskatalogen der Unterrichtsqualität

Die Position Mühlhausens ist originell und beschreibt einige Kritikpunkte, die man solchen Merkmalsauflistungen entgegenbringen kann, durchaus zutreffend. Allerdings liegt ihr ein grundlegendes Missverständnis zugrunde, was die Funktion von Unterrichtsqualitätsmerkmalen anbelangt. Diese sind nämlich gar nicht geeignet, Entscheidungen über das Verhalten in konkreten Unterrichtssituationen zu treffen. Das ist auch gar nicht ihr Anspruch. Schon allein dadurch, dass sie situationsübergreifend erfasst werden, ergibt sich, dass ihre Rolle nicht darin bestehen kann, das Handeln in konkreten Unterrichtssituationen zu leiten. Wie in Abschnitt 5 noch näher erläutert wird, stellen sie *Steuerungswissen* dar, das zur Reflexion und Analyse des Unterrichts dienen kann, nicht aber dazu, das Handeln in konkreten Unterrichtssituation zu lenken. Dieses Missverständnis wird auch an einer anderen Stelle seines Buches deutlich, wenn er in der Person des Lehrers liegende Faktoren anspricht, die das unterrichtliche Handeln zusätzlich zu den situativen Erfordernissen immer auch bestimmen und häufig verhindern, dass die aus den Unterrichtsmerkmalen resultierenden Vorgaben für guten Unterricht umgesetzt werden:

„Lehrer sind nicht immer gut vorbereitet, sondern treten mit ‘Mut zur Lücke’ an. Auch kommen Lehrer häufiger, als ihnen lieb ist, an die Grenzen ihrer Belastbarkeit, sind durch die Gleichzeitigkeit von Ereignissen überfordert, haben schlechte Laune – kurz: Sie stehen sich selbst im Wege. Diese durch den ‘subjektiven Faktor’ gegebenen Einschränkungen sind bedauerlich und kritikwürdig, aber nicht einfach abstellbar. Keine wie immer geartete ‘Professionalisierung’ der Lehrerbildung kann sie auslöschen. Versprechen wie die, dass ein Beachten von acht (oder 10, oder 15, ...?!) Merkmalen ‘guten Unterricht’ garantieren (z. B. Brophy 2000, Helmke 2003, Meyer 2004) nähren unerfüllbare Illusionen. Sie verhindern, dass Lehrer sich realistisch mit der Besonderheit ihrer Tätigkeit auseinandersetzen, um ihre Unzulänglichkeiten und Schwächen einschätzen zu lernen“ (S. 46).

So gesehen nutzt die Position Mühlhausens die Denkfigur des „Pappkameraden“: Es werden Aussagen und Behauptungen kritisiert, die niemand vertritt.

## 5. Reichweite und Grenzen der Aussagekraft von Merkmalen der Unterrichtsqualität: 20 Fragen und Antworten

Die in den vorigen Abschnitten formulierten Ausführungen sind Anlass, im Folgenden einige typische und immer wiederkehrende Kritiken und Fragen zur Bedeutung von Merkmalen der Unterrichtsqualität zu formulieren und zumindest ansatzweise zu beantworten. Gerade während der 2. Phase der Lehrerausbildung im Studienseminar bieten sich vielfältige Gelegenheiten, sich mit diesen Fragen näher auseinanderzusetzen. Abweichend vom traditionellen Format wissenschaftlicher Texte wird dazu der FAQ-Stil („frequently asked questions“ mit zugehörigen kurzen Antworten) gewählt.

### 1. Was ist eigentlich genau gemeint, wenn man von Merkmalen der Unterrichtsqualität spricht?

Merkmale der Unterrichtsqualität sind Aspekte, mittels derer sich Unterricht und Lehrer-Schüler-Interaktionen beschreiben lassen. Diese Merkmale sind nicht direkt beobachtbar, sondern Abstraktionen. Es handelt sich um „Konstrukte“, d. h. gedankliche Ordnungsleistungen, die sich auf bestimmte Regelmäßigkeiten im beobachtbaren Verhalten beziehen. Solche Beschreibungen können mit hohem oder niedrigem Allgemeinheitsgrad vorgenommen werden (zum einen allgemein, abstrakt und situationsübergreifend, zum anderen konkret, situations- und verhaltensbezogen). Wenn hier von *Merkmalen* gesprochen wird, sind Variablen gemeint, also Größen, die in verschiedenen *Ausprägungen* variieren.

## 2. Sollte man statt von Merkmalen der Unterrichtsqualität nicht besser von „Prinzipien“ sprechen?

Es ist nicht einfach, einen Begriff zu finden, der sprachlich gleichermaßen gut für alle in Tabelle 3 aufgeführten zehn Qualitätsmerkmale passt. Die Bezeichnung „Merkmal“ ist relativ voraussetzungslos, legt möglicherweise aber das Missverständnis nahe, es handle sich um ein singuläres, homogenes Merkmal. Deshalb ist die Bezeichnung „Bereich“ eigentlich passender; zumal auch in den Orientierungsrahmen zur Schulqualität von „Qualitätsbereichen“ die Rede ist. Von „Prinzipien“ spricht man im allgemeinen dann, wenn Untersuchungen wiederholt zu gleichen Ergebnissen und Schlussfolgerungen geführt haben und damit regelhafte Beziehungen nachgewiesen werden konnten (Woolfolk, 2008). Innerhalb der Qualitätsbereiche finden sich in der Tat zahlreiche empirisch gut erforschte Prinzipien mit jeweils spezifischem theoretischem Hintergrund.

## 3. Was bedeutet es genau, wenn man von Merkmalszusammenhängen spricht?

Zwei Merkmale hängen zusammen, wenn die Ausprägungen des einen Merkmals mehr oder weniger systematisch mit den Ausprägungen eines anderen Merkmals zusammenhängen („kovariieren“). So kann z. B. eine gute Verständlichkeit des Unterrichts mit einem hohen Lernerfolg der Schüler einher gehen; d. h. Lehrpersonen mit hoher Verständlichkeit haben auch Klassen mit hohem Lernerfolg und Lehrpersonen mit geringer Verständlichkeit haben Klassen mit niedrigem Lernerfolg. Derartige Merkmalszusammenhänge sind allerdings nicht deterministisch, d. h. man kann nicht davon ausgehen, dass bei einer bestimmten Ausprägung des einen Merkmals auch zwingend die exakte Ausprägung des anderen Merkmals resultieren würde. In den Sozialwissenschaften sind vorgefundene Zusammenhänge nahezu immer nur probabilistisch oder statistisch. Es handelt sich dabei um Korrelationen. Wenn man versucht, die Ausprägung eines der beiden Merkmale durch das andere vorherzusagen, so ist die Vorhersage immer mit einem mehr oder weniger großen Unsicherheitspielraum behaftet.

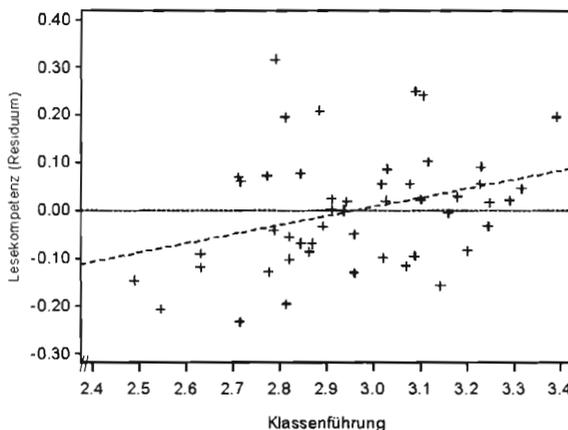


Abb. 2 Klassenführung und Kompetenzzuwachs im Lesen

Ein illustratives Beispiel zeigt Abbildung 2, nämlich den Zusammenhang zwischen dem Unterrichtsmerkmal „Klassenführung“ und dem Lernerfolg (hier Lesekompetenz) in Form eines Streudiagramms (Helmke, A., Helmke, T. Heyne, Hosenfeld, A., Schrader & Wagner, im Druck). Jede einzelne Klasse ist durch ein Kreuz repräsentiert, das die bei dieser Klasse vorfindbaren Ausprägungen des Unterrichtsmerkmals und des Lernerfolgs darstellt. Der gesamte „Punkteschwarm“ (hier: die Gesamtheit der Kreuze) lässt erkennen, dass es einen leicht positiven Zusammenhang gibt:

Je höher der Wert für die Klassenführung ist, umso höher tendenziell auch der erzielte Lernerfolg. Dies wird noch klarer, wenn man eine so genannte Regressionsgerade (die von links unten nach rechts oben verlaufende gestrichelte Gerade) durch den Punkteschwarm legt. Diese Gerade ist so gelegt, dass sie den Punkteschwarm nach einem

mathematischen Kriterium (Minimierung der Summe der quadrierten Abweichungen von dieser Gerade) bestmöglich beschreibt. Die Regressionsgerade umfasst die Werte des Lernerfolgs, die man für die verschiedenen Ausprägungen der Klassenführung jeweils erwarten würde. Wie man sieht, gibt es (a) Klassen, die nahezu den aufgrund ihrer Ausprägung des Unterrichtsmerkmals zu erwartenden Lernerfolg aufweisen, b) Klassen, die einen deutlich höheren Lernerfolg als erwartet erzielen und (c) Klassen, deren Lernerfolg deutlich niedriger als erwartet ausfällt.

Die Höhe des Zusammenhangs lässt sich durch einen Korrelationskoeffizienten kennzeichnen. Dieser ist eine Maßzahl für die Höhe des (linearen) Zusammenhangs zwischen zwei Merkmalen und sagt aus, wie gut sich der Punkteschwarm durch diese Gerade beschreiben lässt. Liegen die einzelnen Punkte alle in der Nähe der Geraden (und hat die Gerade eine von Null verschiedene Steigung), dann ist die Korrelation hoch. Streuen sie sehr weit um diese Gerade, dann ist die Korrelation niedrig. Im Extremfall liegen die Punkte alle auf der Geraden (Korrelation = 1). Im anderen Extremfall gibt es überhaupt keinen regelhaften Zusammenhang; die Gerade verläuft parallel zur x-Achse (Korrelation = 0). Im vorliegenden Fall hat der Korrelationskoeffizient den Wert von .62.

#### 4. Was bedeutet es, wenn man in der empirischen Unterrichtsforschung von erfolgreichem oder lernwirksamem Unterricht spricht?

In der empirischen Unterrichtsforschung spricht man dann von *erfolgreichem* Unterricht, wenn der Zusammenhang zwischen bestimmten Unterrichtsmerkmalen und dem Lernerfolg empirisch nachgewiesen, also durch empirische Daten belegt ist. Dem liegt eine als Prozess-Produkt-Ansatz bezeichnete Denk- und Vorgehensweise zugrunde. Prozessmerkmale sind Merkmale des Unterrichts, insbesondere des Lehrerverhaltens und der Lehrer-Schüler-Interaktion, die meistens durch Beobachtungsverfahren erfasst werden. Produkte sind bei den Schülern gemessene Ergebnisse wie der (meistens mit Tests erfasste) Leistungszuwachs der Schüler. Gelegentlich wurden auch andere Produktmaße wie der Anstieg der Lernfreude oder die Zunahme des Selbstvertrauens der Schüler erfasst. Im Prozess-Produkt-Ansatz werden dann die Prozessmerkmale mit den Produktmaßen in Beziehung gesetzt. Daraus resultieren Korrelationen, also Maße der Stärke des linearen Zusammenhangs zwischen Prozess (Unterrichtsmerkmal) und Produkt (Unterrichtserfolg), die bei *einzelnen* Merkmalen oft nur schwach positiv sind.

