Fachgebiet Erziehungswissenschaft

Inaugural-Dissertation zur Erlangung des akademischen Grades Dr. phil. der Pädagogischen Hochschule Weingarten

Praktika und ihre Lernpotentiale in der Lehrerbildung – Eine längsschnittliche Videostudie zur Untersuchung der Entwicklung unterrichtlicher Handlungskompetenzen in verlängerten Praxisphasen

vorgelegt von

Oliver Küster

aus Flensburg

Weingarten 2008

Erstgutachter: Prof. Dr. G. Dörr
Zweitgutachterin: Prof. Dr. D. Kucharz

Es ist eine gute akademische Tradition, sich nach der Fertigstellung einer langwierigen Arbeit bei den Personen zu bedanken, ohne die diese Arbeit nicht möglich gewesen wäre. Als erstes möchte ich mich bei Günter Dörr und Diemut Kucharz bedanken, die wertvolle Hilfestellungen und Hinweise im Entstehungsprozess der Arbeit gegeben und mich auch persönlich unterstützt haben.
Zu großem Dank bin ich auch den Schweizer Kollegen aus dem Forschungsprojekt „Standarderreichung in der Lehrerinnen- und Lehrerbildung“ verpflichtet, mit denen in einem langwierigen diskursiven Prozess das Beobachtungsinstrumentarium entwickelt wurde. Stellvertretend für die zahlreichen Kollegen sei der Projektleiter Matthias Baer von der Pädagogischen Hochschule Zürich genannt.
Des Weiteren gilt mein Dank den am Modellversuch beteiligten Personen bzw. Institutionen. Besonders hervorheben möchte ich die Kooperationsbereitschaft der Teilnehmer des Praxisjahres, ohne die diese empirische Studie nicht hätte durchgeführt werden können. Gleiches gilt auch für die betreuenden Mentoren bzw. die Rektoren der Schulen und Mitarbeiter des Seminars Laupheim.
Wertvolle Hinweise kamen auch aus dem den Modellversuch begleitenden Kolloquium. Stellvertretend für alle Teilnehmer sei deshalb Bernd Reinhoffer als Leiter des Modellversuches für die Unterstützung gedankt. Dank geht auch an das Rektorat der Pädagogischen Hochschule Weingarten, welches den Modellversuch mit Forschungsmitteln großzügig unterstützt hat.
Besondere Anerkennung haben auch meine unmittelbaren Bürokollegen und Mitstreiter in der Untersuchung des Modellversuches Katharina Müller, Tanja Rosenberger und Robert Barthold verdient.
Zuletzt, aber nicht als Letztes, danke ich meiner Lebensgefährtin Ute Holm für die unermüdliche Unterstützung und Toleranz.

Inhalt

1	Einleitung	S. 5
2	Annäherungen an den Kompetenzbegriff	S. 8
2.1	Etymologische Klärungen des Kompetenzbegriffes	S. 8
2.2	Etablierung des Kompetenzbegriffes im wissenschaftlichen Sprachgebrauch	S. 10
2.3	Kommunikative Kompetenz	S. 10
2.4	Kritische Kompetenz	S. 12
2.5	Schlüsselqualifikationen	S. 13
2.6	Transferpotentiale und Schlüsselqualifikationen	S. 16
2.7	Pädagogische Handlungskompetenz	S. 17
2.8	Kompetenzentwicklung statt Bildung?	S. 19
2.9	Fazit	S. 23
3	Kompetenzen als latente Konstrukte erziehungswissenschaftlicher und psychologischer Forschung	S. 25
3.1	Kompetenzmodellierungen in Abhängigkeit des Abstraktionsniveaus	S. 25
3.2	Dimensionen des Kompetenzbegriffes in der Erziehungswissenschaft und der Psychologie	S. 26
3.3	Definitionen des Kompetenzbegriffes in der Erziehungswissenschaft und der Psychologie	S. 27
3.4	Die Erfassung von Kompetenzen	S. 31
3.5	Kompetenzklassen	S. 32
3.6	Generierung eines Kompetenzmodells	S. 34
3.7	Kompetenz und Expertise	S. 35
3.7.1	Expertiseforschung	S. 35
3.7.2	Expertiseforschung und berufliche Handlungskompetenz	S. 38
3.7.3	Prototypische Modellierung von Expertise	S. 40
3.8	Verfahren zur Kompetenzmessung	S. 41
3.8.1	Kompetenzmessung durch Selbstbeurteilungen	S. 42
3.8.2	Kompetenzmessung durch Fremdeinschätzungen	S. 43
3.8.3	Kompetenzmessung durch Portfolioansätze	S. 43
3.8.4	Kompetenzmessung durch Testverfahren	S. 44
3.9	Fazit	S. 45
4	Entwicklung professioneller didaktischer Handlungskompetenz	S. 46
4.1	Paradigmen empirischer Forschung in didaktischen Handlungsfeldern des Lehrerberufes	S. 46
4.1.1	Paradigma der Lehrerpersönlichkeit	S. 47
4.1.2	Prozess-Produkt-Paradigma	S. 48
4.1.3	Das Experten-Paradigma	S. 48
4.2	Das Angebot-Nutzungs-Modell unterrichtlicher Wirkungen	S. 50
4.3	Der Kompetenzbegriff in der Lehrerbildung	S. 52

4.3.1	Referenzsysteme zur Generierung von Wissen als Grundlage für die Entwicklung von Lehrerkompetenzen	S. 53
4.3.2	Topologien und Typologien professionellen Lehrerwissens	S. 57
4.3.3	Kognitionspsychologische Erforschung des Lehrerwissens und -handelns	S. 59
4.3.4	Generierung von Kompetenzmodellen in der Lehrerbildung	S. 64
4.3.5	Kompetenzentwicklung durch Professionalisierung?	S. 65
4.4	Fazit	S. 70
5	Standardorientierte Kompetenzmodelle der Lehrerbildung	S. 72
5.1	Standardmodelle in den USA	S. 72
5.2	Standardbasierte Untersuchung der Wirksamkeit der Lehrerbildung in der Schweiz (Oser)	S. 76
5.3	Standards für die Lehrerbildung in Deutschland (Terhart)	S. 82
5.4	Vergleich der Standardmodelle von Oser und Terhart	S. 86
5.5	Fazit: Kritik an den Standardmodellen	S. 87
6.	Spezifika des Lernens in Praxislernsituationen des Lehrerberufes	S. 93
6.1	Erfahrungswissen	S. 94
6.2	Didaktisches Handeln des Experten als Ergebnis von Praxislernsituationen?	S. 101
6.3	Kontingenz als strukturelles Merkmal von Praxislernsituationen	S. 104
6.4	Fazit	S. 106
7.	Ergebnisse der empirischen Bildungsforschung	S. 109
7.1	Schülerleistungen	S. 111
7.1.1	BIJU (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter)	S. 111
7.1.2	DESI (Deutsch-Englisch Schülerleistungen International)	S. 112
7.1.3	IGLU (Internationale Grundschul-Leseuntersuchung)	S. 114
7.1.4	Markus (Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext)	S. 116
7.1.5	PISA (Programme for International Student Assessment)	S. 118
7.1.6	TIMMS (Third International Mathematics and Science Study)	S. 119
7.2	Wirksamkeit der Lehrerbildungssysteme	S. 123
7.2.1	Untersuchungen zur Wirksamkeit des Lehrerbildungssystems in der Schweiz	S. 124
7.2.2	Ergebnisse der Wirksamkeit des Lehrerbildungssystems in der Schweiz	S. 127
7.3	Wirksamkeit der Lehrerbildung in den USA im Spannungsfeld von Prozessen der Deregulation und Professionalisierung	S. 132
7.4	Internationale Ergebnisse zur Kompetenzstruktur von Studierenden des Lehrerberufes	S. 135
7.5	Empirische Studien zur Lehrerbildung in Deutschland	S. 136
7.5.1	Untersuchungen zur Entwicklung von diagnostischer Kompetenz in der zweiten Phase der Lehrerbildung	S. 138

7.5.2	Untersuchungen zum Erwerb unterrichtsrelevanten pädagogischen Wissens	S. 142
7.6	Fazit	S. 144
8	Untersuchungen zur Unterrichtqualität als Ausgangspunkt für videobasierte Beobachtungsverfahren	S. 147
9	Anlage der Videostudie zur längsschnittlichen Untersuchung unterrichtsrelevanter Kompetenzen in verlängerten Praxisphasen	S. 154
9.1	Fragestellungen	S. 154
9.2	Design	S. 156
9.3	Beschreibung der Stichprobe	S. 157
10	Entwicklung von videobasierten Beobachtungsverfahren	S. 159
10.1	Möglichkeiten und Grenzen von Videostudien	S. 160
10.2	Problemfelder methodischer Beobachtungsverfahren	S. 164
10.3	Unterrichtsvideos als didaktisches Medium	S. 168
10.4	Entwicklung niedrig-inferenter Kategorien- und Zeichensysteme	S. 170
10.5	Entwicklung hoch-inferenter Beobachtungsverfahren	S. 179
11	Ergebnisse	S. 188
11.1	Interkoderreliabilitäten der niedrig-inferenten Kodierungen	S. 188
11.2	Ergebnisse der niedrig-inferenten Kodierungen über die drei Messzeitpunkte	S. 188
11.2.1	Ergebnisse für die Dimension „Unterrichtsstatus“	S. 188
11.2.2	Ergebnisse für die Dimension „Unterrichtliche Arbeitsformen“	S. 189
11.2.3	Ergebnisse für die Dimension „Allgemeindidaktische Unterrichtsphasen“	S. 191
11.2.4	Ergebnisse für die Dimension „Aktivitäten im Klassenunterricht“	S. 192
11.2.5	Ergebnisse für die Dimension „Unterrichtskommunikation“	S. 194
11.2.6	Ergebnisse für die Dimension „Strukturierung“	S. 195
11.2.7	Ergebnisse für die Dimension „Differenzierung“	S. 196
11.2.8	Diskussion der Kodierungsergebnisse	S. 197
11.3	Interraterreliabilitäten	S. 201
11.4	Reliabilitäten der Skalen	S. 202
11.5	Deskriptive Kennwerte	S. 203
11.6	Überprüfung der Voraussetzung homogener Varianzen	S. 207
11.7	Überprüfung der Voraussetzung der Normalverteilung	S. 207
11.8	Vergleich der Mittelwerte über die Messzeitpunkte (t-Test)	S. 208
11.9	Diskussion der Signifikanzwerte über die Messzeitpunkte	S. 209
12	Zusammenfassende Betrachtung und Diskussion	S. 211
Literatur		S. 217
Anhang:		S. 253
Verzeichnis der Abbildungen		

Verzeichnis der Tabellen
Beschreibung des Modellversuches
Überblick über das Kodiersystem
Überblick über das Raterinventar
Eidesstattliche Versicherung

1 Einleitung

Die Lehrerbildung[1] ist durch verschiedene z.T. internationale, z.T. nationale Schulleistungsuntersuchungen in die Kritik geraten. Prototypisch lässt sich die Kritik anhand der von Oser & Oelkers (2001) vorgelegten empirischen Studie über die Wirksamkeit der Lehrerbildungssysteme in der Schweiz festmachen. Auch wenn die Ausbildung in der Schweiz andere Strukturen als in Deutschland aufweist, so wurde der zentrale Kritikpunkt einer unzureichenden Kompetenz- bzw. Standardorientierung auch auf Deutschland übertragen und führte zur Implementation von Standardkonzepten in die Lehrerbildung (KMK 2004). Diskurse über eine Neuausrichtung der Lehrerbildung sind untrennbar mit der Frage verknüpft, wie Praxiserfahrungen, die sog. schulpraktischen Studien, in das Curriculum integriert werden können oder sollen.
Die vorliegende Studie befasst sich mit der Untersuchung eines Modellversuches, in dem bis zu 16 Studierende die Möglichkeit erhalten, nach Abschluss des zweiten Semesters für ein Schuljahr an einer Biberacher Grund- bzw. Hauptschule schulpraktische Erfahrungen zu sammeln.[2]
Für die eigene Untersuchung wurde die Frage zugrunde gelegt, ob sich die unterrichtlichen Kompetenzen der Studierenden durch und innerhalb dieser frühen Praxisbegegnung entwickeln und welche Tendenzen aufgezeigt werden können. Die Studie ist bewusst explorativ angelegt, weil etwaige Hypothesen über längsschnittliche Verläufe von Kompetenzentwicklungsprozessen nicht dezidiert genug aufgestellt werden konnten. Bislang gibt es keine gesicherten Erkenntnisse darüber, wie die Entwicklungsdynamiken komplexer Kompetenzprofile verlaufen.

Forschungsmethodisch handelt es sich um eine Videostudie, die niedrig- und hoch-inferente Beobachtungsverfahren einsetzt und integriert. Das Design entspricht einer explorativen, längsschnittlich angelegten quasiexperimentellen Interventionsstudie mit Pre- und Posttest. Alle Studierenden der betreffenden Kohorte wurden zu jeweils drei Zeitpunkten mit der Kamera während des Unterrichts gefilmt und aufgenommen, um einen Längsschnitt darstellbar und analysierbar zu machen.
Die Forschungsrichtung fokussiert die Entwicklung von unterrichtlichen und nicht von didaktischen Kompetenzen, indem nur die Daten der Videostudie vorgelegt und im Rahmen der Fragestellung diskutiert werden. Analysen und schlussfolgernde Ergebnisse sind damit nur auf einer Sicht- bzw. Handlungsebene möglich. Um eine didaktische Kompetenzentwicklung zu untersuchen, wäre es erforderlich, die betreffenden kognitiven Planungsprozesse zu berücksichtigen oder aber die Studierenden in Form eines stimulated recall mit dem Videomaterial zu konfrontieren, um etwaige Reflexionsprozesse mit in den Blick zu nehmen.
In der Gesamtmethodik der Untersuchung des Modellversuches wurde dieser Sachverhalt berücksichtigt (vgl. Dörr & Küster 2006; Reinhoffer, Barthold & Küster 2007; Müller i.V.). Insofern stellt die Arbeit eine Teilstudie innerhalb der

[1] Gemäß der Promotionsordnung der Pädagogischen Hochschule Weingarten wird auf die Verwendung von Doppelformen oder andere Kennzeichnungen für weibliche und männliche Personen verzichtet, um die Lesbarkeit und Übersichtlichkeit zu wahren. Mit allen im Text verwendeten Personenbezeichnungen sind stets beide Geschlechter gemeint.
[2] Zur genauen Beschreibung des Modellversuches s. Anhang.

Untersuchung des Modellversuches dar und versteht sich ausdrücklich nicht als Evaluationsstudie über die Wirksamkeit des Modellversuches.
Vor diesem Hintergrund sind die Ergebnisse einzuordnen, eingedenk dessen, dass weitere Forschungsmethodiken zu ergänzenden oder auch divergenten Ergebnissen führen können.