Häufig werden diese Korrelationen vorschnell kausal interpretiert, d. h. die Unterrichtsmerkmale werden als *Ursache* für die Veränderungen auf Schülerseite gesehen. Diese Interpretation ist zwar oft plausibel, aber nicht zwingend und gelegentlich vollkommen ungerechtfertigt. Eine Korrelation zwischen zwei Merkmalen lässt nämlich immer drei Interpretationsmöglichkeiten zu: (1) der Unterricht beeinflusst die Leistung der Schüler; (2) der Leistungsstand der Schüler beeinflusst den Unterricht (z. B. können in einer leistungsstarken Klassen anspruchsvollere Methoden eingesetzt werden); und (3) der Zusammenhang zwischen Unterricht und Leistung hängt von einem dritten Merkmal ab. Beispielsweise stehen in einer gut ausgestatteten Schule Lehrern und Schüler bessere Materialien zur Verfügung, die sowohl den Unterricht als auch das Lernen günstig beeinflussen – ohne dass es dann tatsächlich der verbesserte Unterricht sein muss, der zum besseren Lernen führt.

Statistische Beziehungen kommen deshalb zustande, weil ein bestimmtes Ergebnis, z. B. der Lernerfolg einer Schulklasse, nicht nur von einem einzigen Faktor, sondern immer von einer Vielzahl von Faktoren abhängt, die aber bei einer einfachen Korrelation nicht beachtet werden. Es gibt verschiedene Modelle des Lernerfolgs oder Schulleistung, die verdeutlichen, wie komplex das Wirkungsgefüge der den Lernerfolg beeinflussenden Fak-

toren tatsächlich ist. Insofern sind einfache Korrelationen immer nur ein grober Anhaltspunkt dafür, wie bedeutsam ein Unterrichtsmerkmal für den Lernerfolg ist und welche Rolle es spielt. Sie können aber wichtige Hinweise auf mögliche Stellschrauben liefern, an denen man drehen kann, um den Unterricht zu verbessern.

Abbildung 3 (Helmke, 2009, S. 30) verdeutlicht das Muster verschiedener Determinanten der Schulleistung:

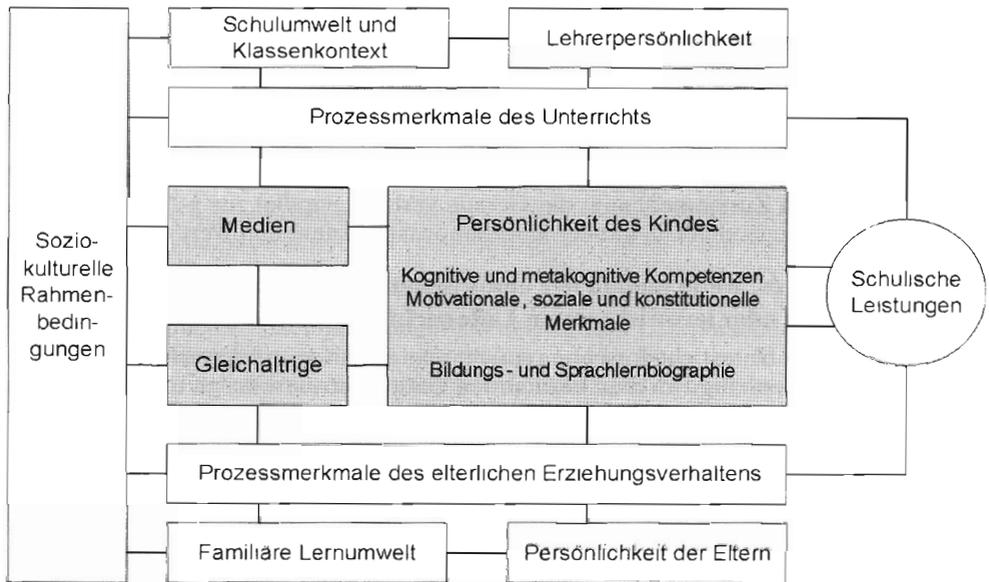


Abb. 3 Makromodell der Schulleistungsdeterminanten (Helmke, 2009, S. 30)

Das Modell zeigt zunächst, dass Prozessmerkmale des Unterrichts nur einer der Bedingungsfaktoren sind, die schulische Leistungen beeinflussen, und dass daneben noch eine Reihe anderer Einflussgrößen wirksam sind. Würde man diese ebenfalls erfassen und berücksichtigen, käme man zu sehr viel genaueren und besseren Vorhersagen der Leistung. Das Modell zeigt ebenfalls, dass manche Faktoren relativ direkt, andere nur sehr indirekt auf die Leistung Einfluss nehmen. So haben Merkmale der Lehrerpersönlichkeit keinen direkten Einfluss auf die Leistung, sondern nur einen indirekten: durch die Gestaltung von Lehr-Lern-Prozessen (Unterrichtsqualität) und durch ihre Persönlichkeit (Lernen am Modell, Beobachtungslernen).

### 5. Sind die meisten Merkmalkataloge nicht insofern anachronistisch, als sie ausschließlich von der Lehrerperspektive ausgehen?

Auch dies ist ein weitverbreitetes Missverständnis. Es kann gar keine Rede davon sein, dass sich die Merkmale ausschließlich auf Lehrpersonen beziehen. An der Realisierung oder Blockierung aller Merkmale können Lehrer wie Schüler gleichermaßen beteiligt sein. Zunächst hängt das Zustandekommen einer bestimmten Merkmalsausprägung nicht allein von der Lehrperson und ihren Kompetenzen, sondern auch von der Klassenzusammensetzung, z. B. dem Leistungsniveau oder dem Bildungshintergrund der Klasse ab. So

kann die (von den Schülern oder einem Beobachter beurteilte) Verständlichkeit unterschiedlich sein, je nachdem, ob es sich um eine leistungsstarke oder leistungsschwache Klasse handelt.

Außerdem lassen sich Kriterien wie Klarheit, Strukturiertheit, Verständlichkeit oder Korrektheit auf Schüleräußerungen ebenso anwenden wie auf Lehreräußerungen. Auch Merkmale wie lernförderliches Klima (z. B. wie Schüler auf Fehler ihrer Mitschüler reagieren) betreffen in hohem Maße das Schülerverhalten. Von Nuthalls (1997) Videostudien wissen wir beispielsweise, dass während einer durchschnittlichen Schulwoche mit unterschiedlichen Sozialformen der sprachliche Input der Peers – jenseits des „offiziellen“ Lehrer-Schüler-Gesprächs – bei weitem dominiert.

Grundsätzlich muss man sich auch klar machen, dass die lernförderliche „Wirkung“ bestimmter Unterrichtsmerkmale nur verstanden werden kann, wenn man den Schüler einbezieht. Die Wirkung resultiert daraus, dass auf Schülerseite bestimmte Lernaktivitäten in Gang gesetzt werden. Unterrichtsmerkmale führen nicht zwangsläufig und automatisch zu bestimmten Lernergebnissen, sondern nur dadurch, dass Schüler das Lehrerverhalten in bestimmter Weise wahrnehmen, interpretieren und für eigene Lernanstrengungen nutzen. Ob und in welchem Maße sie das tun, hängt wiederum davon ab, ob bestimmte Bedingungen wie etwa ausreichende Vorkenntnisse vorliegen. Dieser Sichtweise wird das vom Erstautor (Helmke, 2009) entwickelte Angebots-Nutzungs-Modell gerecht (vgl. Abb. 4).

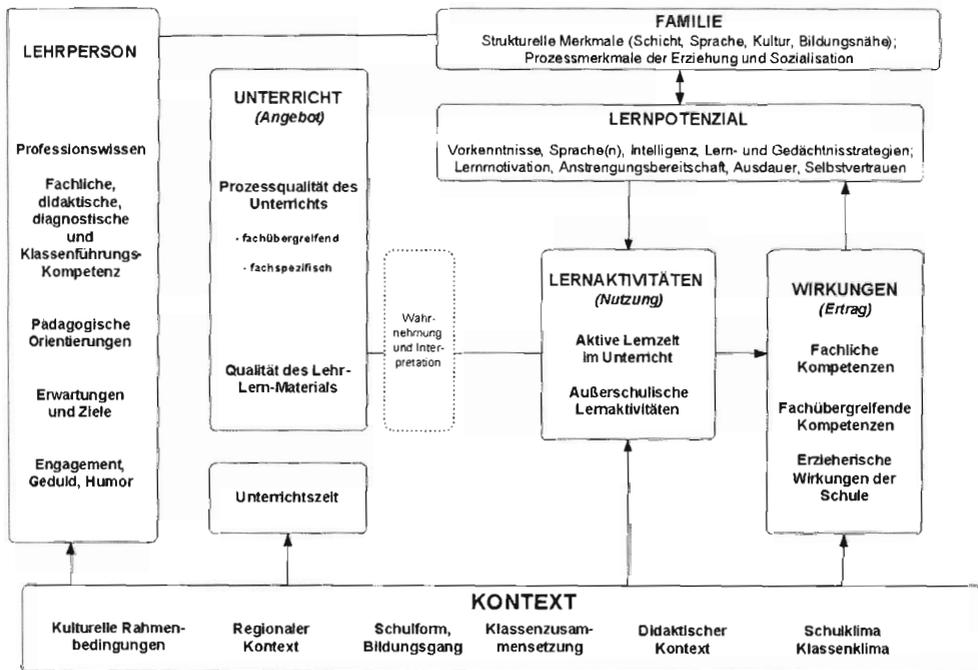


Abb. 4 Angebots-Nutzungs-Modell des Unterrichts (Helmke, 2009, S. 73)

Diesem verbreiteten Modell zufolge ist der Unterricht immer nur ein *Angebot*, dass von den Schülern in geeigneter Weise genutzt werden muss, wenn bestimmte Lernergebnisse

(fachliche oder fachübergreifende Kompetenzen, erzieherische Wirkungen) resultieren sollen. Die *Nutzung* besteht im Einsatz unterrichtlicher oder außerunterrichtlicher Lernaktivitäten, also von kognitiven Aktivitäten, die Grundlage von Lernprozessen sind (Kodieren, Verknüpfen, Vergleichen u. v. m.). Nur wenn diese *Lernaktivitäten* stattfinden, können die angestrebten Wirkungen eintreten, und zwar unabhängig davon, ob der Unterricht eher offen angelegt oder stark lehrergesteuert ist. Neben dem Unterrichtsangebot nehmen auch individuelle *Schülermerkmale* (Lernpotenzial wie Intelligenz, Vorkenntnisse usw.) sowie der *Kontext* Einfluss auf die Lernaktivitäten. Auch hier zeigt sich, dass Personenmerkmale der Lehrkraft nur indirekt wirksam sind. Das Angebots-Nutzungs-Modell verdeutlicht die entscheidende Rolle, die der Schüler selbst im unterrichtlichen Lernprozess spielt und zeigt damit auch, dass sich Unterrichtsangebote nicht eins-zu-eins in Lernerfolge umsetzen lassen.