Die Arbeit gliedert sich in drei Teile. Der erste und Grundlagenteil folgt dem Anliegen, die eigene Untersuchung bezugstheoretisch zu verorten und den zugrunde gelegten Reflexions- und Forschungsstand nachzuzeichnen. Generell erfolgt die Literaturrezeption und -interpretation angesichts der äußerst ausdifferenzierten und zentralen Bezugsdisziplinen der Erziehungswissenschaft und Pädagogischen Psychologie unter dem Selektionskriterium der hinführenden Relevanz für die eigene Fragestellung. Eine interdisziplinäre Bearbeitung der Thematik ist so in der fundamentalen Verwobenheit erziehungswissenschaftlicher Unterrichtsforschung mit kognitions- und lernpsychologischen Konstrukten angelegt. Der zweite Teil stellt die Anlage, Durchführung und Ergebnisse der eigenen Untersuchung dar und ist insofern als methodischer Teil konzipiert, als die besonderen Herausforderungen von Videostudien diskutiert und Spezifika der Methode dargestellt werden. Zentral sind die Unterscheidung von Beobachtungsverfahren, die sich bezüglich des Inferenzgrades differenzieren lassen sowie Einordnungsschwierigkeiten, die sich daraus ergeben. Im dritten Teil werden die Ergebnisse vor dem Hintergrund des bezugstheoretischen Erkenntnisstandes diskutiert und perspektivisch bewertet.

Was die konkrete Kapitelabfolge betrifft, ist in Teil I zunächst eine Abbildung der Diskussion intendiert, die der eigenen Studie zugrunde liegt und ihren Theoriehorizont bildet. Handelt die Fragestellung um unterrichtliche Qualität und daraus abzuleitende Kompetenzen, ist sie im theoretischen Bezugsrahmen der Kompetenzdebatte und Unterrichtsqualität verortet. Diesen Rahmen in seiner Dynamisierung einzufangen meint, ausgehend von einer grundlagentheoretischen Kompetenzbestimmung über die Entfaltung verschiedener – etwa disziplinärer oder bedeutungsbezogener – Stränge der Kompetenzdebatte zu den aktuellen Kompetenzmodellen der empirischen Bildungsforschung zu gelangen. Eine auszumachende Steigerung, die die zumindest partiell zeitchronologisch nachzuvollziehende Kompetenzdiskussion enthält, ergibt sich deshalb, weil Zusammenhänge sich mit andauernder Forschungs- und Diskursintensität als zunehmend komplex und vernetzt zu erkennen geben.
Eine stringente Chronologie kann in der konzeptionellen Umsetzung jedoch nicht vollständig eingehalten werden, weil der Diskurs selbst immer wieder Vor- und Rückgriffe auf vermeintlich bereits abgeschlossene Diskussionsstränge, disziplin- und bereichsübergreifende Verweise vornimmt, um Zusammenhänge herauszuarbeiten und die jeweils eigene Argumentation zu untermauern. Im Sinne der Leserverständlichkeit wird zwar in den Grundlagenkapiteln versucht, entsprechend sprunghafte Skizzierungen zugunsten übergreifender Rekonstruktionslinien möglichst gering zu halten. Aus Gründen der plausiblen Darstellungslogik kann aber nicht gänzlich auf sie verzichtet werden.
Der dargelegten Struktur folgend, sind die Kapitel 2, 3 und 4 dem Kompetenzverständnis in seiner grundlagentheoretischen, multidisziplinären und

empirischen Entfaltung gewidmet. Kapitel 5 thematisiert die Frage einer standardbasierten Untersuchung der Curricula der Lehrerbildung bzw. ihrer Ergebnisse. Die Standarddiskussion ist nicht genuin eine methodologische Debatte auf der Grundlage von Videostudien, jedoch stellen Videodokumentationen einen zentralen Kristallisationspunkt innerhalb des Portfolioansatzes für die Untersuchung von Kompetenzprofilen als Ergebnis der Lehrerbildung dar.
Die kriteriumsorientierte Untersuchung der Ergebnisse im Sinne einer outputorientierten Analyse der Wirkungen der Lehrerbildung ist wesentlich durch den Standard-Ansatz befördert worden, weshalb diese Diskussionen hier dargestellt werden soll.
Verlängerte Praxisphasen innerhalb des Studiums betreffen einen sensiblen Diskussionspunkt der Lehrerbildung in Deutschland, weil die Praxisbegegnung im Phasenmodell institutionalisiert worden ist. Kapitel 6 geht deshalb der Frage nach der Beschaffenheit des Theorie-Praxis-Verhältnisses nach und will klären, welche Lernpotentiale Praxisbegegnungen bieten können bzw. wo man Grenzen vermuten darf. Anschließend werden im Kapitel 7 die empirischen Ergebnisse wichtiger internationaler und deutscher Untersuchungen dargestellt und diskutiert, weil sie einen empirischen Referenzrahmen für die eigene Studie abgeben und wichtige Impulse für die Konzeption der Studie und die Einordnung der eigenen Ergebnisse liefern.
Kapitel 8 erörtert kriterielle Ansätze zur Untersuchung und Definition von Unterrichtsqualität bzw. deren Perspektivenabhängigkeit, weil auf der Basis dieses Diskurses Rückschlüsse auf unterrichtliche Kompetenzen bzw. deren Entwicklung gezogen werden können. Dieses Kapitel markiert gleichzeitig das Ende des theoretischen Teils und bildet die Grundlage für die Entwicklung des Beobachtungsinstrumentariums der empirischen Untersuchung.
In der Reihenfolge der Kapitel des theoretischen Teils, die einer Linie von grundlegenden kategorialen Reflexionen über inhaltliche Fragen der Unterrichtsqualität zu wesentlichen Ergebnissen des Forschungsstands folgt, drückt sich eine sukzessive Annäherung an die Konzeptionalisierung der eigenen Studie und ihrer Methodik aus.
Während in Kapitel 9 die Forschungsfragestellung näher eingegrenzt wird, stellt Kapitel 10 einen methodologischen Überblick der videobasierten Unterrichtsforschung dar und will zudem den Prozess der Entwicklung des Instrumentariums transparent machen und im Theoriebezug darstellen.
Die Ergebnisdarstellung und -reflexion der Studie erfolgt eingehend in Kapitel 11. Den Abschluss der Arbeit bildet Kapitel 12, in dem versucht wird die Anliegen und Ergebnisse der Studie resümierend zusammenzufassen bzw. zu diskutieren und damit ein Fazit, aber auch einen Ausblick als Ergebnis der eigenen Arbeit zu geben.

2 Annäherungen an den Kompetenzbegriff

Zu Beginn der Studie wird der Versuch unternommen den Kompetenzbegriff in verschiedenen Annäherungen zu diskutieren. Es wird darum gehen, begriffsgeschichtliche Stränge herauszuarbeiten und die Begrifflichkeit aus ausgewählten disziplinären Perspektiven aufzuhellen sowie damit verbundene Konnotationen aufzuzeigen. Eine umfassende, alle Disziplinen und Positionen einschließende Begriffsanalyse erscheint hier allerdings kaum möglich, womit ein Anspruch auf Vollständigkeit nicht erhoben werden kann. Hingegen orientieren die begriffsanalytischen Betrachtungen darauf, den kategorialen Theorierahmen voraussetzend für die Konzeptionalisierung der eigenen Untersuchung zu erschließen und systematisch zu beschreiben.
Eine erste Annäherung liegt zunächst in etymologisch fundierten Reflexionen zum Kompetenzbegriff. Weiterhin stehen verschiedene disziplinäre Kompetenzmodelle im Fokus. Der Begriff als wissenschaftlicher Terminus etablierte sich über die sprachwissenschaftlichen Modellierungen von Chomski und wurde von Habermas in einem grundlagentheoretischen Diskurs weiterverarbeitet. Das von ihm geprägte Verständnis der „kommunikativen Kompetenz“ verdichtete Baacke für den Bereich pädagogischer Interaktionen und erweiterte es zugleich. Im Feld der Berufspädagogik hat Geißler wiederum den Begriff der „kritischen Kompetenz“ als eine zentrale Kategorie herausgearbeitet. Auffällig ist zudem, dass die in den 1970er Jahren beginnende Diskussion über Schlüsselqualifikationen Parallelen zur heutigen Kompetenzdiskussion aufweist, weshalb sie in diesem Kapitel ebenfalls nachvollzogen wird, um die historische Begriffsgenese rekonstruieren zu können. Des Weiteren lassen sich zeithistorische Diskurse zur „pädagogischen Handlungskompetenz“ aufzeigen, die den darzustellenden grundlagentheoretischen Rahmen verlassen und gezielt der Frage nachgehen, welche Kompetenzen in der pädagogischen Praxis benötigt werden. Schlussendlich wirft das Kapitel die Frage auf, inwieweit der Kompetenzbegriff den Bildungsbegriff abgelöst hat bzw. ablösen kann.

2.1 Etymologische Klärungen des Kompetenzbegriffes

Ein Blick in ein etymologisches Lexikon (Kluge 2002) liefert für den Kompetenzbegriff herleitend das Adjektiv „kompetent“ und übersetzt es gemäß seiner lateinischen Wurzel für die Zeit bis zum 18. Jh. mit „zusammentreffen, etwas gemeinsam erstreben, gesetzlich einfordern“, für die darauf folgende Zeit auch mit „zustehen, zukommen“. Das Abstraktum „Kompetenz“ kann sogar bis in das 16. Jh. zurückverfolgt werden und meinte ursprünglich „das Recht auf Einkunft“. Ein aktuelles Kompaktlexikon (Brockhaus 2002) fokussiert die Beschreibungen „Sachverstand“ und „Zuständigkeit“. Diese definitorischen Bestandteile lassen die Interpretation zu, dass eine Person Sachverstand benötigt, um Zuständigkeiten zugebilligt zu bekommen. Umgekehrt kann man folgern, dass sich aus Zuständigkeiten die Notwendigkeit von Sachkompetenz ergibt.[3]

[3] Der aus den etymologischen Herleitungen und der lexikalischen Definition ersichtliche Zusammenhang zwischen den Dimensionen „Sachverstand“ und „Zuständigkeit“ soll hier nicht

Frey (2006) diskutiert das lateinische Substantiv „competentia“ als Definitionsgrundlage und nennt, in Überschneidung zu vorangehend dargelegten Bedeutungszuschreibungen, die Übersetzungen „Zusammentreffen“ und „Zuständigkeit“. Zusätzlich benennt er das aus der Juristensprache stammende lateinische Adjektiv „competens“ und übersetzt es mit „maßgebend“ bzw. „befugt“. Ähnlich wie aus der aufgeführten lexikalischen Kurzinterpretation geschlossen werden kann, kommt er zu der Folgerung, dass „Verantwortung“ konstitutiv für den Begriff ist.

„Einer Person kann nur dann Kompetenz zugesprochen werden, wenn sie verantwortungsvoll mit sich selbst und der Umwelt umgeht und berufliche Fähigkeiten besitzt“ (Frey 2006, S. 126).

Aus den Übersetzungen des lateinischen Verbs „competere“ mit „zusammentreffen, zusammenlangen, stimmen, zutreffen, entsprechen, zukommen“ schließt Frey weiterhin, dass eine Person dann kompetent handelt, wenn die Anforderungen einer Situation mit dem individuellen Konglomerat von Fähigkeiten einer Person „zusammentreffen“ (ebd.).
Der Aspektreichtum der Diskussion und jeweilige bedeutungsbezogene Zuschreibungen, die hier anhand lexikalischer Einträge sowie des lateinischen Wortstammes verdeutlicht werden sollten, heben die Vielschichtigkeit des Kompetenzbegriffes und somit die Schwierigkeit einer annähernd einheitlichen Definition hervor. Gleichzeitig sind Definitionsfolien angesprochen, die konstitutiv für den Begriff sind und auch für die folgenden Ausführungen relevant bleiben.
Aufschluss verspricht in diesem Zusammenhang nicht zuletzt das Kompetenzverständnis von Vonken (2005), der den Begriff in der englischsprachigen Version „competence“ bzw. „competency“ diskutiert und zu ähnlichen wie den bereits aufgeführten Aussagen kommt. Auch er findet die Verknüpfung von „Befähigung, Tüchtigkeit und Können“ mit „Zuständigkeit und Befugnis“ (Vonken 2005, S. 17f.). Die im Kompetenzbegriff angelegte implizite Verbindung von „Befähigung“ und „Zuständigkeit“, die ja einen hohen Normativitätsgehalt dahingehend aufweist, wer sich überhaupt auf eigene Kompetenz berufen darf, liefert einen weiteren Hinweis für die heutige Aktualität des Begriffes. Der Wechselbezug von Fähigkeiten und Verantwortung bzw. fähigkeitslegitimierter Zuständigkeit verweist auf die geforderte adäquate Bearbeitung gesellschaftlich definierter Themen und Problemlagen. Diese bleiben eben nicht wie im Bildungsbegriff relativ abstrakt und konstituieren sich nicht erst auf dem Weg individueller Entfaltung. Vielmehr zielen Kompetenzen auf die Bewältigung und Bearbeitung bereits abgesteckter Situationen, die nach einer spezifischen Zuständigkeit verlangen.

überbeansprucht werden. Trotzdem ist es interessant, dass diese, z.T. historischen Verweise in der aktuellen Kompetenzdiskussion wiederum aufscheinen. Man kann die These aufstellen, dass der Kompetenzbegriff für die versuchte Durchsetzung einer begriffstheoretischen Deutungshoheit steht, die sich genau aus dem legitimatorischen Zusammenhang von Sachverstand und Zuständigkeit ergibt. Diese Perspektive verdeutlicht die verbreitet bekundete Attraktivität des Begriffes und seine Karriere.

2.2 Etablierung des Kompetenzbegriffes im wissenschaftlichen Sprachgebrauch

Die Etablierung des Kompetenzbegriffes im wissenschaftlichen Sprachgebrauch und damit als wissenschaftliche Kategorie kann auf den Linguistiker Noam Chomsky zurückgeführt werden. In seiner Sprachtheorie und -philosophie untersuchte er die Syntax menschlicher Sprech- und Verstehensakte und unterschied den Begriff der „Sprach-Kompetenz“ als humanistische Disposition vom eigentlichen Sprechakt als „Sprachverwendung“ (Chomsky 1970). Chomsky wollte mit seinem Modell klären, wie das Individuum zwar ständig grammatikalisch falsche Sätze hört, es ihm aber trotzdem gelingt, diese zu verstehen und damit eine innere Grammatik zu entwickeln, die es ihm erlaubt, neue Sätze zu generieren, die selbst wiederum grammatikalisch falsch sind, von anderen aber dennoch verstanden werden. Er stellte die These auf, dass es immanente und individuelle Kompetenzen gibt, die den Sprecher bzw. Hörer in die Lage versetzen, Sprechakte bzw. Verstehensakte auf der Basis eines regelgeleiteten syntaktischen Systems zu generieren. Die Sprachkompetenz manifestiert sich also in der Sprachperformanz, erschöpft sich damit aber nicht unbedingt. Mit der gesetzten Unterscheidung markierte Chomsky eine Gegenposition zu behavioristischen Ansätzen, die Sprech- und Verstehensakte im Rahmen von Stimulus-Respons-Theorien zu erklären versuchten. Sein Kompetenzbegriff beinhaltet zwar ein Moment von Entwicklung über die Lebenszeit, gleichwohl geht er aber von einem naturbasierten, genetischen Mechanismus aus.
Die Unterscheidung zwischen dispositorischen Kompetenzen und manifesten Performanzen lenkt den Blick sowohl auf das Individuum als kreativen Urheber von Sprech- und Verstehensleistungen als auch auf den sozialen Prozess der Kommunikation als Interaktion. Der linguistische Ansatz von Chomsky lieferte somit wichtige Impulse für die humanistische Psychologie und Erziehungswissenschaft (Oerter 2000). Eine Differenzierung zwischen Kompetenz und Performanz ist seitdem konstitutiv für jede Kompetenztheorie.

2.3 Kommunikative Kompetenz

Der oben dargestellte Kompetenzansatz von Chomsky ist in den Sozialwissenschaften und in der Erziehungswissenschaft weiterentwickelt worden. Die von Habermas (1987a, 1987b) vorgelegte Theorie des kommunikativen Handelns bzw. der Begriff der „kommunikativen Kompetenz“ (Habermas 1990) ergänzen Chomskys Ansatz durch eine stärkere Fokussierung auf das soziale Handlungsgefüge von Kommunikation. Die Theorie der kommunikativen Kompetenz strebt eine Analyse und Rekonstruktion des Regelsystems an, nach dem kommunikative Situationen ermöglicht oder hergestellt werden und will nicht nur die linguistische Syntax von Sprech- und Verstehensakten einer Untersuchung zugänglich machen. Habermas formuliert so einen weiter gefassten Kompetenzbegriff, der individuelle und interaktive Facetten beinhaltet und darüber hinaus soziale Generierungsleistungen enthält. Die Habermassche Position spielt nicht zuletzt aufgrund ihrer idealistisch-abstrakten Ausrichtung in der aktuellen erziehungswissenschaftlichen und psychologischen Kompetenzdiskussion keine

gewichtige Rolle mehr. Es gibt aber Ansätze, seine Theorie auf das Feld Schule und unterrichtliche Lehr-Lernsituationen anzuwenden (vgl. Girmes 2006a, b). Berührungspunkte finden sich auch in soziologischen Rollentheorien oder dem Habitus-Konzept von Bourdieu (1994). Die Betonung des sozialen Rollengefüges bleibt ein Verdienst von Habermas.
Für die Erziehungswissenschaft hat Baacke (1980) die Theorie der kommunikativen Kompetenz weiterentwickelt und fruchtbar gemacht. Ebenfalls in Anlehnung an Chomsky erweitert er das Verständnis von Kommunikation und damit kommunikativer Kompetenz um Elemente des Verhaltens bzw. Handelns. Kommunikation bleibt analog zu Habermas nicht auf Sprech- und Verstehensakte bzw. soziale Interaktion beschränkt, sondern bezieht sich auf das konkrete Verhalten. Dieses wiederum versteht sich nicht nur als intendierte Handlungselemente; hingegen zielt Verhalten ganzheitlich auf die Gesamtschau eines Individuums. Dabei hält auch Baacke an der Trennung von Kompetenz und Performanz fest. Kommunikative Kompetenz betrachtet er als Verbindung von Sprach- und Verhaltenskompetenz (vgl. Baacke 1980, S. 262).
Der Autor geht von einem epistemologischen Subjektbegriff[4] aus, womit auch Verhaltensfreiheiten eingeschlossen sind. Das Individuum ist danach in der Lage seine Verhaltensweisen eigentätig zu generieren, auch wenn diese durch Sozialisationsinstanzen beeinflusst sind und nur jeweils in einem vorgefundenen kulturellen Raum realisiert werden können. Dieses Moment der Selbstbestimmtheit – neben einer sozialisatorischen und interaktionalen Bedingtheit – weist für ein pädagogisches, subjektorientiertes Selbstverständnis eine besondere Attraktivität auf.

„Es wird behauptet, der Mensch könne auch seine Verhaltensschemata ‚generieren', und zwar in der Aktualisierung einer Verhaltenskompetenz, die den inneren motivationalen Lagen des Individuums zur Disposition steht. Ebenso wie Sprache soziokulturellen Überformungen unterliegt, so auch Verhalten in der Kommunikation; aber selbst diese sind nur als das in Systemen institutionalisierte Resultat von kommunikativen Handlungen zu verstehen" (Baacke 1980, S. 261f.; zit. n. Vonken 2005, S. 24).

Speziell für pädagogische Interaktionen, so z.B. Lehr- und Lernprozesse, erscheint Baackes Ansatz weiterführend, indem er Sprach- und Verhaltenskompetenz integriert.

Sowohl Habermas als auch Baacke haben Chomskys Theorie weitergeführt und für die Analyse von pädagogischen Situationen aufgeschlossen. Gemeinsam ist beiden Ansätzen, dass sie kommunikative Kompetenz als Erweiterung von Sprachkompetenz fassen und ein Moment des (kommunikativen) Handelns mit einbeziehen. Danach bedeutet Kompetenz die Fähigkeit des Individuums zur Generierung von kommunikativen Situationen (Habermas) bzw. Verhalten (Baacke), also zur Erzeugung von Interaktion (Vonken 2005, S. 24f.).

[4] Diese Herangehensweise ist konstitutiv für die Erschließung des Kompetenzbegriffes als Erklärungskategorie für pädagogische Prozesse.

2.4 Kritische Kompetenz

Im Feld der Berufspädagogik konnte durch Arbeiten von Geißler (Geißler 1974; Geißler & Orthey 2002) der Begriff der „kritischen Kompetenz“ etabliert werden. Der Autor orientiert sich an Habermas und steht in einer Entwicklungslinie der 1970er Jahre, wo das Verhältnis von beruflich orientierter und subjektorientierter Bildung neu diskutiert wurde. Anlehnungen zur kritischen Theorie der sog. Frankfurter Schule sind hier unverkennbar. In diesem Sinne ist der Kritikbegriff umfassend gemeint und beinhaltet neben der individuellen auch eine gesellschaftliche Dimension mit dem Ziel von gesellschaftlichen Veränderungen.
Kritische Kompetenz lässt sich nach Geißler in die Fähigkeit zur Kritik als interaktionistisches Handlungselement und das Wissen um Verfahrensweisen des Kritisierens unterscheiden. Analog zu Habermas macht erst die Integration beider Dimensionen kritische Kompetenz aus. Die Betonung legt Geißler dabei auf die Differenzierung von Fähigkeits- und Wissensdimensionen.

„Kritikkompetenz ist demnach mehr als nur das Wissen um Verfahrensweisen (wie es z.B. die linguistische Kompetenz im Chomskyschen Sinne darstellt) des Kritisierens. Sie ist gleichsam auch die Fähigkeit zu kritischer Praxis unter der Maßgabe, den eigenen Bedingungsrahmen für Kritik hervorzubringen“ (Geißler 1974, S. 34).

Darüber hinaus unterteilt er Kritikkompetenz in „kritisch-reflexive Kompetenz“, „kritisch-soziale Kompetenz“ und „kritisch-instrumentelle Kompetenz“. Sein Kompetenzbegriff umschließt damit Dimensionen, die auf der individuellen, gesellschaftlichen und methodischen Ebene angesiedelt sind. Insofern ist seine Unterteilung auch heute noch anschlussfähig, wenn auch Aspekte der gesellschaftlichen Transformation im aktuellen Diskurs um schulische Kompetenzen weniger im Vordergrund stehen. Prinzipiell kann in der Rückschau auf Geißlers kritisches Kompetenzverständnis gesagt werden, dass die Frage nach dem Verhältnis zwischen beruflicher Ausbildung und individueller Bildung nach wie vor virulent ist.

Die Kompetenzdimensionen von Chomsky, Habermas, Baacke und Geißler weisen Schnittstellen zueinander auf. Sie betonen individuell-kreative Prozessmerkmale in der interaktiven, zwischenmenschlichen Auseinandersetzung. Während Chomsky als Linguist auf der sprachtheoretischen Ebene verhaftet bleibt, schließt Habermas den Blick für die Genese kommunikativer Prozesse als Kompetenzmerkmal auf. Baacke öffnet den Kommunikationsbegriff wiederum und erweitert ihn um Verhaltensanteile. Geißler stellt den Kritikbegriff in den Mittelpunkt seiner Kompetenzdefinition und macht damit zum einen die gesellschaftliche Eingebundenheit individueller Entwicklung deutlich, zum anderen orientiert er auf die Notwendigkeit gesellschaftlicher Transformation im Sinne der kritischen Theorie.

2.5 Schlüsselqualifikationen

Der Begriff der Schlüsselqualifikation zeugt von Anknüpfungspunkten zum Kompetenzbegriff und wird teilweise synonym verwendet. Auch er scheint einen vermeintlichen Sinnüberschuss zu enthalten, der ein Erklärungsmuster für die Begriffskarriere bietet. Kritisch äußern sich dazu Arnold & Müller (2006, S. 7):

„Vergleicht man den spezifischen Nimbus, der den Schlüsselqualifikationen im Vergleich zu den Fachqualifikationen in den verschiedenen Bereichen des Bildungssystems als eine Art Aura anhängig zu sein scheint, so könnte man die Fachqualifikationen als exakt markierte und leicht erkennbare Schiffe im Meer des Wissens bezeichnen, während die Schlüsselqualifikationen sich darin eher wie nur schwer verortbare und geheimnisumwitterte ‚Fliegende Holländer' bewegen."

Der Begriff „Schlüsselqualifikation" geht auf einen Impulsaufsatz von Mertens (1974) zurück, der damals Leiter des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit war. Mertens wollte mit dem Begriff die Modernisierung beruflicher Ausbildungskonzepte vorantreiben, um den Anschluss an sich verändernde berufliche Anforderungsprofile gewährleisten zu können. Gleichzeitig intendierte er mit seinem übergreifenden Entwurf die Annäherung traditionell getrennter Bereiche der beruflichen, allgemeinen und politischen Bildung, da diese aufgrund gesamtgesellschaftlicher Entwicklungen bezüglich ihrer Lernziele teilweise konvergierten und insofern enger miteinander zu verzahnen waren. Mertens ging von der Ausgangsthese aus, dass Bildungsangebote im Zuge der Anforderungen einer sich verändernden Gesellschaft drei Dimensionen gerecht werden müssen (vgl. Siebert 2006):

- der Schulung zur Bewältigung und Entfaltung der eigenen Persönlichkeit,
- der Schulung zur Fundierung der beruflichen Existenz und
- der Schulung zu gesellschaftlichem Verhalten.

Seiner Auffassung nach sollten alle Bildungs- und Ausbildungscurricula, Lernziele und Lehrinhalte diesen Dimensionen in integrativen Konzepten folgen.

Vergleicht man die Dimensionen von Mertens mit aktuellen Diskussionen, wie sie im 2006 (Klieme et al.) im Auftrag der Kultusministerkonferenz und dem Bundesministerium für Bildung und Forschung herausgegebenen Bildungsbericht festgehalten worden sind, so kommt man zu der erstaunlichen Feststellung, dass dort folgende übergreifende Zielsetzungen des deutschen Bildungssystems benannt werden:

- Selbstregulation
- Human Ressource
- Gesellschaftliche Partizipation

Wenn die Übereinstimmung zum Mertensschen Entwurf angesichts der allgemein proklamierten Entwicklungssprünge und Veränderungen auf den ersten Blick überrascht, so muss man in zweiter Annäherung feststellen, dass mit den genannten

Dimensionen offenbar ein Konsens hergestellt worden ist, welcher sich seit 1974 hinsichtlich seines Bedeutungsgehaltes nicht wesentlich verändert hat. Umgekehrt kann man fragen, ob hiermit ein Anspruch an die Bildungs- und Ausbildungssysteme manifestiert wird, der bis dato nicht eingelöst werden konnte, womit eine Erklärung für die nach wie vor bestehende Aktualität der vor mehr als dreißig Jahren explizierten Einforderungen vorläge. Zudem sind die Zieldimensionen sowohl von Mertens als auch im Bildungsbericht auf einem hohen Abstraktionsniveau formuliert, was einen weiteren Erklärungsfaktor für nahe liegend identifizierbare Überschneidungen ausmacht. Die Notwendigkeit des hohen Abstraktionsniveaus begründete Mertens mit der Dynamisierung gesellschaftlicher Entwicklungen, einem gestiegenen kognitiven Anspruchsniveau in den verschiedenen Berufsfeldern sowie der gewachsenen Obsoleszenzrate des Wissens.

„Es kann die Hypothese vertreten werden, dass das Obsoleszenztempo (Verfallzeit, Veraltenstempo) von Bildungsinhalten positiv mit ihrer Praxisnähe und negativ mit ihrem Abstraktionsniveau korreliert“ (Mertens 1977, S. 109).

Aktuell beschreibt Winkler (2003) das Obsoleszenzphänomen im Rückgriff auf die Formulierung des englischen Soziologen Zygmunt Bauman mit dem Begriff der „liquid society“ (Bauman 2000) und knüpft damit an Mertens an.

„Das beruflich erforderliche Wissen und die nötigen Fähigkeiten tendieren zur Flüssigkeit, wobei sich die Fließgeschwindigkeit zunehmend beschleunigt“ (Winkler 2003, S. 4).

Mertens hat in seinem Schlüsselqualifikations-Konzept ein bis dahin bestehendes Qualifikationsverständnis erweitert. Die sich verändernden Berufsanforderungen, so seine These, forderten zunehmend nicht nur fachliches Wissen, sondern auch überfachliche Qualifikationen wie Teamfähigkeit, Selbstorganisation und lebenslanges Lernen. Konkret benennt Mertens folgende Schlüsselqualifikationen:

„Förderung der Fähigkeit zu lebenslangem Lernen und zum Wechsel sozialer Rollen, Distanzierung durch Theoretisierung, Kreativität, Relativierung, Verknüpfung von Theorie und Praxis, Technikverständnis, Interessenanalyse, gesellschaftswissenschaftliches Grundverständnis, Planungsfähigkeit, Befähigung zur Kommunikation, Dekodierungsfähigkeit, Fähigkeit hinzuzulernen, Zeit und Mittel einzuteilen, sich Ziele zu setzen, Fähigkeit zur Zusammenarbeit, zur Ausdauer, zur Konzentration, zur Genauigkeit, zur rationalen Austragung von Konflikten, zur Mitverantwortung, zur Verminderung von Entfremdung, Leistungsfreude“ (Mertens 1974, S. 40).