## **6. Sind Aussagen zur Unterrichtsqualität, die fachspezifische und fachdidaktische Aspekte ausklammern, nicht unvollständig?**

Ja, eine angemessene Beurteilung der Unterrichtsqualität erfordert in der Tat die Berücksichtigung von beidem: fachübergreifender und fachspezifischer Merkmale. Darüber ist sich die Unterrichtsforschung heute vollkommen einig. Alle großen Unterrichtsforschungsprojekte des letzten Jahrzehntes sind deswegen Kooperationsvorhaben von Pädagogischen Psychologen bzw. Empirischen Erziehungswissenschaftlern einerseits und Fachdidaktikern andererseits. Einen Lehrstoff zu verstehen bedeutet für den Lernenden letztlich immer, Wissensstrukturen aufzubauen, die die im Lehrstoff bestehenden Strukturen hinreichend abbilden. Die Struktur des Lehrstoffs legt also in einem gewissen Grade bereits fest, wie die spätere Wissensstruktur aussehen wird, wenn erfolgreich gelernt wurde. Da schulische Leistungen an bestimmte Fächer gebunden sind, sind von den *Fachdidaktiken* bereitgestellte Erkenntnisse zur Frage, wie fachliche Lehrstoffe organisiert sind und welche Anforderungen sie an das Lehren und Lernen stellen, von großer Bedeutung. Auch innerhalb der Psychologie ist in den letzten Jahrzehnten vielfach belegt worden, welche zentrale Rolle das Vorwissen für den Erwerb neuen Wissens spielt. Bei umfangreichen und gehaltvollen Lernstoffen reichen allgemeine Fähigkeiten, Strategien oder Schlüsselqualifikationen für den Kompetenzerwerb nicht aus; sie sind insbesondere nicht in der Lage, mangelndes Vorwissen zu kompensieren (vgl. Hasselhorn & Gold, 2006, S. 81f.). Ganz offensichtlich ist das bei hierarchisch aufgebauten Wissensgebieten wie dem der Mathematik, bei denen ohne hinreichende Vorkenntnisse weitere Lernschritte kaum möglich sind.

Gegen den Einsatz fachunspezifischer Strategien bei der Wissensvermittlung wird gelegentlich zu Recht der Vorwurf des „Strickens ohne Wolle“ erhoben. Dieser Vorwurf trifft weder für die Allgemeine Didaktik noch für die Empirische Unterrichtsforschung zu, obwohl beide ebenfalls fachübergreifende Zugänge darstellen. Vielmehr ist es nach übereinstimmender Meinung in der internationalen Unterrichtsforschung lohnenswert, ja unabdingbar, über Fächergrenzen und Altersstufen hinweg nach allgemeinen („generischen“) Merkmalen der Unterrichtsqualität zu suchen. Merkmale wie Klassenführung, kognitive Aktivierung, Verständlichkeit oder lernförderliches Klassenklima sind in unterschiedlichen Fächern in sehr ähnlicher Weise bedeutsam, wie viele Befunde der Unterrichtsforschung belegen. Diesem Aspekt wurde in der Fachdidaktik in der Vergangenheit zu Gunsten einer ausschließlich fachspezifischen Sichtweise oft zu wenig Aufmerksamkeit geschenkt.

## 7. Hängt die Beurteilung von Merkmalen der Unterrichtsqualität nicht von der Perspektive des Beurteilers ab?

Ja, dies gehört zu den immer wieder gefundenen Ergebnissen der Unterrichtsforschung: Lässt man ein-und-denselben Unterricht von verschiedenen Schüler/innen beurteilen – egal ob es sich um ein kumulatives Urteil über ein Schuljahr oder über eine konkrete Unterrichtsstunde handelt –, dann finden sich durchweg klasseninterne Streuungen bei den jeweiligen Items, d. h. die einzelne Schüler/innen geben mehr oder weniger voneinander abweichende Urteile ab. Diese Streuungen können darüber hinaus von Item zu Item unterschiedlich ausfallen. Das heißt: Bei manchen Merkmalen dominiert der Konsens (z. B. bei leicht beobacht- und objektivierbaren Sachverhalten), bei anderen Merkmalen zeigt sich starker Dissens innerhalb der Klasse.

Was für Unterschiede zwischen Schülern einer Klasse gilt, trifft erst recht für die Perspektiven verschiedener Urteiler zu, z. B. wenn man die gleiche Unterrichtsstunde sowohl aus Sicht der Klasse (Bildung eines Durchschnittswertes) als auch der unterrichtenden Lehrperson und von Kollegen (via Hospitation) beurteilen lässt. Besonders gründlich ist die Frage der Perspektivenabhängigkeit von Clausen (2002) und kürzlich von Wagner (2008) untersucht worden. Allerdings gibt es zwischen den unterschiedlichen Perspektiven – wenn auch nicht unbedingt hohe – Zusammenhänge, die zeigen, dass es so etwas wie einen gemeinsamen Kern gibt. In der Unterrichtsforschung wird dies insofern berücksichtigt, dass Erhebungen möglichst *mehrperspektivisch* angelegt sind (also Schülerwahrnehmungen, Lehrereinschätzungen und Beobachtungen durch externe Beobachter nach Möglichkeit kombiniert werden).

Für die interne Evaluation in der Schulpraxis ist gerade die Schülerperspektive von nicht zu unterschätzender Bedeutung. Klassenweise aggregierte (d. h. gemittelte) Schülerurteile liefern in den meisten Fällen durchaus zutreffende und nützliche Angaben. Da solche Angaben leicht zu erheben und schnell auszuwerten sind, sind sie gerade für die Unterrichtsentwicklung ein praktikabel einsetzbares Instrument, siehe Helmke, Piskol, Piskovsky & Wagner (im Druck) und den letzten Abschnitt dieses Artikels. Sehr viel aufwändiger, aber für Unterrichtsentwicklung ebenfalls sehr wertvoll und in der schulischen Praxis noch viel zu selten genutzt ist die Videografie des Unterrichts. Wenn der aufgezeichnete Unterricht beobachtet und ggf. gemeinsam mit Kollegen/innen analysiert wird, kann das viele wertvolle Einsichten vermitteln und Prozesse der Unterrichtsentwicklung in Gang setzen.

## 8. Basieren die Merkmale der Unterrichtsqualität auf einer Theorie?

Nein, zumindest nicht im strengen Sinne, wenn man unter einer Theorie ein logisch geordnetes System von Aussagen versteht, die auf wenige Grundannahmen und Grundbegriffe zurückführbar sind. Solche einheitlichen Theorien sind in Psychologie und Pädagogik ohnehin nur selten. Die Merkmale der Unterrichtsqualität sind zwar nicht aus einer kohärenten pädagogisch-psychologischen oder didaktischen Theorie ableitbar; ihre Bedeutung für den Lernerfolg und andere Zielkriterien ist aber durch umfangreiche Forschungsbemühungen empirisch gut abgesichert. Diese Forschungen wurden weniger aus theoretischen Gründen, sondern mit einer eher praktischen Zielsetzung betrieben, nämlich zu prüfen, ob und in welchem Maße Unterrichtsmerkmale mit dem Lernerfolg zusammenhängen (vgl. Frage 4).

In den in Abschnitt 2 dargestellten Übersichten werden in pragmatischer Weise Prinzipien und Qualitätsbereiche zusammengestellt, die empirisch mehr oder weniger gut bestätigt sind und über deren Relevanz daher aus wissenschaftlicher oder bildungspolitischer Perspektive wenig Zweifel bestehen. Dieser mittlerweile recht umfängliche Wissenskorpus stellt zwar keine Theorie im strengen Sinne dar, sondern ist eklektischer Natur. Er ist des-

wegen aber nicht theorieelos, denn *innerhalb* jedes Qualitätsbereichs finden sich bereicherspezifische Theorien, etwa aus der Pädagogischen Psychologie und der Kognitionspsychologie, die z. B. Fragen des Lernens und Gedächtnisses, des Selbstkonzeptes oder der Motivation zum Gegenstand haben; und viele der gefundenen Zusammenhänge lassen sich vor dem Hintergrund solcher Theorien sehr gut interpretieren.

### 9. Welchen Sinn haben solche Qualitätsmerkmale im konkreten Schulalltag angesichts der Unberechenbarkeit des Unterrichtsverlaufs?

Die Kenntnis dieser Prinzipien und das Verständnis ihrer Wirkungsweise ist eine notwendige, aber nicht hinreichende Voraussetzung für erfolgreichen Unterricht; es handelt sich nicht um Unterrichtsrezepte, sondern um *Steuerungswissen*. Dieses kann genutzt werden, um Unterricht zu reflektieren und zu analysieren. Anders als z. B. von Mühlhausen (2008) unterstellt, eignen sich die Qualitätsmerkmale nicht für die Planung und Gestaltung einer konkreten Unterrichtsstunde, sondern eher für langfristige Weichenstellungen und Schwerpunktsetzungen im Rahmen der Unterrichtsentwicklung. Bei diesen Bemühungen geht es nicht darum, eine einzelne Unterrichtsstunde zu optimieren, sondern darum, allgemeine Gewohnheiten und Routinen zu verändern.

Unterricht ist ein hochkomplexes Geschehen. Der amerikanische Unterrichtsforscher Doyle (1979) hat das Unterrichtsgeschehen durch *Multidimensionalität* (multidimensionality; Vielzahl von Ereignissen von Wahrnehmungen), *Gleichzeitigkeit* (simultaneity; gleichzeitiges Auftreten vieler Ereignisse), *Unaufschiebbarkeit* (immediacy; beschränkte Möglichkeit, Reaktionen aufzuschieben), *Unvorhersagbarkeit* (unpredictability; begrenzte Vorhersagbarkeit vieler Abläufe) und *Relevanz für künftiges Handeln* (history; Präzedenzfallcharakter vieler Entscheidungen) charakterisiert. Diese Komplexität kann von der Lehrperson zumindest mittel- und langfristige nur dadurch bewältigt werden, dass Verhaltensroutinen aufgebaut und Vorgehensweisen automatisiert werden, durch die eine kognitive Entlastung geschaffen wird. Solche Maßnahmen können nur längerfristig aufgebaut und damit auch nur längerfristig verändert werden. Genau dazu ist Steuerungswissen wichtig.

Voraussetzung für erfolgreiche Maßnahmen der Unterrichtsentwicklung ist eine *empirische Orientierung*, d. h. die Bereitschaft und Fähigkeit, den eigenen Unterricht auf den Prüfstand zu stellen und eine datengestützte Bestandsaufnahme des Unterrichts und des Lernerfolgs vorzunehmen. Diese umfasst die regelmäßige Diagnose sowohl der Lernvoraussetzungen und des Leistungsstandes der Schülerinnen und Schüler als auch (z. B. durch kollegiale/virtuelle Hospitation oder durch Schülerfeedback) des eigenen Unterrichts. Die so entstehenden, empirisch basierten (oder „evidenzbasierten“) Bestandsaufnahmen sind Anlässe, sich gezielt mit bestimmten Merkmalen der Unterrichtsqualität zu beschäftigen und mittelfristig Schwerpunkte zu setzen – z. B. einem Qualitätsbereich (z. B. der Motivierung) besondere Aufmerksamkeit zuzuwenden und andere, momentan weniger relevante Qualitätsbereiche vorübergehend nur am Rande zu verfolgen.

### 10. Heißt „guter Unterricht“, dass alle Merkmale maximal ausgeprägt sein müssen?

„Guter Unterricht“ im Sinne eines nachweislich erfolgreichen Unterrichts ist keineswegs identisch mit einer optimalen Ausprägung aller Merkmale. Dazu kommt: Nicht immer ist die *maximale* Ausprägung eines Merkmals optimal. Oft sind es eher *mittlere* Ausprägungen, also weder eine zu hohe noch eine zu niedrige Ausprägung. So ist z. B. öfter eher ein mittlerer Grad an Strukturierung günstig, die den Schülern nicht alle eigenen Strukturierungsbemühungen abnimmt, sie auf der anderen Seite aber nicht orientierungslos lässt.

Welches Maß der Strukturierung nötig ist, hängt ganz entscheidend von den Lernvoraussetzungen der Schüler ab.

Es kann auch durchaus so sein, dass bis zu einem gewissen Grad Defizite in einem Bereich durch Stärken in anderen Bereichen kompensierbar sind. Wann immer in der empirischen Unterrichtsforschung nicht nur auf Aggregatebene operiert wurde (Zusammenfassung vieler erfolgreicher Klassen zu einer Gruppe, deren Durchschnittswert in die weiteren Analysen eingeht), sondern zusätzlich Unterrichtsprofile der einzelnen Klassen beschrieben wurden, wurde gezeigt, dass es sehr unterschiedliche Wege zum Erfolg gibt, zuletzt von Mayr (2008). So auch in der Grundschulstudie SCHOLASTIK (Weinert & Helmke, 1997), wie Abbildung 5 zeigt: Nachweislich erfolgreiche Klassen (überdurchschnittliche Entwicklung der Fachleistungen in Mathematik und im Lesen und zugleich überdurchschnittliche Entwicklung der Lernfreude) weisen sehr unterschiedliche Unterrichtsprofile auf. Zwar befinden sich die Profile erwartungsgemäß auf der „positiven“ rechten Seite, d. h. die Variablen der Unterrichtsqualität haben meistens überdurchschnittliche Ausprägungen. In einzelnen Fällen finden sich jedoch auch unterdurchschnittliche Werte. Bemerkenswert ist auch, dass bei zwei der seinerzeit geprüften Qualitätsmerkmalen *alle* Klassen überdurchschnittliche Werte aufweisen, also alle rechts von der Durchschnittslinie liegen: Klarheit/Strukturiertheit und Klassenführung. Das unterstreicht die besondere Rolle dieser beiden Unterrichtsmerkmale, insbesondere die Bedeutung der Klassenführung als eines für erfolgreichen Unterricht notwendigen, wenn auch nicht hinreichenden Merkmals (s. u.).

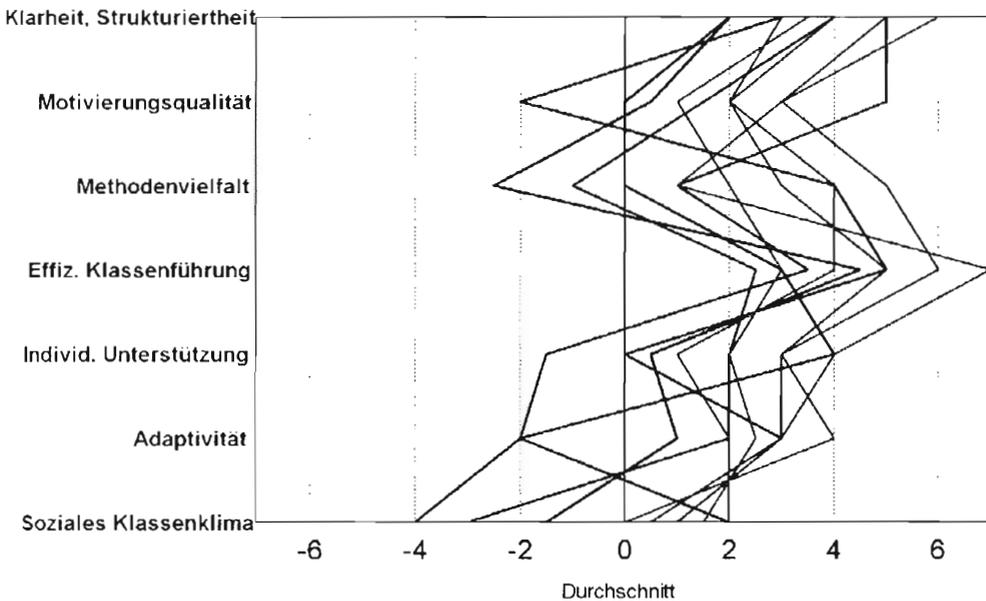


Abb. 5 Unterrichtsprofile von Optimalklassen (Projekt SCHOLASTIK)

## 11. Sind die Merkmale der Unterrichtsqualität miteinander vereinbar, oder widersprechen sie sich?

*Prinzipiell* sind die meisten – aber nicht alle – Merkmale der Unterrichtsqualität miteinander vereinbar, das heißt: Eine hohe bzw. günstige Ausprägung eines Merkmals ist nicht notwendig mit einer niedrigen Ausprägung eines anderen Merkmals verknüpft. So harmonisiert ein hohes Maß an Motivierung im allgemeinen mit einer hohen Ausprägung von Schülerorientierung und lernförderlichem Klima. Es gibt aber auch eingebaute Inkompatibilitäten, beispielsweise zwischen der Zeitnutzung (für Lernen und Unterricht) und der bewussten, pädagogisch gewollten – Zeit beanspruchenden – Investition von Unterrichtszeit in extracurriculare Aktivitäten, die dem lernförderlichen Klima oder auch der Verdeutlichung von Regeln dienen können. Ähnliches gilt für die (beschränkte) Kompatibilität zwischen hoher Strukturiertheit und hoher Ausprägung selbstständigen Lernens. Dieser Gesichtspunkt verdeutlicht nochmals, dass das Ziel der simultanen *Maximierung* aller Qualitätskriterien unangemessen ist, da es auf die optimale Platzierung, Dosierung und – auf das Timing ankommt – entsprechend den jeweiligen Lernvoraussetzungen der Schüler und dem fachlichen Kontext.

In einer *konkreten Unterrichtsstunde* ist die gleichzeitige Berücksichtigung geschweige denn Maximierung unterschiedlicher Merkmale schon aus Gründen begrenzter zeitlicher Ressourcen ausgeschlossen: Welche der zehn Merkmale in einer konkreten Unterrichtsstunde besondere Aufmerksamkeit erfordern, welche vorübergehend ignoriert oder vernachlässigt werden können, dies hängt von der Situation, insbesondere vom unterrichteten Fach und der Klassenstufe, vor allem aber von der Klassenzusammensetzung und dem aktuellen Niveau fachlicher und überfachlicher Kompetenzen ab. Hier liegt die Quelle für ein weitverbreitetes Missverständnis, nicht nur in Lehrerkreisen, sondern auch in der Wissenschaft.

## 12. Wie viele Merkmale der fachübergreifenden Unterrichtsqualität gibt es?

Es lässt sich keine eindeutige Zahl benennen. Das hat damit zu tun, dass Merkmale der Unterrichtsqualität keine in der Realität einfach vorfindbaren Gegebenheiten sind, sondern gedankliche Ordnungsleistungen (vgl. Frage 1) darstellen, die sich auf einem je nach Erkenntnisinteresse und Zielsetzung unterschiedlichen Auflösungslevel bewegen können. Hinzu kommt, dass die verschiedenen Merkmalsbereiche nicht nur facettenreich sind, sondern auch unterschiedliche Hierarchieebenen aufweisen, die von allgemeinen situationsübergreifenden Aspekten bis hin zu konkreten Verhaltensweisen reichen können (vgl. Frage 1). Allgemeine Merkmale sind über verschiedene Hierarchieebenen mit konkreten Verhaltensweisen verbunden, die dann als Indikatoren für die allgemeinen Merkmale fungieren. Die entsprechenden Merkmalsbeschreibungen sind theoretisch mehr oder weniger gut ausgearbeitet.

So können bei einem allgemeinen Merkmal wie Effizienz der Klassenführung Facetten wie die Qualität eines Regelsystems, die Organisation und Flüssigkeit der Abläufe oder der Umgang mit Disziplinproblemen unterschieden werden. Auf einer konkreteren und verhaltensnäheren Ebene wären dann verschiedene Arten des Umgangs mit Disziplinproblemen denkbar wie Überreaktionen, zeitliche Fehleinschätzungen oder Adressierungsfehler (die falschen Schüler sanktionieren), die man bis auf die Ebene einzelner Verhaltensweisen in spezifischen Situationen weiterverfolgen könnte. Je nachdem, auf welcher Ebene man sich bewegt, kommt man zu einer unterschiedlichen Anzahl von Merkmalen. Dabei kommt hinzu, dass sich Konstrukte bis zu einem gewissen Grad überlappen können. Z. B. könnte sich ein verhaltensnahes Merkmal wie Pünktlichkeit sowohl als Indikator für Klassenführung als auch für Zeitnutzung ansehen lassen.

In gewissem Sinne ist somit jede Klassifikation ein stückweit arbiträr. Mit entsprechender Begründung könnte man daher meistens auch eine größere oder geringere Anzahl von Qualitätsbereichen vertreten, indem man theoretisch und empirisch verwandte Merkmale zu umfassenderen Bereichen zusammenfasst. Den in Tabelle 5 aufgeführten Merkmalen 2–4 ist beispielsweise gemeinsam, dass sie sich direkt auf die Förderung der *Informationsverarbeitung* beziehen, die Merkmale 5–7 richten sich dagegen primär auf die Förderung der *Lernbereitschaft*, und die Merkmale 9–10 tragen dem Sachverhalt der *Unterschiedlichkeit von Bildungszielen, fachlichen Inhalten und individuellen Lernvoraussetzungen* Rechnung. Natürlich heißt das nicht, dass Unterrichtsmerkmale völlig beliebig sind. Man muss vielmehr genaue Regeln festlegen, mit denen die Merkmale eingegrenzt, voneinander abgegrenzt und mit beobachtbaren Indikatoren verknüpft werden. Diese Konstruktion muss hinreichend präzise sein, damit etwa verschiedene Beobachter oder Urteiler zu den gleichen oder zumindest sehr ähnlichen Ergebnissen kommen können. Nur dann ist die Merkmalsbildung hinreichend fundiert, um Klassen oder Lehrkräfte vergleichen zu können.

### 13. Welche Merkmale der Unterrichtsqualität sind am wichtigsten, welche weniger wichtig?

Das hängt vom betrachteten Bildungsziel ab. Je nach dem, ob es um die Förderung fachlicher Kompetenzen, überfachlicher Schlüsselkompetenzen oder um erzieherische Wirkungen der Schule geht (Sozialisierungseffekte), haben Merkmale der Unterrichtsqualität ein unterschiedliches Gewicht. Die empirische Unterrichtsforschung hat allerdings gezeigt, dass es einige Schlüsselmerkmale der Unterrichtsqualität gibt, die sich in Forschungen immer wieder als bedeutsam herausgestellt haben. Dies sind Klassenführung und Zeitnutzung, Schülerorientierung, lernförderliches Klima und Motivierung, Klarheit/Verständlichkeit/Strukturiertheit, Übung/Wiederholung/Konsolidierung, Aktivierung, Umgang mit Heterogenität (der Lernvoraussetzungen) und angemessene Methodenvielfalt sowie Wirkungs- und Kompetenzorientierung (Helmke, Göllner, Kleinbus, Schrader & Wagner, 2008).