Es handelt sich um einen sehr breiten Anforderungskatalog, der individuelle Entwicklungsmöglichkeiten berücksichtigt, aber doch eine starke Affinität zu beruflichen Anforderungsprofilen aufweist.
Vor dem Hintergrund dynamisierter Prozesse gesellschaftlichen und wirtschaftlichen Wandels ging es nicht mehr darum den Einzelnen an die Bedingungen anzupassen, sondern es sollte ihm selbst die Fähigkeit zur Anpassung an den Wandel vermittelt

bzw. ermöglicht werden.[5] Dem Konzept der Schlüsselkompetenz haftete damit der Ruf der Instrumentalisierung von Lehr-Lernprozessen an und es wurde entsprechend kritisch gewürdigt (Kösel & Scherer 2006; Pätzold & Wortmann 2006). Umgekehrt musste anerkannt werden, dass das Konzept sich nicht der beruflichen Praxis verschließt. Zudem macht es noch in gegenwärtigen Lesarten auf die subjektive Seite von Qualifikationsprozessen aufmerksam. Winkler (2003) spricht in diesem Zusammenhang von einer Subjektivierung beruflicher Anforderungen, da hier immer stärker Handlungssituationen vorzufinden seien, die zunehmend hoch flexibler und verantwortlicher Subjekte bedürfen.
Die Subjektivierung beruflicher Anforderungsprofile und eine darauf antwortende subjektivierende Ausbildungskonzeption in Form der Konzentration auf Schlüsselqualifikationen enthalten Spannungselemente hinsichtlich der Gewichtung individueller und sozialer Entwicklungen. Ansätze von Schlüsselqualifikationen stehen somit im Spannungsfeld von individueller Subjektentwicklung und fremdgesteuerter Anforderungsdynamik. Inwieweit die Pole einen gelingenden Ausgleich im Sinne von erweiterter Persönlichkeitsentwicklung zulassen, kann hier nicht beantwortet werden. Schon frühzeitig hat Mertens auf die selbst gestellte Frage „Was ist aus den damaligen Anstößen geworden?“ eher resigniert festgestellt:

„Eigentlich nicht sehr viel mehr als eine abstrakte Diskussion, viel verbale Zustimmung mit wenig Umsetzung, überhaupt wenig Handfestes“ (Mertens 1988, S. 43).

Die grundsätzliche Tendenz zur Förderung der Ausbildung von Fähigkeiten und Fertigkeiten mit Richtung auf zu erweckende Transferpotentiale verweist auf den Kern der Schlüsselqualifikationsdebatte, nämlich auf die Ermöglichung von Erschließungskompetenzen bzw. von Lernfähigkeit an sich (Tietgens 1992). Diese Transferpotentiale, näher bestimmt als laterale (horizontale) oder vertikale Kompetenzen (Ausubel et al. 1980, S. 206; Gagné 1975, S. 264), konnten bislang in lern- und bildungspraktischen Kontexten zu wenig eingelöst werden.
Vor dem Hintergrund der älteren Diskussion um „formale Bildungsprozesse“ (die deshalb formal genannt wurden, weil sie inhaltsübergreifend gedacht waren) überraschen diese Ergebnisse weniger (Schwenk 1989, S. 218f.; Weinert & Schrader 1997, S. 302ff.). In der Schlüsselqualifikationsdebatte wurde die Möglichkeit formaler Bildungsprozesse rekontextualisiert, woraus man schließen kann, dass es bestimmte pädagogische Topoi gibt, die sich überzeitlich halten und von neuen Generationen mit neuen Begrifflichkeiten appellativ formuliert werden. Nicht bestritten werden soll, dass laterale Transfereffekte (Gelerntes wird in demselben Gegenstandsbereich genutzt) bzw. vertikale Transfereffekte (Gelerntes wird in neuen Kontexten oder Situationen genutzt) beobachtbar sind und auch gefördert werden können (Dörr 1991a, b). Die Vorstellung extrem breiter Lern- bzw. Transferwirkungen, wie sie konstitutiv für die Theorie der formalen Bildung ist und auch in der Schlüsselqualifikationsdebatte durchscheint, ist mittlerweile jedoch wissenschaftlich obsolet geworden (Schmid 2006, S. 284).

[5] Mit verlagerten Verantwortlichkeiten für Aneignungsprozesse auf die Seite der Lernenden bahnt sich das an, was später und unter dem Einfluss konstruktivistischer Ausrichtungen als Ermöglichungsdidaktik proklamiert wird (vgl. Arnold 2007).

2.6 Transferpotentiale und Schlüsselqualifikationen

Der Erschließungsbegriff als Kernkategorie von Schlüsselqualifikationen lenkt den Blick, wie angedeutet, auf damit verbundene Transferpotentiale. Mit dieser Konstruktion einhergehende Schwierigkeiten sind jedoch z.T. unterschätzt worden.
Messner (1978) macht auf die Notwendigkeit aufmerksam, zwischen Transfer und Transfereffekt zu unterscheiden. Transfer meint den Prozess der Übertragung von Lernleistungen auf unbekannte Aufgabenformen. Der Transfereffekt markiert das Ergebnis dieser Übertragung. Man kann dabei wiederum zwischen spezifischen und unspezifischen Transfereffekten unterscheiden. Unspezifische Transfereffekte können dann festgestellt werden, wenn insgesamt eine Verbesserung von Gedächtnisleistungen, logischem Denken und Problemlösefähigkeit auszumachen ist. Spezifische Transfereffekte bedingen eine inhaltliche (deklarative) oder strukturelle (prozedurale) mentale Wissensrepräsentation, die Analogien zwischen – bekannten und neuen – Aufgabentypen abbildet. Strukturen im Rahmen von strukturellen Präsentationen können als Elemente und deren wechselseitige Relationen ausgemacht werden. Als zentrale heuristische Strategie ist hier die Analogiebildung zu nennen. Kondakov (1983, S. 28; zit. n. Dörr 1991a, S. 29) beschreibt die Analogiebildung als

„das Schließen vom Modell auf das Original. … Der Gegenstand, der unmittelbares Untersuchungsobjekt ist, wird Modell genannt, und der Gegenstand, auf den die durch Untersuchung des Modells erlangte Information übertragen wird, heißt Objekt, Original oder Prototyp. Davon ausgehend wird Analogie als Relation zwischen einem Modell und seinem Original, seinem Prototyp definiert."

Die von Kondakov herausgestellten Relationen in der Analogiebildung lassen eine Unterscheidung in horizontale und vertikale Relationen zu (Hesse 1970). Horizontale Relationen stellen solche der Identität oder Unähnlichkeit dar und lassen sich allgemein als Ähnlichkeitsrelationen beschreiben (Dörr 1991a, S. 35). Vertikale Relationen stehen dagegen für die Verknüpfung von Elementen innerhalb eines Systems, die kausal attribuiert werden können. Bei maximaler Ähnlichkeit (Identität) bzw. Unähnlichkeit (es können keine gemeinsamen Merkmale gefunden werden) fallen die beiden Relationsdimensionen als Analogietypen zusammen und sind nicht voneinander zu unterscheiden.
Die Identifikation von Analogien zwischen zwei Systemen verweist auf das kognitionstheoretische Konzept des „Schemas". Hierunter werden generalisierende Abstraktionen verstanden, die auf die zu vergleichenden Systeme bezogen sind und Aussagen über strukturelle Ähnlichkeits- oder Unähnlichkeitskonzepte enthalten.
Schemata können durch verschiedene externe Modelle oder interne mentale Modelle repräsentiert werden. In der Beschreibung des Zusammenhanges von mentalen Modellen und Schemata als Grundlage für Analogiebildungen kommt Dörr (1991a, S. 43) zu dem Schluss, „dass Analogiebildung keine Prozedur darstellt, die zur Bildung mentaler Modelle dient, sondern auf der Ebene der Repräsentation dazu dient, von einzelnen Modellen zu abstrahieren, um abstrakte Konzepte (Schemata) zu generieren". In drei empirischen Pilotstudien konnte er den Nachweis führen, dass

Transfereffekte zwischen inhaltlich unterschiedlichen, jedoch strukturell isomorphen Problemen nur dann zustande kommen, wenn ein hinreichend abstraktes Schema aufgebaut worden ist. Sind die Schemata bezüglich ihrer abstrakten Reichweite eingeengt oder werden nur mentale Modelle des Problemlösungsprozesses gebildet, gelingt der Transfer nicht. Die Notwendigkeit der Herausbildung abstrakter Schemata als Grundlage für die Möglichkeit von Transferprozessen erklärt, warum das übergreifende Konzept der Schlüsselqualifikationen letztlich kaum neue Ergebnisse hervorgebracht hat; schließlich verweist die Entwicklung von Schemata auf vorgeschaltete komplexe Lernprozesse, die eine hohe Bereichsspezifik aufweisen. Auch Wellenreuther (2004, S. 112) sieht die praxisorientierte und übungsintensive Automatisierung von Schemata als wesentliche Voraussetzung für gelingenden Transfer.

Die Forschungsergebnisse von Anderson (vgl. Kap. 3.7) zu sog. ACT[6]-Theorien lassen an der Gewichtigkeit von Schlüsselqualifikations-Konzepten zweifeln. Der Autor hat ein hierarchisches Stufenmodell vorgelegt, das die Entwicklung prozeduralen Wissens[7] aus deklarativem Wissen als schrittweise Prozeduralisierung darstellt (Anderson 1982, 1983, 1987). Ein zentrales Ergebnis dieser Forschungsrichtung lautet, dass kognitive Transferleistungen ausschließlich in Form von Prozeduren, nicht aber in Form von Faktenwissen auf neue Lernsituationen übertragen werden können. Prozedurales Wissen basiert damit auf komplexen Lernprozessen, für die deklaratives Wissen immerhin eine wichtige Voraussetzung darstellt. Die eher zweifelhaften Erfolge des Schlüsselqualifikationskonzeptes lassen sich hypothetisch mit diesem Zusammenhang erklären. Deklarative Wissenstypen bieten nur ein geringes, allerdings ein durchaus auszumachendes Transferpotential, wie Franzke (1996) in ihrer Kritik an Anderson festhält.
Dieses geringe Transferpotential rechtfertigt aber nicht ein so ambitioniertes Konzept wie das der Schlüsselqualifikationen. Es erscheint theoretisch und empirisch evident, dass prozedurale Wissenstypen ein höheres Transferpotential im Vergleich zu deklarativen Wissenstypen aufweisen, verfügen sie doch über elaboriertere Wissenskonzepte bzw. abstraktere Schemata, womit erst die „Flüssigkeit grundlegender Fertigkeiten“ (Wellenreuther 2004, S. 140) ermöglicht wird.

Das Konzept der Schlüsselqualifikationen darf vor dem Hintergrund dargelegter Befunde nicht unbedingt als gescheitert, jedoch als überholt angesehen werden. Die damit verbundenen Zielsetzungen und Potentiale, insbesondere in Form von Transferleistungen, sind aber im Kontingenztheorem (vgl. Kap. 6.3) aufgegangen und beeinflussen insofern noch die aktuelle Debatte.

2.7 Pädagogische Handlungskompetenz

Ausgangspunkt für die Formulierung von pädagogischen Handlungskompetenzen seit Ende der 1970er, Anfang der 1980er Jahre war das Theorie-Praxis-Problem in der Erziehungswissenschaft. Aus Sicht der erziehungswissenschaftlichen

[6] Adaptive Control of Thought.
[7] Hier als analoger Terminus zum Begriff des „Schemas“ zu verstehen.

Theoriebildung erscheint die pädagogische Praxis häufig defizitär. Vor diesem Hintergrund ergab sich eine wechselseitige Auseinandersetzung zwischen pädagogischer Praxis und wissenschaftlicher Theoriebildung.
Einen Überblick zu Konzepten pädagogischer Handlungskompetenz gibt Dieterich (1983). Die verschiedenen Ansätze konzeptualisieren je spezifische Schwerpunktsetzungen. Übergreifend lässt sich aber eine Anlehnung an den von Baacke formulierten Kompetenzbegriff ausmachen, mit einer Schwerpunktsetzung auf Fragen der Interaktion zwischen kompetenten Pädagogen und einer näher zu bestimmenden Klientel. Die Frage des Theorie-Praxis-Problems spiegelt sich in der Verbindung von Wissen und Können, die allen Ansätzen anhaftet. Pädagogische Handlungskompetenz erscheint hier nicht voraussetzungslos, sondern wird an der Art des Wissens und seiner Verwendung festgemacht. Untersuchen lässt sich der mehr oder weniger kompetente Pädagoge, die pädagogische Situation oder die Interaktion.

Dieterich formuliert unter expliziter Bezugnahme auf Bedingungen pädagogischer Berufspraxis einen Kompetenzbegriff, der sich aus den Facetten „personenexterne", „personeninterne" und „interaktionistische" Bestandteile zusammensetzt. Personenexterne Faktoren zielen auf einen jeweils vorgefundenen Handlungsspielraum, der unter bestimmten Bedingungen und Zuständigkeiten gestaltet werden muss. An dieser Facette verdeutlicht Dieterich, dass die Anlegung von Gütemaßstäben an das Handeln von Pädagogen nicht ausschließlich normativ durch theoretisch formulierte Konzepte, sondern als „Best Practice" aus dem Handeln selbst heraus erfolgt. Diese pragmatische Sichtweise genügt zwar nicht den Ansprüchen wissenschaftlicher Theoriebildung, fokussiert aber den Blick auf reale pädagogische Praxis, macht sie beobachtbar und bietet damit Möglichkeiten der Verbesserung. Die personenexternen Facetten des Kompetenzkonstruktes stellen so einen Analyserahmen zur Verfügung, liefern jedoch keine deduktiven Aussagen zu Kompetenzen pädagogischen Handelns im Sinne von „Befähigung" (Vonken 2005, S. 28).
Personeninterne Kompetenzkonstrukte gehen davon aus, dass sich Kompetenzen an zugrunde liegenden Kognitionen festmachen lassen, die sich im Handeln manifestieren.

„Kompetenz ließe sich dann beschreiben als innerpsychische Voraussetzungen für kompetentes Handeln, als handlungsleitende Kognition" (ebd.).