Vor allem der *Klassenführung* kommt eine zentrale Bedeutung zu. Während Klassenführung in Deutschland häufig unterschätzt wird, hat sie in den USA (Classroom Management) auch in der Ausbildung einen festen Platz gefunden, wie eine Reihe von Handbüchern belegen. Eine effiziente Klassenführung scheint den Status einer notwendigen Voraussetzung für guten Unterricht zu haben: Ohne eine geordnete Klassenführung ist ein lernwirksamer Unterricht kaum möglich. Wenn es nicht gelingt, einen geordneten Rahmen für die Unterrichtsaktivitäten herzustellen, sind anspruchsvolle didaktische Maßnahmen kaum realisierbar. Auf der anderen Seite garantiert eine geordnete Klassenführung aber nicht bereits, dass ein anspruchsvoller kognitiv aktivierender und lernförderlicher Unterricht resultiert. Natürlich gibt es hier vielfältige Wechselbeziehungen mit anderen Unterrichtsqualitätsmerkmalen. So erleichtert ein gut vorbereiteter und flüssig durchgeführter, gut auf die Vorkenntnisse und Interessen der Schüler abgestimmter Unterricht die Aufgaben der Klassenführung, wohingegen ein schlecht vorbereiteter, mit Leerlauf und Stockungen verbundener, die Schüler über- oder unterfordernder Unterricht es erschwert, einen geordneten Unterrichtsablauf zu realisieren. Eine wenig effiziente Klassenführung ist häufig ein Anlass für das Auftreten von Disziplinstörungen, die oft aufgrund von Leerlauf, Langeweile oder geringer Aufmerksamkeit entstehen. Gute Klassenführung ist mehr als ein effektiver Umgang mit Disziplinschwierigkeiten. Dies ist eher ein Nebenaspekt. Entscheidend ist vielmehr, dass der Unterricht so angelegt wird, dass Störungen möglichst gar nicht erst auftreten. Dann verringert sich auch die Wahrscheinlichkeit, dass Disziplinschwierigkeiten entstehen.

Dass ein positives, vertrauensvolles und der Motivation zuträgliches Unterrichtsklima – um nur eines der oben genannten Merkmale noch hervorzuheben – ebenfalls eine große Bedeutung hat und praktisch immer wichtig ist, liegt auf der Hand. Wie Untersuchungen zeigen, bedeutet ein freundliches und vertrauensvolles Unterrichtsklima keinesfalls den Verzicht auf Leistungsanforderungen. Es ist vielmehr eine wichtige Grundlage für die Lernbereitschaft. Ein Unterricht, der die Schüler anspricht, sie motiviert, ihnen Kompetenzen vermittelt und Erfolgserlebnisse ermöglicht, hat deshalb auch Rückwirkungen auf das Unterrichtsklima. Diese knappen Ausführungen machen im Grunde auch deutlich, dass es weniger darauf ankommt, einzelne Unterrichtsmerkmale für sich zu betrachten, sondern ihr Zusammenwirken und das gesamte Beziehungsgefüge in den Blick zu nehmen. Man spricht deshalb auch von der „Choreographie“ und der „Orchestrierung“ (Oser & Baeriswyl, 2002). Letztlich kommt auf diese Weise auch wieder die häufig geforderte Ganzheitlichkeit der Betrachtung von Unterricht ins Spiel, jetzt aber gewissermaßen auf einer höheren Stufe, bei der die einzelnen Aspekte, die zu dieser Ganzheitlichkeit beitragen, sehr viel klarer zutage treten.

#### **14. Hängen die Merkmale der Unterrichtsqualität zusammen, oder sind sie unabhängig voneinander?**

Wie bereits erwähnt, sind die einzelnen Merkmale meistens keine homogenen Konstrukte, sondern Bezeichnungen für Bereiche („Schubladen“), die inhaltlich verwandte, aber durchaus heterogene Facetten und Forschungstraditionen enthalten. In der pädagogisch-psychologischen Forschung versucht man häufig, voneinander unabhängige Faktoren oder Dimensionen zu finden, die miteinander nicht zusammenhängen, deren Interkorrelation also bei Null liegt. Miteinander zusammenhängende Merkmale sind insofern redundant, als sich die Ausprägung eines Merkmals bereits ein Stück weit aus der Kenntnis des anderen vorhersagen lassen. Das widerspricht Prinzipien wie Sparsamkeit und Ökonomie, ist aber auch aus praktischer Sicht wichtig. Die zeitlichen und materiellen Ressourcen sind schließlich immer beschränkt, so dass man vermeiden möchte, Unterrichtsmerkmale zu erheben, über die man ansatzweise bereits durch die Kenntnis anderer Merkmale informiert ist. Die Interkorrelationen zwischen verschiedenen Unterrichtsmerkmalen liefern immer auch einen Hinweis darauf, wie differenziert der Unterricht wahrgenommen wird, ob eher als global und undifferenziert (hohe Korrelationen) oder eher differenziert und vielschichtig (niedrige Korrelationen). Dieses Problem stellt sich besonders bei den in der Unterrichtsforschung häufig verwendeten Schülerurteilen: Wenn die einzelnen Merkmale hoch korrelieren, deutet das darauf hin, dass sich die Beurteilung möglicherweise an globalen Eindrücken wie Sympathie und Antipathie gegenüber dem Lehrer orientiert.

Das statistische Werkzeug, um die Korrelationen zwischen verschiedenen Merkmalen zu beschreiben und auf grundlegendere Dimensionen zurückzuführen, sind so genannte Faktorenanalysen. Hoch miteinander korrelierende Merkmale lassen sich gewissermaßen zu Merkmalsbündeln oder Dimensionen zusammenfassen, und es lässt sich prüfen, ob diese Merkmalsbündel ihrerseits miteinander zusammenhängen. Der Versuch, unabhängige Dimensionen zu finden, ist in der Forschung zur Unterrichtsqualität bislang jedoch lediglich im Fach Mathematik ansatzweise gelungen. So gehen Klieme und Baumert von folgenden Basisdimensionen der Unterrichtsqualität aus: schülerorientiertes Unterrichtsklima, kognitive Aktivierung und effiziente Klassenführung (Klieme, 2002). Die Ergebnisse der DESI-Studie zeigten jedoch für die bislang noch kaum untersuchten Fächer Deutsch und Englisch völlig andere Strukturen (Klieme, Jude, Rauch, Ehlers, Helmke, Eichler, Thomé & Willenberg, 2008; Helmke, A., Helmke, T., Schrader, Wagner, Klieme, Nold & Schröder, 2008). Zwar konnten hier separate Dimensionen ermittelt werden; gleichwohl fanden sich substantielle korrelative Zusammenhänge *zwischen* den Dimensionen.

### 15. Welche Merkmale sind empirisch gut erforscht, welche weniger?

Einige Prinzipien sind empirisch sehr gut fundiert, insbesondere Klassenführung, Klarheit und Strukturiertheit, Konsolidierung und Sicherung, Aktivierung und Motivierung. *Die Ergebnisse stammen aus empirischen Studien der Bildungsforschung, bei denen Unterrichtsmerkmale erfasst und mit Wirkungen (Schulleistungen oder andere Zielkriterien) im Zusammenhang gebracht wurden.* Die eingangs erwähnten Übersichtsdarstellungen *beziehen sich auf* in Deutschland durchgeführte Studien wie TIMSS, PISA, DESI, Münchner Studie, SCHOLASTIK, VERA – Gute Unterrichtspraxis. Neben diesen im deutschen Sprachraum entstandenen Studien wurden auch die klassischen Unterrichtsstudien berücksichtigt, die überwiegend im angloamerikanischen Sprachraum entstanden. Daneben wurden die verfügbaren Monographien, Enzyklopädien, Handwörterbücher zur Thematik herangezogen, mit einem Schwerpunkt auf dem Handbook of Research on Teaching und dem Handbook of Educational Psychology. Schließlich wurden Metaanalysen (zusammenfassende systematische Analysen unterschiedlicher empirischer Studien zu einem Gesamtbild) verwendet, beispielsweise Haertel et al. (1983), Walberg (1986), Fraser et al. (1987), Scheerens & Bosker (1997) und kürzlich Seidel & Shavelson (2007). Andere Qualitätsmerkmale als die oben erwähnten, die bereits vielfach untersucht wurden (wie die Kompetenzorientierung) stehen erst seit kurzem auf der Agenda der Bildungspolitik und Bildungsforschung; hier befindet sich steht die Forschung erst am Anfang.

### 17. Geht bei der Zerlegung des Unterrichts in Einzelmerkmale nicht die Ganzheitlichkeit verloren?

Für viele Pädagogen ist die *analytische* „Zerlegung“ eines Gesamtbildes in separate Prinzipien oder Variablen ungewohnt und entspricht nicht den bei ihnen verbreiteten ganzheitlich orientierten Denkgewohnheiten. Es ist jedoch unabdingbar, um ein differenziertes Bild des Unterrichts, seiner Eigenheiten, Stärken und Schwächen zu gewinnen. Nach diesem ersten analytischen Blick muss jedoch ein zweiter, *synthetischer* Blick auf das Gesamtmuster erfolgen, andernfalls besteht die Gefahr von Kurzschlüssen: Ob eine niedrige Ausprägung eines Qualitätsmerkmals gravierend und beachtlich ist, hängt eben auch von der Gestalt des Gesamtprofils ab. Der in der traditionellen empirischen Unterrichtsforschung vorherrschende Ansatz lässt sich als *variablenorientiert* charakterisieren.

Es gibt auch in der modernen Forschung einen legitimen *personenorientierten* Zugang zur Frage des erfolgreichen Unterrichtens, nämlich den *Best-Practice-Ansatz*. Man sucht dabei zunächst nach Gruppen von Lehrpersonen, die einander ähnlich sind, und versucht dann diese Gruppen anhand verschiedener Merkmale zu beschreiben und in Form von Profilen anzuordnen. Dieses Vorgehen entspricht der im Alltag verbreiteten Sichtweise, dass es unterschiedliche Typen von Personen gibt, die sich klar voneinander abgrenzen lassen. In wissenschaftlichen Untersuchungen erweisen sich solche Typen dann zwar meistens als weitaus weniger prägnant voneinander abgrenzbar, als dies die Alltagsbetrachtung nahelegt. Im Bereich der Unterrichtsforschung ist man insbesondere an besonders erfolgreichen „Meisterlehrern“ (man spricht *gelegentlich auch* von best practice) interessiert und versucht, durch detaillierte Untersuchungen deren Erfolgsgeheimnis auf die Spur zu kommen. Solche Lehrpersonen können dann auch als Modelle dienen, an denen sich insbesondere künftige Lehrer orientieren können. In den letzten Jahren hat der Experten-Ansatz starke Beachtung gefunden. Für verschiedene Berufe und Tätigkeitsbereiche wurde nachgewiesen, dass sich Experten von Novizen vor allem in der Menge und Qualität des bereichsspezifischen Wissens unterscheiden, das bei Experten Ergebnis einer sich meistens über Jahre erstreckenden Auseinandersetzung mit den in ihrem Wissensbereich anfallenden Aufgabenstellungen ist. Der Expertise-Ansatz ist auch Grundlage von Forschungen zur Lehrertätigkeit (siehe etwa Bromme, 2008).