Eine solche kognitionspsychologische Herangehensweise berücksichtigt sowohl fachwissenschaftliche Kenntnisse als auch ihre Synthese im Sinne anwendungsrelevanten Wissens. Ähnlich wie bei Chomsky kann hier zwischen Kompetenz und Performanz unterschieden werden. Angenommen wird damit auch, dass es ein kognitives Regelsystem gibt, welches Handlungen – indem situationsadäquat mit dem Regelsystem umgegangen wird – bezüglich der ihnen zugrunde liegenden Kompetenzen unterscheidbar macht.
Neben kognitionsorientierten Ansätzen finden sich in der Literatur auch Konstrukte, die Kompetenzdefinitionen an Persönlichkeitsmerkmale anlegen. Der Begriff der „Lehrerpersönlichkeit" (vergl. Kap. 4.1.1) hat sich bis heute gehalten. Grundsätzlich

erscheint eine Trennung von Kognitionen und Persönlichkeitsmerkmalen als Tellbereiche personeninterner Kompetenzbereiche nicht unproblematisch. Allerdings öffnet sie den Blick dafür, dass personeninterne Kompetenzkonstrukte in kognitiven Ansätzen durch Persönlichkeitsmerkmale beeinflusst sind. Dies zeigt sich z.B. an der Realisierung oder Nichtrealisierung bestimmter Performanzen.
Dieterich präsentiert auch Ansätze, die eine Integration von personenexternen und -internen Positionen intendieren. Er bezeichnet sie als „interaktionistische" Ansätze, weil sie die Verwobenheit und den wechselseitigen Bezug von feldspezifischen Handlungsmöglichkeiten und individuell gesteuerten Handlungen untersuchen.

„Sie stellen somit eine Verbindung aus kognitiven Prozessen und Verarbeitung der Umwelt durch den Handelnden dar. Hierbei wird betont, dass kompetentes Handeln neben dem kognitiven Aspekt der Abarbeitung eines (isolierten) Handlungsplanes auch das Verarbeiten von Umwelteinflüssen jenseits von Reiz-Reaktions-Schemata als antizipiertes Einschätzen von Handlungsfolgen bedeutet" (Vonken 2005, S. 29).

In der genaueren Beschäftigung mit personenexternen und personeninternen Merkmalsbereichen für die Konstruktion pädagogischer Handlungskompetenz sind wesentliche Unterschiede in den Betrachtungsweisen festzustellen. Erstere schärfen eher den Blick für Verhaltensweisen, die auf Fremd- oder Selbstattribuierungen basieren und damit den Charakter von Projektionen und Initiationen bekommen, ohne dass Aussagen über theoriebasierte Gütemaßstäbe getroffen werden können. Fortführende Analysen dieses Zugangs finden sich in der sozialwissenschaftlichen Rollentheorie. Personeninterne Merkmalsbereiche unterliegen in der Untersuchung von Kognitionen und Persönlichkeitsmerkmalen als Einflussgrößen auf kompetentes Handeln in der aktuellen Kompetenzdiskussion einer nach wie vor bestehenden Bedeutsamkeit. Es hat diesbezüglich eine verfeinerte Konstruktentwicklung in der empirischen Bildungsforschung stattgefunden. In Fortführung der genannten interaktionistischen Konzepte werden hier auch Felddaten, z.B. sozioökonomische Daten, erhoben, um den Systemcharakter von Schule und Unterricht und die Eingebundenheit in eine soziale Umwelt zu berücksichtigen (Rutter et al. 1980).

2.8 Kompetenzentwicklung statt Bildung?

Die bereits in den vergangenen Jahrzehnten und aktuell besonders deutlich auszumachende Durchsetzungskraft der Kompetenzdebatte ist auch in ihrem Verhältnis und in Absetzung zu vorangehend dominierenden Bildungskonzepten zu sehen. Danach tritt, so kann vermutet werden, der Kompetenzbegriff an die Stelle des Bildungsbegriffes und setzt gegenüber diesem etwa Lernziele sowie Anforderungen an pädagogisch tätige Personen und Lernende neu. Als – öffentlich, bildungspolitisch, erziehungswissenschaftlich und didaktisch diskutierte – Begründungen einer solchen Begriffsentwicklung halten gesellschaftliche Dynamisierungs- und Globalisierungsbewegungen und in diesem Zusammenhang der Anpassungsdruck an international durchgesetzte Beschreibungs- und Messgrößen her. Verwiesen sei hier nur auf das in Deutschland durch die PISA-Studien an Selbstverständlichkeit gewonnene „Literacy-Konzept" (vgl. z.B. Richter

2006). Trotz der in der Erziehungswissenschaft bereits auf die empirische Wende zurückgehenden Abwendung von Bildungsverständnissen durchdringen sie aufgrund ihres starken traditionellen und disziplinbezogenen identifikatorischen Gehaltes nach wie vor den aktuellen Kompetenzdiskurs.

Der Bildungsbegriff ist seit den 1960er Jahren, auch gerade in der Pädagogik, kritisch hinterfragt worden. Er war unter Ideologieverdacht geraten, womit Bildung vorrangig noch als funktionales Mittel zur Sicherung der Privilegien eines Bildungsbürgertums erschien. In entsprechend instrumentalismuskritischer Absicht prägte Adorno bereits 1962 den Begriff der „Halbbildung". Prinzipiell können über Bildung kaum überzeitliche Aussagen getroffen werden (Benner & Brüggen 2004, S. 174), sondern das jeweilige Verständnis ist abhängig vom kulturellen Kontext, von weltanschaulichen sowie Orientierungen hinsichtlich wissenschaftlicher Richtungen und Methoden.[8] Tenorth (1997, S. 972) konstatiert, dass Bildung seit Humboldt ein Thema habe, nämlich die „Subjekt-Welt-Relation". Diese weite thematische Fassung verweist auf die Heterogenität des Bildungsbegriffes in verschiedensten disziplinären Annäherungen. In der Konzentration auf pädagogische Zielgruppen können dennoch einige übergreifende Feststellungen getroffen werden.
Auch wenn Schorb (1975, S. 41f.) der Auffassung ist, dass Bildung hauptsächlich im Jugendalter stattfindet, trifft man vorrangig auf ein Verständnis, Bildung gehe über die Schulbildung hinaus und stelle einen lebenslangen Prozess dar (vgl. u.a. Schröder 1992, S. 38; Köck 1997, S. 99). Anders gesagt ist Bildung Ergebnis eines zeitlich unbegrenzten Bemühens des einzelnen Menschen (vgl. Schwenk 1989, S. 216). Im Unterschied zur Erziehung stellt Bildung so nicht nur einen Prozess, sondern auch ein Resultat dar (Langewand 2002, S. 69; Koch 1999, S. 78; Keller & Novak 1993, S. 64; Köck 1997, S. 98; Ehrenspeck 2004, S. 66).
Nach Schaub & Zenke (2000, S. 95) erhält der Bildungsbegriff noch heute seine Bedeutung durch die Aufklärungsphilosophie und die Emanzipation des Bürgertums von feudaler und kirchlicher Bevormundung. Äquivalent zum Erziehungsbegriff basiert ein Bildungsverständnis so auf Annahmen, dass der Mensch vernunftbegabt und zu Freiheit und Mündigkeit aufgerufen ist. Entsprechend steht Bildung bei Langewand (2002, S. 69) für die Verschränkung der Individualität von Personen mit „abstrakten Formen vernünftiger, philosophischer, wissenschaftlicher oder ästhetischer Selbst- und Weltdeutung", wobei sich ein Selbst- und Weltverständnis universal durch Kategorien wie Vernunft, Rationalität, Humanität, Sittlichkeit oder Werte bestimmt. Bildungsziel ist die Entfaltung der Persönlichkeit, die sich kritisch und verantwortungsvoll in Bezug zu sich selbst und zu seiner Umwelt verhält (vgl. Keller & Novak 1993, S. 63f.; Köck 1997, S. 99). Voraussetzend geht es in der Bildungsarbeit, wie in der Erziehung, um die Förderung von Urteilskraft und Handlungsfähigkeit.
Konkrete Bildungsinhalte ergeben sich aus zeitkulturellen und lebensgeschichtlichen Zusammenhängen (vgl. Schaub & Zenke 2000, S. 96). Sie werden bspw. durch Wissenschaft, Technik, Kunst, Literatur, Politik oder Ökonomie vorgegeben. Dabei gilt die Möglichkeit eines verbindlichen Bildungskanons in einer pluralen Gesellschaft als verworfen (vgl. Köck 1997, S. 99).

[8] Ein Sachverhalt, der auch auf den Kompetenzbegriff zutrifft.

Während für die Aufklärungszeit Bildung didaktisch noch vorwiegend mit Anleitung einhergeht, verweist ein neuhumanistisches Bildungsverständnis auf die Selbstentfaltung und -gestaltung (vgl. Schröder 1992, S. 36). Pädagogik formt nicht mehr, sondern befördert die Selbstbildung in Auseinandersetzung mit der Kultur (vgl. Koch 1999, S. 78). Jedoch fassen etwa Keller & Novak Bildung auch als individuelle Entfaltung in Folge intentionalen Unterrichts (vgl. Keller & Novak 1993, S. 63), was dann analog zum Erziehungsbegriff eine Differenzierung in intentionale und funktionale Bildungsprozesse zulässt. Nach Köck (1997, S. 99) wird die Fremdbestimmung im Bildungsvorgang zugunsten der Selbstbestimmung nach und nach abgebaut. In der Konzentration auf den Selbstbildungsgedanken kann Bildungsarbeit hingegen niemals als gestaltende Einflussnahme im Sinne der „Formung nach einem vorgegebenen Bilde gedacht sein, sondern allein als Angebot von außen und Bereitschaft von innen zum Dialog über das Sein des Einzelnen wie der Menschheit“ (Schaub & Zenke 2000, S. 95).

An dieser Stelle soll jedoch keine systematische Darstellung des Bildungsbegriffes in seiner historischen Genese geleistet werden. Stattdessen wird er im Weiteren dem Kompetenzbegriff gegenübergestellt, womit die These verbunden ist, dass der Kompetenzbegriff den Bildungsbegriff in vielerlei Hinsicht abgelöst, nicht jedoch verdrängt hat, wobei prinzipiell auch eine Umkehrung dieses Entwicklungsprozesses möglich erscheint (Faulstich 2002).
In der vom Bundesministerium für Bildung und Forschung vorgelegten Expertise „Zur Entwicklung nationaler Bildungsstandards“ (Klieme et al. 2007) wird deutlich, dass man auch hier am Bildungsbegriff festhalten möchte, allerdings in der Koppelung mit dem Begriff „Standard“. Die Expertise beschreibt wiederum Kompetenzen und Kompetenzmodelle als Grundlagen für die Entwicklung von „Bildungsstandards“ und verweist auf damit zusammenhängende Konstruktions- und Legitimationsprobleme (vgl. Klieme et al. 2007, S. 55 ff.). In der Expertise werden bestehende Vorwürfe gegen die Implementation nationaler Bildungsstandards antizipiert und mit den Begrifflichkeiten „Reduktionismusvorwurf, Vorwurf der Vereinheitlichung und Nivellierung“ sowie „Legitimationsdefizit“ diskutiert. Allerdings sehen die Autoren der Expertise hierin zwar ernstzunehmende Bedenken, nicht jedoch einen systematischen Einwand, der gegen die Einführung von nationalen Bildungsstandards sprechen könnte. Vielmehr insistieren sie auf der Notwendigkeit einer pragmatischen Annäherung und erkennen in der Formulierung von Bildungsstandards und Kompetenzmodellen eine Antwort auf das Problem von Bildungszielen.
Kritischer betrachtet etwa Regenbrecht (2005) den Begriff der Bildungsstandards und plädiert dafür, besser von „Leistungsstandards“ zu sprechen, weil die Bildungsstandards nicht Bildung, sondern schulische Leistungen thematisieren, wobei diese Leistungen eine notwendige, aber eben nicht hinreichende Bedingung für Bildung darstellen würden (Regenbrecht 2005, S. 17):

„Bildung setzt sich nicht aus einer Vielzahl von sachbezogenen Kompetenzen zusammen, sondern bedeutet Identität, Authentizität und Autonomie der Person. Bildung ist immer individuell. Es gibt sie nur in der Einzahl und daher lässt sie sich grundsätzlich nicht an objektivierten Standards vergleichend messen.“

Benner (2002) versucht im Unterschied dazu, eine Anschlussfähigkeit von Bildungsbegriff und Kompetenzdiskussion zu erzielen. In diesem Zusammenhang eruiert er die Struktur der Allgemeinbildung im Kerncurriculum moderner Bildungssysteme. Mit der Formulierung einer „dreifach horizontalen Grundstruktur moderner Bildung und ihrer Inhalte und Fragestellungen“ möchte er letztlich bestimmte Reflexionsstrukturen aufzeigen. Auf der Ebene der Schulstufe unterscheidet er:

- Elementarstufe des allgemeinbildenden Schulunterrichts,
- Schulstufe der über Wissenschaft vermittelnden Weltkunden,
- Schulstufe wissenschaftspropädeutischer Wissens- und Reflexionsformen.

Diesen Schulstufen ordnet Benner folgende Aufgaben zu:

- Lernen, mit Weltinhalten und Mitmenschen im Medium der Schriftsprache umzugehen,
- Aneignung elementarer, an die Beherrschung der Schriftsprache gebundener und ohne Schulbildung nicht tradierbarer Kulturbereiche,
- Ausdifferenzierung szientifischer historischer und praktischer Wissensformen und Einübung in Unterscheidungen einer nicht-fundamentalistischen Kritik.

Darüber hinaus ordnet Benner den Schulstufen fachliche Inhalte und sog. „Blickwechsel“ zu, die jeweils geleistet werden sollen. Man kann unschwer erkennen, dass es sich bei dieser Kernstruktur moderner Bildung um eine sehr ambitionierte Herangehensweise handelt, die Reflexion, kulturelle Werte und Wissenschaftlichkeit in den Mittelpunkt stellt, dabei aber eher von einem konservativ-idealistischen Bildungsbegriff ausgeht.

Auch disziplinübergreifend wird das Verhältnis von Bildung und Kompetenz kontrovers diskutiert. Für den Bereich der Erwachsenenpädagogik stellt etwa Faulstich (2002, S. 15) einen ständigen Wechsel der Leitbegriffe fest, die sich von „Bildung zu Qualifikation, über Kompetenz zum Lernen“ entwickelt haben. Er betrachtet diese Entwicklung als eine Suchbewegung, die ein empirisch brauchbares und praktisch verwertbares Instrumentarium finden und sich von „ideologiegeladenen Kategorien wie Bildung“ absetzen will. Faulstich plädiert aber letztendlich doch dafür den Bildungsbegriff wieder als zentrale Kategorie zu betrachten, weil er in ihm die beste Orientierungsfunktion für die Bewältigung zukünftiger Aufgaben sieht.
Auch Brödel (2002, S. 39) schreibt dem Kompetenzbegriff zum einen positive Potentiale zu, weil der Begriff den professionellen Blick auf Lernprozesse in Praxissituationen des gesellschaftlichen Lebens und Arbeitens lenken und zudem individuelle Problemlösepotentiale mitberücksichtigen würde. Zum anderen betont der Autor in ideologiekritischer Perspektive die inflationäre Verwendung des Kompetenzbegriffes, der insofern häufig nicht zweckfrei eingesetzt, sondern spezifischen Interessen (z.B. finanzpolitischer Art) dienen würde. Im Vergleich des Bildungs- mit dem Kompetenzbegriff macht Brödel darauf aufmerksam, dass sich die Begriffe aus unterschiedlichen Referenzsystemen speisen. Während der

Bildungsbegriff den kollektiven gesellschaftlichen Zusammenhang reflektiert, verweisen Kompetenzen auf eine individuelle Dimension.