### **18. Geht ein effektiver lernförderlicher Unterricht nicht auf Kosten anderer Lernziele?**

In der empirischen Unterrichtsforschung ist vor allem untersucht worden, welche Unterrichtsmerkmale zum Lernerfolg der Schülerinnen und Schüler beitragen. Häufig ist damit die Befürchtung verbunden, dass ein auf Leistungsziele ausgerichteter Unterricht auf Kosten anderer Ziele geht, etwa die Lernfreude beeinträchtigt oder zu erhöhter Ängstlichkeit führt. Von Bildungstheoretikern sind neben der Leistung schon immer auch andere Ziele wie Mündigkeit, Selbstständigkeit, Kooperationsfähigkeit oder moralische Werte proklamiert worden. Diese Ziele haben auch Niederschlag in Schulgesetzen gefunden. Für die Unterrichtsforschung ergibt sich daraus die Frage, ob diese verschiedenen Ziele miteinander vereinbar sind oder ob Zielkonflikte unvermeidlich sind. Die Frage der Multikriterialität von Unterricht ist bislang erst Gegenstand weniger Untersuchungen gewesen. Die dort gewonnenen Ergebnisse zeigen, dass es durchaus möglich ist, verschiedene scheinbar disparate Ziele gleichzeitig zu erreichen (Gruehn, 1995; Helmke & Schrader, 1990; Kunter, 2005; Schrader, Helmke & Dotzler, 1997).

### **19. Gibt es objektive Kriterien dafür, was „guter Unterricht“ ist?**

Regelrechte Prozess-Standards, die den Professionalitätsstandards der KMK entsprechen, sind damit allerdings allenfalls in Ansätzen verfügbar. Wünschenswert wären Mindestausprägungen im Sinne einer kriterialen Bezugsnorm, die exakt angeben, wie häufig z. B. bestimmte Unterrichtsverhaltensweisen auftreten müssen, damit man von gutem Unterricht sprechen kann. Solche exakte Kriterien kann es allerdings aus verschiedenen Gründen nicht geben: (1) Ob ein Unterrichtsmerkmal wirksam ist, hängt auch immer davon ab, welche anderen Unterrichtsmerkmale vorliegen. (2) Unterrichtsmerkmale sind allgemeine Merkmale, die sich zwar an beobachtbaren Verhaltensweisen verankern lassen, aber oft auf ganz unterschiedliche Weise. (3) Aussagen über die Wirksamkeit von Unterrichtsmerkmalen basieren auf Forschungen, bei denen Klassen und Lehrkräfte verglichen werden, so dass man im Grunde immer nur Aussagen über die Qualität relativ zu anderen Klassen und Lehrkräften treffen kann (soziale Bezugsnorm). Nichtsdestoweniger haben Prozess-Standards eine wichtige Orientierungsfunktion, indem sie die Aufmerksamkeit auf zentrale Merkmalsbereiche lenken und Verbesserungen in diesen Bereichen anregen.

### **20. Kann man das Wissen über lernwirksame Unterrichtsmerkmale überhaupt für die Verbesserung und Weiterentwicklung des Unterrichts nutzen?**

Wie bereits angedeutet, bewegen sich die Erkenntnisse der empirischen Unterrichtsforschung zum Zusammenhang zwischen Unterrichtsmerkmalen und Produkten (Leistungen, Lernerfolg) auf einer vergleichsweise abstrakten Ebene. Es ist nicht immer unmittelbar ersichtlich, welche Lehrerverhaltensweisen und Lehrer-Schüler-Interaktionen einem Unterrichtsmerkmal wie hoher Strukturierung oder effizienter Klassenführung zugrunde liegen, zumal das gleiche Ziel oft auf unterschiedliche Weise erreicht werden kann. Daher ist auch oft nicht klar, welche Verhaltensweisen in realen Unterrichtssituationen einzusetzen sind, um Veränderungen auf einer Merkmalsdimension zu erreichen. Dies sei für das Merkmal „lernförderliches Klima“ durch Abbildung 6 illustriert:

Bekommt eine Lehrkraft die Rückmeldung, dass in ihrem Unterricht das lernförderliche Klima nur niedrig ausgeprägt ist, dann muss sie sich zunächst einmal fragen, durch welche Verhaltensindikatoren dieses abstrakte Merkmal gekennzeichnet ist, um Verbesserungen in die Wege zu leiten. Sie könnte sich dabei zunächst erst einmal klarmachen, dass ein lernförderliches Klima verschiedene Facetten aufweist wie einen wertschätzenden Umgangs-

ton oder einen Lernchancen bietenden und verständnisvollen Umgang mit Schülerfehlern. Auch diese Facetten sind noch *nicht konkret* genug, um Verhaltensänderungen vornehmen zu können. Zu diesem Zweck müsste sie die einzelnen Facetten noch weiter auf die Ebene konkreter Verhaltensweisen herunterbrechen. Sie könnte sich zunächst einmal auf die Facette „Lernchancen bietender Umgang mit Schülerfehler“ konzentrieren und sich klar machen, dass sie beim Auftreten von Schülerfehlern verschiedene Verhaltensmöglichkeiten hat (*Schülern* mitteilen, warum etwas falsch ist; Hinweise geben, die das Verständnis fördern usw.). Diese Beschreibungen sind vermutlich konkret genug, um in Situationen, in denen Schülerfehler auftreten, geeignete Verhaltensweisen einsetzen zu können.



Abb. 6 Hierarchieebenen der Unterrichtsqualität

Um eine systematische und kontrollierte Verhaltensänderung einzuleiten, ist es aber in der Regel nicht ausreichend, sich einfach vorzunehmen, in solchen Situationen ein anderes Verhalten zu zeigen. Vermutlich hat die Lehrkraft Routinen aufgebaut, wie sie sich in solchen Situationen üblicherweise verhält, und vielleicht sind ihr diese weitgehend automatisierten Reaktionsweisen gar nicht einmal hinreichend bewusst. Ein wichtiger Schritt ist daher, zunächst einmal durch kontrollierte Beobachtungen herauszufinden, wie sie sich in solchen Situationen tatsächlich verhält. Die Lehrkraft könnte etwa ein einfaches Beobachtungssystem mit den angenommenen Verhaltensweisen und anderen, die sie bei sich vermutet, als Kategorien erstellen und sich dann eine Zeitlang selbst beobachten und mit Hilfe einer einfachen Strichliste vermerken, welche Verhaltensweisen sie tatsächlich wie häufig ausführt. Wahl, Weinert und Huber (1997) machen Vorschläge, wie solche einfachen Beobachtungen im Unterricht durchgeführt werden können. Wenn genügend Beobachtungen vorliegen, könnte sie sich dann ein konkretes Ziel setzen, welches Verhalten in solchen Situationen gezeigt werden soll. Hilfreich ist es dabei, die zu zeigenden Verhaltensweisen möglichst konkret festzulegen, sich vielleicht sogar mögliche Formulierungen zu überlegen. Auch das neue Verhalten sollte (eine Zeitlang) beobachtet werden, bis es sich hinreichend stabilisiert hat. Schließlich geht es darum, mehr oder weniger eingeschlifene Gewohnheiten durch neue Reaktionsweisen zu ersetzen, und dies ist immer aufwändig.

Die in Abbildung 6 vorgenommene Aufschlüsselung zeigt aber auch, dass einem Merkmal wie lernförderlichem Klima auf ganz unterschiedliche Verhaltensweisen zugrunde liegen können. Das erschwert nicht nur die (Fremd- und Selbst-) Diagnose, weil man eine Vorstellung davon haben muss, welche Verhaltensindikatoren hinreichend äquivalent sind, sondern auch die Verhaltensänderung. In der Regel ist es sicher nicht sinnvoll, in einer Situation stereotyp mit ein und demselben Verhalten zu reagieren, sondern eine gewisse Variabilität und Flexibilität zu zeigen. Dafür ist aber Voraussetzung, dass man ein gewisses Repertoire geeigneter Verhaltensweisen verfügbar hat.

Hier wurde davon ausgegangen, dass die Lehrkraft den Prozess der Merkmalsaufschlüsselung selbst vornimmt. Differenziert ausgearbeitete Beurteilungsverfahren können hierfür allerdings eine wesentliche Hilfestellung geben, indem sie die Merkmalsaufschlüsselung schon in das Instrument selbst verlagern. Dies ist etwa in dem vom Erstautor für die Qualitätsagenturen im Land Schleswig-Holstein (Helmke, 2007) und Rheinland-Pfalz (Helmke, 2009; <http://aqs.rlp.de/>) entwickelten Verfahren „Einblick in die Lehr-Lernsituation (ELL)“ der Fall.

Noch schwieriger ist es, Veränderungen bei einem Merkmal mit Veränderungen bei anderen ebenfalls wichtigen Merkmalen zu koordinieren. Die empirische Unterrichtsforschung gibt dafür keine direkt umsetzbaren Empfehlungen (und kann es aufgrund ihres Ansatzes auch gar nicht tun), sondern stellt Kriterien bereit, anhand derer man Unterricht reflektieren und analysieren kann. Letztlich müssen dann Überlegungen angestellt werden, welche Verhaltensweisen und Interaktionen dazu beigetragen haben könnten, dass bei einem Unterrichtsmerkmal eine bestimmte Ausprägung resultiert, und welche Änderungen im Verhalten erforderlich sind, um sich auf einer Merkmalsdimension zu verbessern.

Es sei noch einmal darauf hingewiesen, dass es nicht darum geht, das Verhalten in einer konkreten Situation oder einer Unterrichtsstunde zu optimieren. Das Unterrichtsverhalten basiert auf eingeschliffenen Routinen und Automatismen, die entwickelt wurden, um mit der komplexen Unterrichtssituation überhaupt zurecht zu kommen (s. Frage 9). Als ungünstig und problematisch erkannte Routinen zu verändern, ist immer mühsam und aufwändig. Da Routinen auch immer Verhaltenssicherheit geben, sind Änderungsversuche zunächst häufig mit Unsicherheit verbunden. Änderungen müssen daher systematisch angegangen werden, und dabei ist Hilfestellung und Unterstützung, z. B. durch gleichgesinnte Kolleginnen und Kollegen, wichtig. Solche Kooperation sind oft allerdings nur gegen erhebliche Widerstände zu erreichen, da sich in Deutschland Lehrende aufgrund ihrer beruflichen Sozialisation noch immer meistens als Einzelkämpfer verstehen und die Tätigkeit im Klassenzimmer nach außen abgeschottet ist.

Wie schwierig und langwierig solche Prozesse der Unterrichtsentwicklung sein können, zeigen Untersuchungen zur Rezeption von Ergebnismeldungen aus Forschungsprojekten. Die Rückmeldung von Leistungsergebnissen ist mittlerweile Bestandteil der meisten großen Leistungsstudien (vgl. etwa Kohler & Schrader, 2004; Schrader, in Druck). Mittlerweile gibt es auch erste Versuche, Rückmeldungen zum Unterricht zu geben (A. Hosenfeld, in Vorb.). Die mit der Rezeption, Nutzung und Umsetzung von Informationen über den Unterricht und seine Effekte lassen sich anhand eines vom Erstautor entwickelten Modells veranschaulichen (vgl. Abbildung 7).

Damit Informationen über den Unterricht zu nachhaltigen Veränderungen führen, muss eine Kette von mehreren Schritten durchlaufen werden: Die Information muss zunächst beim Betroffenen ankommen und dann von ihm verstanden werden (*Rezeption*). Dann ist es in der Regel nötig, weitere Überlegungen anzustellen und ggf. zusätzliche Informationen einzuholen (*Reflexion*), z. B. durch Bildung und Überprüfung von Hypothesen (Welche Verhaltensweisen tragen zu den festgestellten Merkmalsausprägungen bei? Welche müssen verändert werden, um Veränderungen auf diesen Merkmalsdimensionen zu erzielen?). Dann müssen Entscheidungen getroffen, Veränderungen geplant und umgesetzt werden (*Aktion*). Dies ist oft nicht einfach, da in der Regel Routinen und eingeschliffene Gewohnheiten überdacht und geändert werden müssen. Wenn Entschlüsse gebildet und Handlungen eingeleitet worden sind, müssen diese gegen andere Einflüsse (Widerstände, konkurrierende Motive u. a.) abgeschirmt werden. Handlungen müssen bis zur Erreichung des Ziels aufrechterhalten werden. Wie diese Handlungen aussehen und welche Maßnahmen zu ihrer Umsetzung ergriffen werden (z. B. Fortbildung, Training, Super-

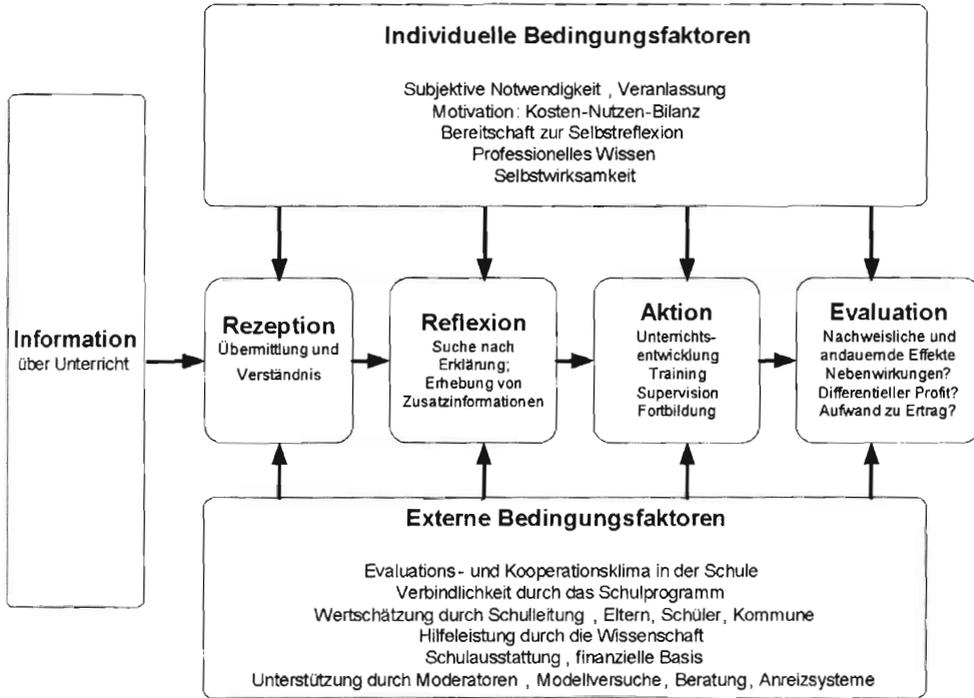


Abb. 7 Ein Sequenzmodell der Unterrichtsentwicklung und ihrer Bedingungen (Helmke, 2009, S. 309)

vision), hängt von der Art des Problems, dem Problemdruck und den verfügbaren Ressourcen ab. Wie die Volitionsforschung gezeigt hat, reichen die anfängliche Motivation und individuelle Bedingungs-faktoren dafür oft nicht aus. Es ist daher wichtig, dass externe Unterstützung vorhanden ist und genutzt werden kann. Durchgeführte Maßnahmen einer *Evaluation* zu unterziehen, ist wichtig, um festzustellen, ob und in welchem Umfang man erfolgreich war und was Ursachen für suboptimale Ergebnisse gewesen sein könnten. Erlebte Fortschritte, die in einer Evaluation sichtbar werden, sind eine wichtige Motivationsquelle. Überhaupt sollte man sich die im Sequenzmodell dargestellte Handlungskette nicht als einmaliges Geschehen, sondern eher als zyklische Abfolge vorstellen, da auch eine erfolgreiche Beseitigung von Defiziten nicht unbedingt dauerhafte Veränderungen garantiert und nach Aufarbeitung einer Schwachstelle vielleicht andere in den Vordergrund treten. Die fortlaufende Überprüfung und Verbesserung der eigenen Handlungs-kompetenzen ist ein Merkmal beruflicher Professionalität.

## 6. Konsequenzen für die Lehrerbildung

Die Reflexion über die Qualität und Wirkungsweise unterrichtlicher Prozesse macht zweifellos den Kern der Professionalisierung aus und dürfte vor allem – aber nicht nur – in der zweiten Phase der Lehrerbildung im Mittelpunkt stehen. Dazu können Erkenntnisse der empirischen Unterrichtsforschung einen wichtigen Beitrag leisten. Allerdings sind Lehrkräfte aufgrund ihrer Ausbildung häufig noch wenig darauf vorbereitet, diese Erkenntnisse zu nutzen. Lehrkräften die dazu nötigen Kenntnisse und Kompetenzen zu vermitteln,

müsste daher zu einer wichtigen Aufgabe der Lehreraus- und -weiterbildung werden. Außerdem müssen geeignete Formen der Unterstützung gefunden und eingeübt werden, etwa in Form kollegialer Zusammenarbeit oder Beratung und Supervision durch Experten. Neben dem Lernen von Experten, dem Lernen an guten Beispielen und durch Beobachtung benötigt Reflexion über Unterricht vor allem geeignete Werkzeuge, etwa in Gestalt von Fragebögen, Checklisten und Unterrichtsbeobachtungsbögen. In den Servern der Landesinstitute sowie in Österreich und in der Schweiz stehen hierzu bereits vielfältige Werkzeuge zur Verfügung.

Abschließend soll auf eine Neuentwicklung zur *Unterrichtsdiagnostik* hingewiesen werden, die künftig allen interessierten Lehrpersonen – in Ausbildung und Beruf – zur kostenlosen Nutzung zur Verfügung stehen wird. Es handelt sich dabei um einen *Schülerfragebogen* zu Merkmalen der Unterrichtsqualität und als Tandempartner einen äquivalenten *Lehrerfragebogen* zur Selbsteinschätzung einer konkreten Unterrichtsstunde. Dieses vom Erstautor entwickelte Fragebogen-Duo wurde in Klassen der Sekundarstufe I empirisch getestet (Piskol, 2008) und wird im Heft 1/1009 der „Lernenden Schule“ ausführlich beschrieben (Helmke, Piskol, Pikowsky & Wagner, im Druck). Ab 2009 werden diese Instrumente zum kostenlosen Herunterladen auf der Homepage des Erstautors zur Verfügung stehen<sup>4</sup>.

Tab. 8: Ausblick: Lehrer- und Schülerfragebogen für die Selbstevaluation und Unterrichtsreflexion (Helmke, Piskol, Pikowsky & Wagner, im Druck)

Der Fragebogen lässt sich mit einem – ebenfalls herunterladbaren – EXCEL-Programm auswerten (Entwicklung: Wolfgang Wagner). Zunächst werden die Schüler- und Lehrer-Fragebogendaten in die EXCEL-Maske eingegeben. Nach der Dateneingabe stellt das Programm, unterteilt nach den einzelnen Inhaltsbereichen, verschiedene Kennwerte zur Verfügung (Mittelwert und Streuung der Schülerurteile, prozentualer Schüleranteil mit jeweils zustimmenden Einschätzungen, Anzahl der vorliegenden Schülerurteile pro Aussage, Differenz zwischen Lehrer- und Schülerurteil. Diese Kennwerte lenken den Fokus bei der Interpretation der Daten auf jeweils unterschiedliche Aspekte:

- Der (Klassen-) *Mittelwert* gibt Auskunft über die durchschnittlichen Einschätzungen der Schülerinnen und Schüler, während die *Streuung* den Blick auf Konsens bzw. Dissens innerhalb der Klasse lenkt.
- Der prozentuale Schüleranteil mit zustimmenden Einschätzungen fokussiert auf den Kontrast *ungünstiger vs. günstiger Einschätzungen*.
- Was die Anzahl der vorliegenden Schülerurteile anbelangt, so sollten diese möglichst die gesamte Klasse (mit Ausnahme von Schülern, die nicht teilnehmen konnten) umfassen. Fehlen unerwartet viele Angaben, dann sind die Ergebnisse zu dieser Unterrichtseinheit nur eingeschränkt beurteilbar.
- *Differenzen zwischen Lehrer- und Schüler-Einschätzungsprofil* zeigen, welche Unterrichtsmerkmale von der Lehrkraft günstiger oder weniger günstig eingeschätzt werden als von ihren Schülerinnen und Schülern.

Zusätzlich zu den numerischen Kennwerten bietet das Programm auch grafische Darstellungen der Ergebnisse an (s. Abbildung 8).

<sup>4</sup> <http://www.uni-koblenz-landau.de/landau/fb8/einrichtungen/entwicklungspsychologie>

Zunächst einmal ist den Abbildungen auf einen Blick zu entnehmen, wo Stärken und Schwächen (aus Schüler- und Lehrersicht) der beurteilten Unterrichtseinheit liegen.

Für jede Frage wird die *Verteilung* der Schülerantworten auf die einzelnen Antwortalternativen – also die Anzahl der Einschätzungen, die auf die einzelnen Kategorien entfallen – in Form eines sogenannten Balkendiagramms veranschaulicht. Daraus lässt sich entnehmen, wie die als Kennwerte berichteten Streuungen zustande kommen: Große Streuungen können aus einer 'breiten', gleichmäßigen Verteilung der Einschätzungen über die Antwortalternativen resultieren, aber auch daraus, dass zwei in sich homogene Schülergruppen existieren, die sich in ihrer Unterrichtswahrnehmung aber deutlich unterscheiden.

Schließlich werden die Profile der von Lehrkräften vs. Schüler/innen (Klassenmittelwerte) eingeschätzten Ausprägungen der einzelnen Unterrichtsmerkmale kontrastiert. So lässt sich erkennen, ob etwaige Differenzen über alle Merkmale hinweg eher konstant verlaufen (also permanente Über- oder Unterschätzungen vorliegen), oder ob sie bereichs- oder sogar itemspezifisch sind.