Die eingangs gestellte Frage, inwieweit Prozesse der Kompetenzentwicklung die Zieldimensionen von Bildung abgelöst haben, lässt sich unter Hinzuziehung von unterschiedlichen und differenten Lagern zuzuordnenden Autorenreflexionen nicht abschließend beantworten. Bisher überwiegen Abwägungen, Annäherungen, aber auch polare Differenzierungen in der Beschäftigung mit einem zu charakterisierenden Verhältnis der Leitbegriffe. Der Bildungsbegriff weist auf Grund seiner historischen Tiefendimensionen, die sich unter Verzicht auf einen mythisch anmutenden Überschuss allerdings kaum noch ausloten lassen, einerseits eine Reihe von Unschärfen auf, die für zu beschreibende Zielsetzungen komplexer Lehr-Lernprozesse eine Umorientierung in der Begriffsverwendung und zumindest dem Begriffsverständnis erforderlich scheinen lässt. Skeptisch beschrieben erscheint Bildung in der modernen bzw. postmodernen Gesellschaft als ein rückwärtsgewandtes Konstrukt mit elitären Konnotationen. Andererseits sorgen genau diese historischen Tiefendimensionen, festzumachen an vielfältigen und multidisziplinären Diskursen, dafür, dass der Bildungsbegriff nach wie vor für einen der meistgebrauchten Begriffe in der Pädagogik steht und als pädagogische Kategorie nicht austauschbar ist. Davon abgegrenzt wird der Kompetenzbegriff teilweise als Innovationsbegriff verwendet, weil er noch relativ unbelastet auf verschiedene Handlungsfelder angewendet werden kann und sich besonders für die Beschreibung beruflicher Anforderungen bewährt hat. Für die eigene Studie nimmt er deshalb eine grundlegende Stellung ein.

2.9 Fazit

Der noch vergleichsweise junge Begriff „Kompetenz“ wurde in seinem historischen Entwicklungsverlauf skizziert, um seine theoretischen Voraussetzungen und Implementationen zu klären. Den Ansätzen von Chomsky und Habermas lag ein transformationstheoretisches Verständnis von Kompetenz zugrunde. Danach entsteht subjektive Handlungsfähigkeit aus einem dialektischen Wechselspiel von Kompetenz und Performanz, von Tiefen- und Oberflächenstruktur, wobei sich das Subjekt generativ im eigenen Handeln erzeugt (Brödel 2002). Die Fortführungen von Geißler und Baacke lehnen sich noch stark an ein solches grundlagentheoretisches Verständnis an, schließen den Begriff jedoch spezifisch für die pädagogische Diskussion auf. Die Thematisierung pädagogischer Handlungskompetenzen verlässt den grundlagentheoretischen Rahmen schließlich vollends und fokussiert auf Fragen der Ausbildung und der Anforderungen, die sich aus der Beziehung zwischen dem kompetent handelnden Pädagogen und dem von ihm vorgefundenen beruflichen Handlungsfeld bzw. darin sich vollziehenden Interaktionen ergeben.
Grundlegend zur theoretischen Fundierung der eigenen Fragestellung ist zudem die Diskussion um Schlüsselqualifikationen einzubeziehen, weil sie Parallelen zur aktuellen Kompetenzdiskussion aufweist und zu dieser hinführt. Beide Diskurse zeugen von einem Überschusspotential in den Begriffsverständnissen, welches nur ansatzweise begründungs- und handlungslogisch einholbar erscheint. Zudem hat der

Kompetenzbegriff die Diskussion um Schlüsselqualifikationen abgelöst und wird z.T synonym verwendet, z.B. als „Schlüsselkompetenz“. Die Fragestellung, inwieweit „Erschließungskompetenzen“ herausgebildet werden können, die einen Transfer zwischen verschiedenen Problemstellungen erlauben, wurde im Rückgriff auf Theorien, Modelle und Ergebnisse der Kognitionspsychologie diskutiert. Transferpotentiale können dort ausgemacht werden, wo die Entwicklung abstrahierender Schemata gelingt.
Die Betrachtungen mündeten vor dem Hintergrund einer intendierten terminologischen und kategorialen Grundlegung der eigenen Studie in eine vergleichende Gegenüberstellung von Kompetenz- und Bildungsbegriff. Obwohl beide Begriffe völlig unterschiedliche Referenzsysteme aufweisen und einen jeweils spezifischen Geneseprozess durchlaufen haben, lassen sich doch auch Konvergenzen aufzeigen. Man kann vermuten, dass der Kompetenzbegriff angetreten ist, um den Bildungsbegriff abzulösen (bzw. dass bestimmte gesellschaftliche Gruppen und Institutionen den Kompetenzbegriff mit dieser Absicht lanciert haben). Teilweise ist das, mit sowohl klärenden als aber auch instrumentalisierenden Auswirkungen, gelungen.
Betrachtet man die aktuelle vom Bundesministerium für Bildung und Forschung in Auftrag gegebene Expertise „Zur Entwicklung nationaler Bildungsstandards“, stehen die Begriffe „Bildung“, „Standards“ und „Kompetenzen“ seltsam versöhnlich nebeneinander. Was angesichts der lange ausgefochtenen Zuschreibungs- und Positionskämpfe einerseits paradox anmutet, bietet andererseits die Möglichkeit, die in den Begriffen aufscheinenden Potentiale miteinander zu verzahnen. Es gibt mittlerweile einen – wiederum der Tradition folgenden – Konsens in weiten Teilen der Erziehungswissenschaft und auch unter Vertretern der empirischen Bildungsforschung, dass schulische Bildung nicht allein zweckfreie Dimensionen ansprechen kann und damit zielgerichtete Bildung eben immer noch „Bildung“ bleibt. Zur Beschreibung von Bildungszielen hat sich der Kompetenzbegriff auch im schulischen Feld durchgesetzt, weil er vielfältige Anschlussmöglichkeiten liefert und empirisch modellierbar erscheint (Benner 2007).
Spätestens seitdem sich Deutschland an den internationalen Schulleistungsvergleichen beteiligt, wurde auch die Bedeutung von standardisierten Testverfahren stärker herausgestellt. Die in der angeführten bildungspolitischen Expertise verfolgte Linie einer Orientierung an Kompetenzmodellierungen, die eine standardisierte Überprüfung ermöglichen soll und sich trotzdem nicht vom Bildungsbegriff (und seinen Implikationen) verabschiedet, werte ich als eine sinnvolle und weiterführende Entwicklung für die deutsche Erziehungswissenschaft in ihrem Bezug zu pädagogischen Handlungsherausforderungen und ihrer internationalen Anschlussfähigkeit.
Im Folgenden sollen aktuelle Kompetenzdefinitionen, -konstrukte und -modelle dargestellt werden, die Gegenstand erziehungswissenschaftlicher und psychologischer empirischer Forschung sind.

3 Kompetenzen als latente Konstrukte erziehungswissenschaftlicher und psychologischer Forschung

Der Kompetenzbegriff hat in der erziehungswissenschaftlichen bzw. psychologischen Forschung eine erstaunliche Karriere absolviert und steht im Zentrum aktueller Diskurse, die nicht zuletzt eine Umsteuerung des Bildungssystems anstreben (Klieme 2003; KMK 2004). Die oben skizzierten historischen Entwicklungslinien des Kompetenzbegriffes als strukturtheoretisch-anthropologische Markierungspunkte werden in der empirischen Forschung nur rudimentär diskutiert. Gleichwohl liefern sie Ansatzpunkte für die Entwicklung von Untersuchungsbereichen. Im Folgenden soll es um Formen der Modellierung von Kompetenzen gehen, die operationale Definitionen vornehmen, damit empirisch anschlussfähig sind und insofern der Forschung zugänglich gemacht werden. Diesbezüglich gibt es unterschiedliche Herangehensweisen und Modelle, die sich nicht unbedingt gegenseitig ausschließen, allerdings jeweils unterschiedliche Schwerpunktsetzungen vornehmen und damit für mögliche Perspektivwechsel stehen. Die grundsätzliche Herausforderung bei der Formulierung von Kompetenzkonstrukten, das auf Chomsky zurückgehende Verhältnis von Kompetenz und Performanz, bleibt in nachfolgenden Ausführungen weiterhin berücksichtigt, indem der Kompetenzbegriff als Dispositionsbegriff über latente Variablen aufgeschlossen werden soll.

3.1 Kompetenzmodellierungen in Abhängigkeit des Abstraktionsniveaus

In der empirischen Forschung hat es verschiedene Vorschläge darüber gegeben, auf welchem Abstraktionsniveau Kompetenzen operational definiert werden können. Die Wahl des jeweiligen Niveaus hat dabei großen Einfluss auf die empirische Anschlussfähigkeit. Wählt man ein höheres Abstraktionsniveau, so gewinnt die Definition an theoretischer Plausibilität und erscheint anschlussfähiger an historische Diskurse. Die Möglichkeiten einer empirischen Untersuchung werden dabei jedoch eingeschränkt. Umgekehrt verlieren Kompetenzmodellierungen, die auf einem zu geringen Abstraktionsniveau vorgenommen werden, ihren Begründungszusammenhang hinsichtlich des Kompetenzbegriffes, wenn sie sich z.B. lediglich auf die Überprüfung deklarativer Wissensbestände beschränken. Aus psychologischer Warte diskutiert Wollersheim (1993, S. 92; zit. n. Vonken 2005, S. 30) die Frage des Abstraktionsniveaus im Bezug auf ausgewählte Kompetenzkonzepte unter der Perspektive von Bewältigungsprozessen:

„Die erste Gruppe von Autoren definiert Kompetenz auf höchster Abstraktionsebene: als Fähigkeit zur Kontrolle, als Fähigkeit zur Akkommodation und Assimilation, als Fähigkeit zur Auseinandersetzung mit oder Bewältigung von (neuen) Situationen, als erwiesene Fähigkeit zum Handeln, als Fähigkeit zur selbständigen und eigenverantwortlichen Lebensführung. Die zweite Gruppe thematisiert demgegenüber inhaltliche Komponenten, die einer oder mehreren der formalisierten Definitionen der ersten Gruppe zuzuordnen sind, und zwar in der Regel in der Weise, dass hier Aspekte explizit betont werden, die bei der üblichen und nahe liegenden Explikation von Kompetenz durch epistemische Kategorien unbeachtet bleiben: Zur Kompetenz im Sinne einer umfassenden Fähigkeit zur

Auseinandersetzung mit oder Bewältigung von Lebenssituationen gehören neben Wissen eben auch die Fähigkeit zur Pflege sozialer Kontakte, die Fähigkeit zur Antizipation und Verarbeitung von Belastungen, die Erfahrung der Transformation eigener Bewältigungsmöglichkeiten und -techniken im Laufe der Biographie."

Das gewählte Abstraktionsniveau kann somit nicht nur als Herausforderung der empirischen Forschung betrachtet werden, sondern damit verbunden sind auch grundsätzliche mit dem Kompetenzbegriff assoziierte Ausrichtungen. Abstraktere Definitionen verweisen eher auf Fähigkeitsdimensionen, die dem Individuum in vielfältigen lebensweltlichen Bereichen zuzuschreiben sind. Der Kompetenzbegriff rückt dort in die Nähe zum Bildungsbegriff und dem Konzept der Schlüsselqualifikationen, womit sich nach Faulstich alle „Messbarkeitsillusionen" eben als Illusionen erweisen würden (Faulstich 2002). Dem Autor ist insofern zuzustimmen, als der Kompetenzbegriff auch in Konnotationen und Verwendungsweisen Einsatz findet, die sehr unscharf und empirisch nicht unbedingt überprüfbar sind. Einer generellen Messbarkeitsillusion soll hier aber widersprochen werden. Aufgabe von Erziehungswissenschaft und Psychologie ist die operationale Definition des Kompetenzbegriffes, der dann unter Einhaltung von Gütekriterien methodisch zugänglich ist.

3.2 Dimensionen des Kompetenzbegriffes in der Erziehungswissenschaft und der Psychologie

Die Diskussion um verschiedene Abstraktionsniveaus zur Modellierung von Kompetenzen führt zu der Überzeugung, dass sie sich vermutlich am besten auf einem mittleren Abstraktionsniveau beschreiben und erfassen lassen. Sie behalten ihren Potentialcharakter, indem sie als Dispositionsbegriff entworfen und über latente Variablen erschlossen werden. Kompetenzen als psychische Konstrukte lassen sich dabei jeweils nur über bereichsspezifische äußere Tätigkeiten als latente Verweise bzw. Performanzen untersuchen. Hascher & Thonhauser (2004) machen zudem darauf aufmerksam, dass Kompetenzen nicht als statische Phänomene zu betrachten sind, hingegen sollte man den dynamischen Entwicklungscharakter berücksichtigen, womit auch psychische Verlaufsqualitäten eingeschlossen sind.
Der spezifische Mehrwert des Kompetenzkonzeptes erschließt sich nur, wenn verschiedene Dimensionen gebündelt werden, um dem Konzept gerecht zu werden. Ein solches wissenschaftliches Bedeutungsverständnis steht im Gegensatz zur allgemeinsprachlichen Verwendung des Begriffes und erhält dann erst seine Berechtigung. Kompetenzen können somit nicht als verallgemeinerte, kontextfreie Fähigkeitsdimensionen gedacht werden, sondern wesentlich von Anforderungen der Lebenswelt und des Arbeitsfeldes her (Klieme 2004). Jedoch muss eine derartige in bestimmter Weise funktional gedachte Zielperspektive zwischen der Entwicklung von bestimmten domänenspezifischen Kompetenzen und der summativen Messung unterscheiden. Dem Kompetenzbegriff ist so die Kombination von „inhaltsübergreifender" und zugleich „anforderungs- und situationsbezogener" Charakteristik eigen (ebd.) – er erhält eine theoretische Schärfe, die ihn vom Bildungsbegriff und dem Konzept der Schlüsselqualifikationen abgrenzt. Vor dem

Hintergrund einer empirischen Messbarkeit erscheinen damit verbundene Einschränkungen bezüglich einer Anschlussfähigkeit an strukturtheoretische Diskurse nur im geringen Maße ausschlaggebend, wenn auch nicht völlig problemlos. Ein bestehender Vorwurf des Pragmatismus lässt sich nicht völlig verwerfen.
Dargelegte strukturelle Beschreibungen umreißen einen modellierten psychologischen Kompetenzbegriff, der grundsätzlich auch empirisch untersucht werden kann (verbunden allerdings mit verschiedenen Schwierigkeiten und Anforderungen); in diesem Zusammenhang lässt er sich folgendermaßen auffächern. Kompetenzen sind:

- dem Individuum zugehörig und enthalten also biographische Anteile,
- in ihrer Entwicklung dynamisch und nicht linear,
- psychische Konstrukte, die nur über verschiedene Handlungsformen (Fertigkeiten) als Performanzen erschlossen werden können,
- inhalts- und domänenspezifisch, zugleich aber auch inhaltsübergreifend,
- wissensbasierte Dispositionen in Form von Befähigungen,
- Konstrukte, die psychische Verlaufsqualitäten enthalten (z.B. Motivation und Volation),
- erfahrungsbasiert und zeigen sich in Handlungsformen, auch sozialer Art.

Die Auflistung unterstreicht die Breite vorgenommener Dimensionalisierungen des Begriffes und erhebt keinen Anspruch auf Vollständigkeit. Aufgeführte Dimensionen erscheinen zwar vielperspektivisch gefasst, dennoch sind sie prinzipiell empirisch zugänglich. Die Problematik der Multidimensionalität liegt eher in Fragen, inwieweit alle Dimensionen realistischer Weise gleichzeitig berücksichtigt werden können bzw. wie groß ein diesbezüglicher Aufwand auszusehen hätte.

3.3 Definitionen des Kompetenzbegriffes in der Erziehungswissenschaft und der Psychologie

Unter der Fragestellung „Was ist Kompetenz? – ein theoretischer Rahmen mit Blick auf die beruflichen Fähigkeiten von Lehrerinnen und Lehrer“ formuliert Plöger (2006) zehn Thesen, die eine Annäherung an den Kompetenzbegriff aus pädagogischer Sicht erlauben.

1. These: Im Begriff der Kompetenz werden Wissen und Handeln adäquat aufeinander bezogen.
2. These: Kompetentes Wissen und Handeln bedürfen der reflexiven Legitimation.
3. These: Kompetentes Handeln und Denken zeichnet sich durch Sinnkohärenz aus.
4. These: Kompetentes Denken und Handeln stellen sich der Öffentlichkeit.
5. These: Kompetenz basiert auf einem differenzierten und strukturierten Wissensbestand, der sich auf die konstitutiven Aspekte des jeweiligen Handlungsfeldes bezieht.

6. These: Kompetenz ist Ausdruck „erarbeiteter Identität“.
7. These: Kompetenzentwicklung ist nur in geringem Maße durch Systemsteuerung beeinflussbar.
8. These: Kompetenz ist vorwiegend Ausdruck individueller Fähigkeiten (die sich allerdings durch gemeinsame Prozesse der Qualitätsentwicklung von Schulen erheblich steigern können).
9. These: Kompetenzen entwickeln sich im Rahmen von subjektiven Theorien.
10. These: Eine inhaltliche Bestimmung von Kompetenz ergibt sich aus einem historisch-systematischen Verständnis von Schulpädagogik und den Anforderungen des pädagogischen Alltags.

Die von Plöger vorgestellten Thesen sprechen für sich und sollen hier nur eine kurze Einordnung erfahren. Zum einen kann man eine Orientierung an bildungstheoretischen Positionen in der Betonung von reflexiver Legitimation, erarbeiteter Identität, Sinnkohärenz, fehlenden Möglichkeiten der Systemsteuerung und dem historisch-systematischen Verständnis von Schulpädagogik erkennen. Zum anderen fallen in stärker anwendungsbezogener Ausrichtung die Integration von Wissen und Handeln, die Berücksichtigung sozialer Aspekte und kognitiver Wissensbestände sowie eine pragmatische Orientierung an Berufsanforderungen auf. Dabei ist keine Gewichtung in dem polaren Verhältnis erkennbar.
Insgesamt verbleiben Plögers Aussagen auf der noch relativ unscharfen Ebene eines Thesenpapieres und entsprechen keiner eigentlichen für den Untersuchungszusammenhang benötigten Definition.

In der psychologischen Forschung wurden verschiedene Versuche unternommen, Kompetenzen zu definieren. Sie unterscheiden sich im Grad der Umfassendheit und der Bereichsspezifität. Erpenbeck & Rosenstiel (2003, S. IX) halten fest:

„Davon ausgehend muss verwundern, wie wenig klar ‚Kompetenz’ gegenwärtig begrifflich gefasst und messend zugänglich gemacht werden kann. Aber nur was vergleichend beschrieben, qualitativ charakterisiert und womöglich quantitativ verglichen werden kann, wird wirklich begriffen, bleibt nicht bloß Begriff. Allerdings wäre es verfehlt, auf ein einheitliches Verständnis zu hoffen, auf allgemein verbindliche Charakterisierungen und Messverfahren zu warten. Dazu sind die Phänomene, auf die der Begriff verweist, zu komplex und die Gebiete, in denen er Bedeutung gewinnt, zu vielfältig. Doch zeichnen sich in den letzten Jahren zunehmend verbindende Überlegungen und gemeinsame Vorgehensweisen ab.“

Die genannten Schwierigkeiten einer umfassenden Charakterisierung und Definition haben bis heute nicht an Gewichtigkeit verloren. Sie sind im Kompetenzbegriff und in dahinter liegenden Konzepten implizit angelegt. Der Prozess einer Verständigung und Einigung ist zumindest in der psychologischen Forschung, und auch der erziehungswissenschaftlichen, so sie sich auf diese Konzepte bezieht, jedoch weiter vorangeschritten. Frey (2006, S. 126) erarbeitet eine Synopse aus verschiedenen Kompetenzdefinitionen und kommt zu folgendem Ergebnis:

„Besitzt eine Person Kompetenz, so kann sie etwas, ist handlungsfähig und übernimmt für sich und andere Verantwortung. Sie besitzt die Kompetenz, so tätig zu werden, dass sie eine Absicht, ein Ziel

oder einen Zweck unter Beachtung von Handlungsprinzipien, Werten, Normen und Regeln, mit Bezug auf konkrete, die jeweilige Handlungssituation bestimmende Bedingungen zu erreichen vermag. Wer Kompetenz besitzt, ist erfolgreich und vernünftig tätig. Somit kann man Kompetenz als ein Bündel von körperlichen und geistigen Fähigkeiten bezeichnen, die jemand benötigt, um anstehende Aufgaben oder Probleme zielorientiert und verantwortungsvoll zu lösen, die Lösungen zu bewerten und das eigene Repertoire an Handlungsmustern weiterzuentwickeln."

Frey hebt hervor, dass in dieser umfassenden Definition verschiedene Kompetenzklassen aufscheinen, die nicht isoliert voneinander betrachtet werden können, sondern die eine vernetzte Struktur abgeben, d.h. sich gegenseitig positiv oder negativ beeinflussen und erst in der Gesamtschau die Handlungskompetenz eines Individuums ausmachen.

Maag Merki & Grob (2003, S. 125) weisen ebenfalls auf eine nicht einheitliche Verwendung des Kompetenzbegriffes hin und betrachten den diesbezüglichen Diskurs gleichwohl als anschlussfähig an langjährige Forschungstraditionen z.B. in der Intelligenzforschung oder der Persönlichkeitsforschung. Im Versuch einer begrifflichen Klärung und in der Bezugnahme auf Rychen & Salganik (2000) und Weinert (2001) kommen sie zu folgender Definition:

„Unter Kompetenzen sollen Eigenschaften verstanden werden, welche ihre Träger/innen in die Lage versetzen, bestimmte Aufgaben erfolgreich zu verwenden. Kompetenzen entsprechen dem Potential, in konkreten Situationen erfolgreich handeln zu können. Weil konkretes Handeln (Performanz) immer auch durch Merkmale der konkreten Situation beeinflusst wird, besteht zwischen Kompetenz und Performanz nicht ein deterministischer, sondern ein vermittelter, stochastischer Zusammenhang. Kompetenzen weisen den Charakter von relativ stabilen Personmerkmalen auf. Sie entsprechen unterschiedlich komplexen kognitiv-motorischen Schemata, welche in emotionale und motivationale Regelkreise eingebunden sind und sukzessive angeeignet, aufgebaut bzw. adaptiv modifiziert werden" (Maag Merki & Grob 2003, S. 125).

Die Definitionen von Frey und Maag Merki & Grob weisen Überschneidungen auf, verdeutlichen aber auch, dass solch breiten Begriffsverständnisse, obwohl theoretisch überzeugend, forschungspraktisch kaum in Form valider Konstrukte operational umgesetzt werden können.

Dieser Schwierigkeit Rechnung tragend hat Weinert (1999) im Auftrag der OECD ein einflussreiches Gutachten erstellt, in dem er eine Auswahl und Definition von Kompetenzen als Grundlage für internationale Schulleistungsstudien entwickelt. Dort nimmt er eine Aufteilung in zwei Begriffsbestimmungen vor (zit. n. Klieme 2004, S. 10 f.).

1. Kompetenzen in einem allgemeinen Sinn verstehen sich als intellektuelle Fähigkeiten und Dispositionen, die eine Person in die Lage versetzen, unterschiedliche, im Schwierigkeitsgrad differierende Situationen erfolgreich zu bewältigen.

2. Kompetenzen in einem funktionalen Sinn beziehen sich auf bestimmte Kompetenzklassen:

- Auf Situationen und Anforderungen bezogene kognitive Leistungsdispositionen, die sich psychologisch als Kenntnisse, Fertigkeiten, Strategien, Routinen oder auch bereichsspezifische Fähigkeiten beschreiben lassen,
- Kompetenz im Sinne motivationaler Orientierungen, die Voraussetzungen sind für die Bewältigung anspruchsvoller Aufgaben,
- Handlungskompetenz als Begriff, der die genannten Kompetenzkonzepte umschließt und sich jeweils auf die Anforderungen und Aufgaben eines bestimmten Handlungsfeldes, z.B. eines Berufes, bezieht,
- Metakompetenzen als Wissen, Strategien oder Motivationen, die Erwerb und Anwendung von Kompetenzen in verschiedenen Inhaltsbereichen erleichtern,
- Schlüsselqualifikationen, die über eine vergleichsweise breite Spanne von Situationen und Aufgabenstellungen hinweg einsetzbar sind; Weinert zählt hierzu unter anderem muttersprachliche und mathematische Kenntnisse und Fertigkeiten sowie die Inhalte einer basalen Allgemeinbildung.

Die von Weinert unterschiedenen Kompetenzbestimmungen lassen es zu, unter inhaltlicher Begründung eine Auswahl vorzunehmen. Für die internationalen Schulleistungsstudien hat er nach gründlichen Abwägungsprozessen in seinem Gutachten die zweite genannte Begriffsbestimmung vorgeschlagen. Klieme (2004, S. 13) folgert daraus:

„Kompetenzen sind funktional bestimmt, aber rein auf das Kognitive beschränkt. Sie sind bereichsspezifisch, d.h. auf einen begrenzten Sektor von Kontexten und Situationen bezogen, aber doch als – begrenzt – verallgemeinerbare gedacht, d.h. als Dispositionen."

Weinerts vorgeschlagene Dimensionalität des Kompetenzbegriffes in einer Fokussierung auf funktional-kognitive Bereiche mag auf den ersten Blick als Verkürzung anmuten. Für den Bereich schulischer Leistungsstudien erscheint eine Konzentration auf kognitive Leistungsdimensionen aber nachvollziehbar. Dennoch hat Weinert zu einem späteren Zeitpunkt seine ursprünglich vorgenommene Eingrenzung wieder aufgehoben. Kompetenzen sind demnach (Weinert 2001, S. 27f.)

„die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können."

Diese im Vergleich zum früheren Entwurf breitere Definition hat sich in der Erziehungswissenschaft durchgesetzt und kann als Referenzdefinition bezeichnet werden. Sie überzeugt im Hinblick auf die umfassende Berücksichtigung, aber doch präzise Beschreibung intrapsychischer Verlaufsmuster sowie den

kompetenzbezogenen Potentialcharakter als Bedingung für die Bearbeitung von Problemstellungen.
Fähigkeiten und Fertigkeiten sind gemäß Weinerts Verständnis dem Kompetenzbegriff untergeordnet. Sie werden durch Lernprozesse erzeugt und benötigen voraussetzend Wissens- und Könnensbasen. Kompetenzen werden im Problemhandeln generiert und modelliert. In ihrer Genese verlangen sie aber auch Entscheidungen durch das Individuum im sozialen Prozess. Frey (2006) hebt denn auch die Bedeutung von bewusst gesteuerten Tätigkeiten als Problemhandeln heraus, die Fertigkeiten, Fähigkeiten und schließlich die Entwicklung von Kompetenzen ermöglichen.

Mit seiner Definition hat Weinert eine Präzisierung geliefert, die prinzipiell die Möglichkeit der empirischen Bearbeitung eröffnet. Neben der genannten Referenzdefinition hat er generell wichtige Klärungen zum Kompetenzbegriff beigetragen, die besonders für die Anwendung in didaktischen Handlungsfeldern hervorzuheben sind (aber auch auf andere Berufsfelder angewandt werden können). So spricht er sich dafür aus, nur dann den Kompetenzbegriff zu verwenden, wenn damit auf die Bewältigung von komplexen Anforderungen verwiesen wird und so der konstitutive Zusammenhang zwischen psychologischer Anforderungsstruktur und Kompetenzstruktur herausgestellt wird. In diesem Kompetenzverständnis – und hier weist Weinert durchaus Verbindungen zu Plöger (2006) auf – beinhaltet die Bewältigung komplexer Anforderungen sowohl kognitive als auch motivationale, ethische, volitionale und soziale Komponenten. Kompetenzen sind dann nicht als „skills“, also als automatisierte Fertigkeiten zu verstehen, sondern als Ergebnis komplexer Lernprozesse.
Darüber hinaus verweist Weinert auf die Notwendigkeit einer konzeptuellen Trennung von Schlüsselkompetenzen und Metakompetenzen, die im wissenschaftlichen Sprachgebrauch häufig synonym verwendet werden. Er versteht Schlüsselkompetenzen als Kompetenzfeld, das sich auf verschiedene, aber gleichrangig gelagerte Dimensionen von Anforderungsbereichen bezieht. Dagegen zielt der Begriff der Metakompetenzen auf deklaratives bzw. prozedurales Wissen über die eigenen Kompetenzen (Erpenbeck & Rosenstiel 2003, S. XXXI).
Nach Rupp, Leucht & Hartung (2006) bleibt der Kompetenzbegriff von Weinert sehr weit gefasst, indem er nicht nur kognitionspsychologische (Fähigkeiten und Fertigkeiten), sondern auch lernpsychologische Elemente (motivationale, volitionale und soziale Bereitschaften) beinhaltet. Zwar ist die Möglichkeit geschaffen entsprechende Testbatterien und Fragebögen zur Leistungs-, Persönlichkeits- und Verhaltensdiagnostik auf den Begriff zu beziehen und die entsprechenden Konstrukte in Beziehung zueinander zu setzen. Die Anforderungen an die Forschergruppe im Hinblick auf den zu betreibenden Aufwand sind hier aber sehr hoch gesteckt.