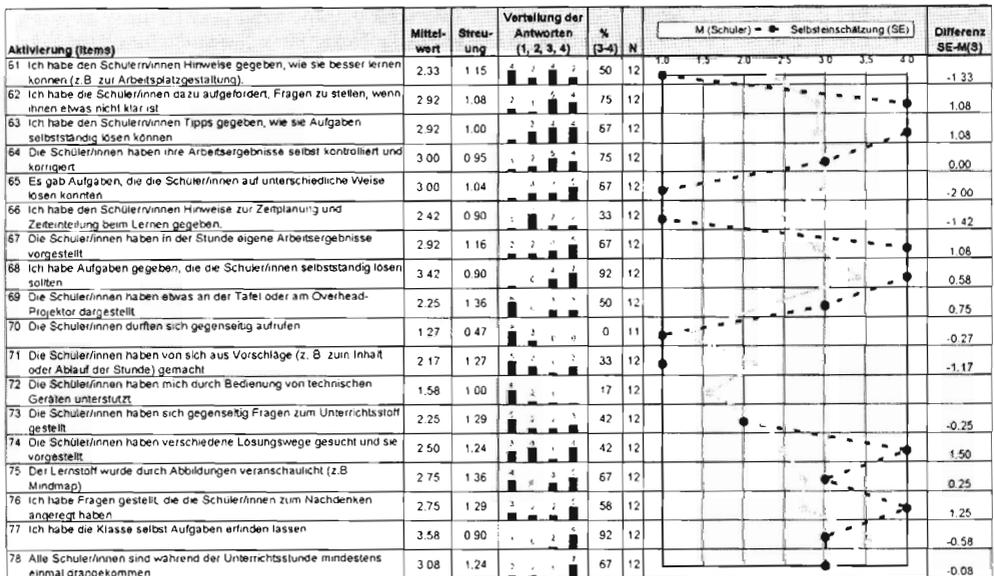


Abb. 8 Ausschnitt aus einer exemplarischen Ergebnisdarstellung für einen Teilbereich (hier: Aktivierung) des Lehrer- und Schülerfragebogens für eine konkrete Unterrichtsstunde

### 7. Literatur

Arnold, K.-H., Koch-Priewe, B. & Lin-Klitzing, S.: Allgemeine Didaktik, Fachdidaktik und Unterrichtsqualität. In K.-H. Arnold (Hrsg.), *Unterrichtsqualität und Fachdidaktik* (S. 19–49). Bad Heilbrunn: Klinkhardt 2007.

Borich, G.D.: *Observation Skills for Effective Teaching* (5. Aufl.). Upper Saddle River, NJ: Merrill Prentice Hall 2007.

Bromme, R.: Lehrerexpertise. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (Handbuch der Psychologie, Band 10, S. 159 -167). Göttingen: Hogrefe 2008.

Brophy, J.E.: *Teaching* (Educational Practices Series, Vol. 1). Brussels: International Academy of Education & International Bureau of Education 2008 ([www.ibe.unesco.org](http://www.ibe.unesco.org)).

- Brunnhuber, P.: *Prinzipien effektiver Unterrichtsgestaltung*. Donauwörth: Auer 1971.
- Clausen, M.: *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann 2002.
- Ditton, H.: Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick über den 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* (S. 73–92). Weinheim: Beltz 2000.
- Fraser, B.J., Walberg, H.J., Welch, W. & Hattie, J.A.: Syntheses of educational productivity research. *International Journal of Educational Research, 11*, 1987, 145–252.
- Gruehn, S.: Vereinbarkeit kognitiver und nicht-kognitiver Ziele im Unterricht. *Zeitschrift für Pädagogik, 41*, 1995, 531–553.
- Haertel, G.D., Walberg, H.J. & Weinstein, T.: Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research, 53*, 1983, 75–91.
- Hasselhorn, M. & Gold, A.: *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren*. Stuttgart: Kohlhammer 2006.
- Helmke, A.: Was wissen wir über guten Unterricht? *Pädagogik, 2*, 2006, 42–45.
- Helmke, A.: Einblick in die Lehr-Lern-Situation. In T. Riecke-Baulecke (Hrsg.), *SchulleitungPlus: Schule und Unterricht erfolgreich gestalten* (S. 88–101/208–209). München: Oldenbourg 2007.
- Helmke, A.: *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung der Unterrichtsqualität*. Seelze: Klett-Kallmeyer 2009.
- Helmke, A., Goellner, R., Kleinbub, I., Schrader, F.-W. & Wagner, W.: Standards – Motor oder Bremse der Unterrichtsentwicklung? *SEMINAR – Lehrerbildung und Schule. Kompetenzerwerb in der Lehrerbildung, 14* (1), 2008, 44–61.
- Helmke, A., Helmke, T., Heyne, N., Nordheider, I. & Schrader, F.-W.: Analyse und Bewertung von Fremdsprachenunterricht mit Hilfe von standardisierten Erhebungsinstrumenten. In U.O.H. Jung (Hrsg.), *Praktische Handreichung für Fremdsprachenlehrer* (4. vollst. neu bearb. Aufl., S. 537–541). Frankfurt am Main: Lang 2006.
- Helmke, A., Helmke, T., Heyne, N., Hosenfeld, A., Schrader, F.-W. & Wagner, W.: Effiziente Klassenführung als Schlüsselmerkmal der Unterrichtsqualität – ein Untersuchungsbeispiel aus der Grundschule. In C. Spiel, R. Reimann, B. Schober & P. Wagner (Hrsg.), *Bildungspsychologie*. Göttingen: Hogrefe (in Druck).
- Helmke, A., Helmke, T., Schrader, F.-W., Wagner, W., Klieme, E., Nold, G. & Schröder, K.: Wirksamkeit des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 382–397). Weinheim: Beltz 2008.
- Helmke, Piskol, Pikowsky & Wagner: Schüler als Experten von Unterricht: Unterrichtsqualität aus Schülersperspektive. *Lernende Schule*. (in Druck)
- Helmke, A. & Schrader, F.-W.: Zur Kompatibilität kognitiver, affektiver und motivationaler Zielkriterien des Schulunterrichts – Clusteranalytische Studien. In M. Knopf & W. Schneider (Hrsg.), *Entwicklung. Allgemeine Verläufe – Individuelle Unterschiede – Pädagogische Konsequenzen* (S. 180–200). Göttingen: Hogrefe 1990.
- Hosenfeld, A.: *Führt Unterrichtsrückmeldung zu Unterrichtsentwicklung? Die Wirkung von videographischer und schriftlicher Rückmeldung bei Lehrkräften der vierten Jahrgangsstufe*. Dissertation, Universität Koblenz-Landau, Campus Landau, Fachbereich Psychologie.
- Huwendiek, V. & Bovet, G. (Hrsg.): *Leitfaden Schulpraxis. Pädagogik und Psychologie für den Lehrberuf*. Berlin: Cornelsen 2005.
- Klieme, E.: Was ist guter Unterricht? Ergebnisse der TIMSS-Videostudie im Fach Mathematik. In W. Bergsdorf, J. Court, M. Eckert & H. Hoffmeister (Hrsg.), *Herausforderungen der Bildungsgesellschaft. 4. Ringvorlesung der Universität Erfurt*. (S. 89–113). Weimar: Rhino Verlag 2002.
- Klieme, E., Jude, N., Rauch, D., Ehlers, H., Helmke, A., Eichler, W., Thomé, G. & Willenberg, H.: Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 319–344). Weinheim: Beltz 2008.

- Kohler, B. & Schrader, F.-W. (Hrsg.): *Ergebnisrückmeldung und Rezeption. Von der externen Evaluation zur Entwicklung von Schule und Unterricht* (Empirische Pädagogik/Themenheft, Vol. 18/1). Landau: VEP 2004.
- Kunter, M.: *Multiple Ziele im Mathematikunterricht*. Münster: Waxmann 2005.
- Mayr, J.: Klassen kompetent führen. Ergebnisse aus der Forschung und Anregungen für die Lehrerbildung. SEMINAR – Lehrerbildung und Schule, 14, Heft 1/2008, 76–87.
- Meyer, H.: *Was ist guter Unterricht?* Berlin: Cornelsen 2004.
- Mühlhausen, U.: *Schüleraktivierung im Schulalltag – Band 1 – Ungewöhnliche Unterrichtsmethoden in der Sekundarstufe* (Begleit-DVD mit 10 Hannoveraner Unterrichtsbildern). Baltmannsweiler: Schneider Verlag Hohengehren 2008.
- Nuthall, G.: Understanding Student Thinking and Learning in the Classroom. In B.J. Biddle, T.L. Good & I.F. Goodson (Hrsg.), *International Handbook of Teachers and Teaching*. (Vol. 2, S. 681–768). Dordrecht, Netherlands: Kluwer Academic Publishers 1997.
- Oser, F. & Baeriswyl, F.J.: Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (4. Aufl., S. 1031–1065). Washington: American Educational Research Association 2002.
- Piskol, K.: *Unterrichtsqualität aus der Schülerperspektive: Ein Verfahren zur Unterrichtsentwicklung*. Unveröffentlichte Diplomarbeit, Universität Mannheim 2008.
- Scheerens, J. & Bosker, R.J.: *The foundations of educational effectiveness*. Oxford: Pergamon 1997.
- Scheltwort, P.: *Lehrerhandeln wahrnehmen, erfassen, bewerten. Theoretische und empirische Analysen zum systemanalytischen Beobachtungsverfahren „Lehrerhandeln im Unterricht“*. Unveröffentlichte Dissertation, Pädagogische Hochschule Heidelberg 2006.
- Schmoll, L.: Allgemeindidaktische Modelle in der Lehrerbildung. *Pädagogik*, 60, Heft 10/Okttober 2008, 36–39.
- Schrader, F.-W.: Leistungsmessung und Rückmeldung. In T. Bohl, W. Helser, H.G. Holtappels & C. Schnelle (Hrsg.), *Handbuch Schulentwicklung*. Bad Heilbrunn: Klinkhardt (in Druck).
- Schrader, F.-W., Helmke, A. & Dotzler, H.: Zielkonflikte in der Grundschule: Ergebnisse aus dem SCHOLASTIK-Projekt. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 299–316). Weinheim: Psychologie Verlags Union 1997.
- Seidel, T. & Shavelson, R.J.: Teaching effectiveness research in the last decade: Role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77, 2007, 454–499.
- Slavin, R.E.: *Educational Psychology* (5. Aufl.). Boston: Allyn and Bacon 1997.
- Terhart, E.: Fremde Schwestern. Zum Verhältnis von Allgemeiner Didaktik und empirischer Lehr-Lern-Forschung. *Zeitschrift für Pädagogische Psychologie*, 16 (2), 2002, 77–86.
- Wagner, W.: *Methodenprobleme bei der Analyse der Unterrichtswahrnehmung und -wirksamkeit – am Beispiel der Studie DESI (Deutsch Englisch Schülerleistungen International) der Kultusministerkonferenz*. Dissertation. Universität Koblenz-Landau, Campus Landau, Fachbereich Psychologie 2008.
- Wahl, D., Weinert, F.E. & Huber, G.L.: *Psychologie für die Schulpraxis* (6. Aufl.). München: Kösel 1997.
- Walberg, H.J.: Syntheses of research in teaching. In M.C. Wittrock (Hrsg.), *Handbook of research on teaching* (3. Aufl., S. 214–229). London: Macmillan 1986.
- Walberg, H.J. & Paik, S.J.: *Effective educational practices*. Brussels: International Academy of Education & International Bureau of Education ([www.ibe.unesco.org](http://www.ibe.unesco.org)) 2000.
- Weinert, F.E. & Helmke, A. (Hrsg.): *Entwicklung im Grundschulalter*. Weinheim: Psychologie Verlags Union 1997.
- Wellenreuther, M.: *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht* (2. korrigierte und überarbeitete Aufl.). Baltmannsweiler: Schneider Verlag Hohengehren 2005.
- Woolfolk, A.: *Pädagogische Psychologie* (10. Aufl.). München: Pearson 2008.
- Prof. Dr. Andreas Helmke und Dr. Friedrich-Wilhelm Schrader,*  
Universität Koblenz-Landau, Fachbereich Psychologie, Arbeitseinheit Entwicklungspsychologie und Bildungsforschung  
(<http://www.uni-koblenz-landau.de/landau/fb8/einrichtungen/entwicklungspsychologie>)