3.4 Die Erfassung von Kompetenzen

Erpenbeck & Rosenstiel (2003, S. X) werfen die Frage auf, inwieweit Kompetenzen überhaupt beobachtbar sind, wenn sie doch innere Voraussetzungen, Dispositionen

des selbstorganisierten Handelns einer Person implizieren. Sie schlussfolgern, dass sie nur anhand der tatsächlichen Performanz, der Anwendung und des Gebrauches von Kompetenzen, aufzuklären sind. Das Verhältnis zwischen Kompetenz und Performanz stellt sich als Teil-Ganzes-Beziehung dar (Huber 2001). Kompetenzen als psychische Konstrukte können nur im Rahmen von theoriegeleiteten Verfahren beobachtet werden. Sie sind damit an Attribuierungen geknüpft, die sich aus dem Theoriekonzept und dem Urteil von Beobachtern ergeben. In Abhängigkeit der Fragestellung kann die Beobachtung auch durch Selbsteinschätzungen erfolgen. Kompetenzen in einem darüber hinaus gehenden Sinne entziehen sich der Beobachtung. Es ergibt sich damit die Notwendigkeit einer Differenzierung von Kompetenzklassen, die unter Verwendung spezifischer Verfahren einer Beobachtung zugänglich gemacht werden können.

3.5 Kompetenzklassen

Erpenbeck & Rosenstiel (2003) machen einen Vorschlag, wie Kompetenzen in Form von Klassen differenziert werden können und knüpfen damit an die von Roth (1971) und Frey (2002) bzw. Frey & Balzer (2003) entwickelten Kompetenzklassen an. Sie unterscheiden:

- Personale Kompetenzen
- Aktivitäts- und umsetzungsorientierte Kompetenzen
- Fachlich-methodische Kompetenzen
- Sozial-kommunikative Kompetenzen

Kompetenzen verstanden als Dispositionen selbstorganisierten Handelns verweisen auf selbstreflexive Handlungen der Person, auf Aktivitäten und Entscheidungen, die inhalts- und könnensbasiert im Interaktionsgeschehen einer gegenstands- und personenbezogenen Umwelt stattfinden und sich im sozialen Prozess konstituieren.

Erpenbeck & Rosenstiel (2003, S. XVI) definieren ihre formulierten Kompetenzklassen genauer.

„Personale Kompetenzen: Als die Dispositionen einer Person, reflexiv selbstorganisiert zu handeln, d.h. sich selbst einzuschätzen, produktive Einstellungen, Werthaltungen, Motive und Selbstbilder zu entwickeln, eigene Begabungen, Motivationen, Leistungsvorsätze zu entfalten und sich im Rahmen der Arbeit und außerhalb kreativ zu entwickeln und zu lernen.

Aktivitäts- und umsetzungsorientierte Kompetenzen: Als die Dispositionen einer Person, aktiv und gesamtheitlich selbstorganisiert zu handeln und dieses Handeln auf die Umsetzung von Absichten, Vorhaben und Plänen zu richten – entweder für sich selbst oder auch für andere und mit anderen, im Team, im Unternehmen, in der Organisation. Diese Dispositionen erfassen damit das Vermögen, die eigenen Emotionen, Motivationen, Fähigkeiten und Erfahrungen und alle anderen Kompetenzen – personale, fachlich-methodische und sozial-kommunikative – in die eigenen Willensantriebe zu integrieren und Handlungen erfolgreich zu realisieren.

Fachlich-methodische Kompetenzen: Als die Dispositionen einer Person, bei der Lösung von sachlich-gegenständlichen Problemen geistig und physisch selbstorganisiert zu handeln, d.h. mit fachlichen und instrumentellen Kenntnissen, Fertigkeiten und Fähigkeiten kreativ Probleme zu lösen, Wissen sinnorientiert einzuordnen und zu bewerten; das schließt Dispositionen ein, Tätigkeiten, Aufgaben und Lösungen methodisch selbstorganisiert zu gestalten, sowie Methoden selbst kreativ weiterzuentwickeln.

Sozial kommunikative Kompetenzen: Als die Dispositionen, kommunikativ und kooperativ selbstorganisiert zu handeln, d.h. sich mit anderen kreativ auseinander zu setzen, sich gruppen- und beziehungsorientiert zu verhalten und neue Pläne, Aufgaben und Ziele zu entwickeln."

Die hier vorgestellten Kompetenzklassen basieren auf einer allgemeinen theoretischen Vorstellung über menschliche Auseinandersetzung mit Handlungs- und Problemlösungsherausforderungen. Sie bilden die Prozesshaftigkeit geforderter Bewältigungen, mit einer Schwerpunktsetzung auf der Selbstorganisation, umfassend ab.

Freys (2002, 2006) Ansatz entspricht vorgestellter Kompetenzklassifikation prinzipiell, wobei der Autor von Fach-, Methoden-, Sozial- und Personalkompetenzklassen spricht. Speziell wendet er seine Klassifikation auf die Formulierung beruflicher Handlungskompetenz an. Fachkompetenzen werden als disziplin- bzw. fachorientierte Fähigkeitskonzepte ausgelegt, die der Spezialisierung einer Person dienen und die Grundlage für die Erfüllung einer beruflichen Tätigkeit bilden. Methodenkompetenzen befähigen eine Person, sich in einem genau bestimmbaren Sachbereich als denk- und handlungsfähig zu erweisen. Damit sind Fähigkeitskonzepte beschrieben, die genauer als Strukturierung, Analyse und Reflexion bezeichnet werden können. Fähigkeitskonzepte, die es einer Person erlauben in Kooperation mit anderen ein angestrebtes Ziel zu erreichen, werden wiederum Sozialkompetenzen genannt. Hierzu zählen soziale Verantwortung, Kritikfähigkeit, Führungsfähigkeit u.a. Unter Personalkompetenzen versteht Frey Fähigkeitskonzepte, die Einstellungen oder Eigenschaften beinhalten, aufgrund derer eine Person in die Lage versetzt wird, verantwortlich und motiviert zu handeln. Zur Spezifizierung greift Frey (2006) auf den Tugendbegriff zurück und spricht von gewonnenen Einsichten, die für das Individuum „lebensführend" geworden sind und damit als sittliche Grundhaltungen bezeichnet werden können.
Für die Formulierung von Konstrukten spezifischer beruflicher Handlungskompetenz müssen die vier Kompetenzklassen weiter ausdifferenziert werden. Hinsichtlich der Ausbildung von Erzieherinnen und Erziehern bzw. Studierenden des Lehrerberufes macht Frey (2002, 2004, 2006) Vorschläge, wie solche Kompetenzstrukturen aussehen und wie sie erhoben werden können (vgl. Kap. 4).

„An dieser Stelle kann zusammenfassend formuliert werden, dass Kompetenz und Fähigkeit zwei theoretische Gebilde sind, die nur über Fertigkeiten strukturiert und diagnostiziert werden können. Indikatoren für Kompetenz stellen somit Fertigkeiten dar" (Frey 2006, S. 130).

3.6 Generierung eines Kompetenzmodells

Die oben dargestellten Klassifikationen und Strukturierungen erlauben es Kompetenzmodelle zu entwickeln. Solche Kompetenzmodelle müssen die Möglichkeit bieten, eine Unterscheidung zwischen Kompetenzklassen zu treffen sowie Relationen zwischen den Klassen zu beschreiben. Breit rezipiert wurde der Ansatz von Roth (1966, 1971). Sein Entwurf einer „pädagogischen Anthropologie“ zielte auf Mündigkeit durch Bildung ab; auf intellektuelle Mündigkeit durch „Sachkompetenz“, soziale Mündigkeit durch „Sozialkompetenz“ und moralische Mündigkeit durch „Selbstkompetenz“. Roth formulierte modellhaft das Zusammenspiel der drei Kompetenzbereiche, das dem Individuum im Spannungsfeld von Selbst- und Weltbezug mündiges Handeln ermöglicht. Dieses Modell stellt die Grundlage obiger Klassifikationen dar und wurde von Frey durch die Kompetenzklasse „Methodenkompetenz“ ergänzt.
Genauer bezieht sich Frey (2006) auf das von Roth vorgelegte Modell und erweitert es zu einem hypothetischen Modell beruflicher Kompetenz (vgl. Abb. 3-1).

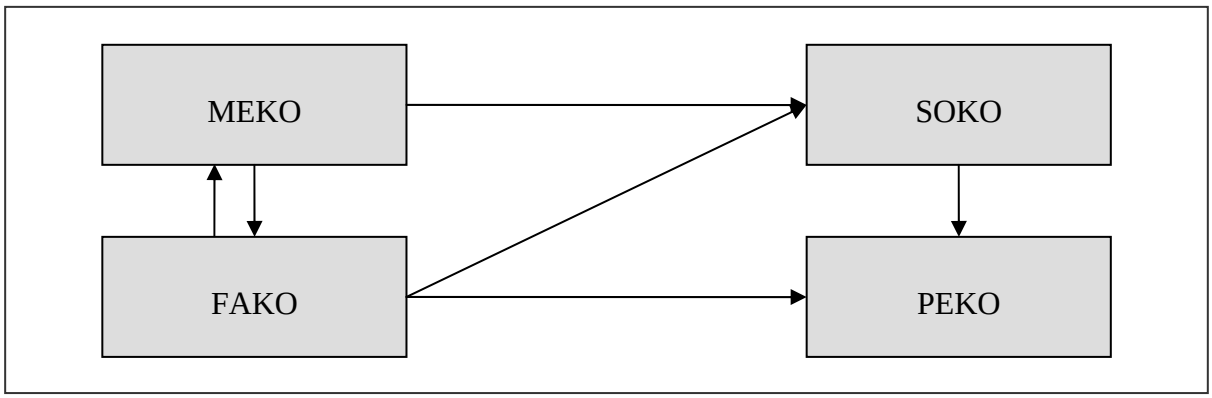

Abbildung 3-1: Hypothetisches Modell beruflicher Kompetenz (Frey 2006): MEKO: Methodenkompetenz, SOKO: Sozialkompetenz, FAKO: Fachkompetenz, PEKO: Personalkompetenz.

Es besteht aus den vier oben herausgestellten Kompetenzklassen. Kompetenzen entwickeln sich dabei über Aktivitäten und Tätigkeiten und beziehen sich auf Fähigkeitsmuster, die durch Fertigkeiten entwickelt werden. In Freys Modell besteht ein enger, wechselseitiger Bezug zwischen Fach- und Methodenkompetenzen, die jeweils mit Sozialkompetenzen verbunden sind. Darüber hinaus hängen die Fach- und die Sozialkompetenzklasse strukturell mit der Personalkompetenzklasse zusammen. Besondere Betonung liegt auf der Fachkompetenzklasse, die bei Frey eine zentrale Position einnimmt. Es wird ein Netzwerk der vier Kompetenzklassen aufgespannt, welches inhaltlich plausibel erscheint, aber erst einmal auf einer hypothetischen Ebene verbleibt. Eine empirische Überprüfung auf der Basis der Selbsteinschätzungen von Studierenden des Lehramtes zeigte heterogene Ergebnisse (Frey 2004). Das Modell konnte nur bei Studierenden bestätigt werden,

die in allen vier Kompetenzklassen hohe Selbsteinschätzungen[9] aufwiesen. Diese Befunde können als Indiz dafür interpretiert werden, dass die Kompetenzklassen sich in Abhängigkeit der Güteausprägung wechselseitig beeinflussen. Die inkonsistenten Ergebnisse lassen Fragen offen, inwieweit das Modell in seiner Betonung der Fachkompetenzen als Referenzgröße für die anderen Kompetenzklassen empirische Relevanz beanspruchen kann. Durch den Einsatz anderer, objektivierender Verfahren (z.B. Fremdeinschätzungen oder Testverfahren) wäre dem weiter nachzugehen. Da das Modell jedoch nicht umfassend falsifiziert wurde, erfährt es weiterhin eine breite Rezeption.

3.7 Kompetenz und Expertise

Entwickelt man Indikatoren, die Fertigkeitsdimensionen abbilden und so die Untersuchung von Kompetenzen erlauben, gelangt man zur Unterscheidung von aufeinander aufbauenden Kompetenzklassen oder auch -stufen. Hier werden Kompetenzen als Kontinuum gedacht, in dem auch „Inkompetenz“ enthalten ist. Die detailliertesten Untersuchungen zur Kompetenzentwicklung stammen aus der Expertiseforschung (Gruber 1997, 1998, 2000; Chi 1978; Chi, Feltovich & Glaser 1981; Chi, Glaser & Rees 1982; Chi, Glaser & Farr 1988; Mandl, Gruber & Renkl 1992; Mandl & Gruber 1996).

3.7.1 Expertiseforschung

Der Expertisebegriff[10] wird als Obergriff für die Kennzeichnung erfolgreicher Lernprozesse bzw. Kompetenzentwicklung verwendet. Eine eindeutige Definition erweist sich aber als schwierig, weil diesbezüglich viele Kriterien berücksichtigt werden müssen. So umfasst Expertise in einer bestimmten Domäne immer eine umfangreiche Wissensbasis, metakognitive Kontrolltechniken, Erfolg beim Lösen von Problemen, Fehlerfreiheit etc. Die Liste ließe sich noch problemlos um andere Kriterien erfolgreichen Lernens und Handelns erweitern. In einer psychologischen Annäherung an den Begriff sind sowohl allgemeinpsychologische (z.B. umfangreiche Wissensbasis) als auch differentialpsychologische Kriterien (z.B. herausragender Erfolg) auszumachen, wodurch die Formulierung einer klaren Definition erschwert wird (Gruber & Mandl 1996, S. 584).
Ziel der in den 1970er Jahren aufkommenden Expertiseforschung war die Analyse von Determinanten, die sich auf die Kompetenzentwicklung auswirken. Dabei konnte etwa mit Bezug auf Schachspieler gezeigt werden (de Groot 1965), dass Expertise im hohen Maße inhalts- bzw. domänenspezifisch ausgebildet ist und sie dem Experten nicht notwendigerweise explizit bewusst sein muss; d.h. Experten können

[9] Die Kompetenzerhebung über Selbsteinschätzungen stellt methodisch keine unproblematische Variante dar. So kann nicht immer vorausgesetzt werden, dass die Probanden in der Lage sind, die Konstrukte objektiv und valide einzuschätzen.
[10] Experten unterscheiden sich von Novizen u.a. durch hervorragende Gedächtnisleistungen. Die Rolle des Gedächtnisses und diesbezügliche Modelle, die verschiedene Ebenen und Funktionen des Gedächtnisses konzeptualisieren, sind in der Expertiseforschung weitgehend erforscht. Sie werden hier aber nicht weiterverfolgt, weil sie im Kern nicht Gegenstand videobasierter Untersuchungen sind.

selbst nicht unbedingt Begründungen ihrer Expertise anführen. Die Ergebnisse zeigen, dass keine allgemeine Transferfähigkeit existiert, die in speziellen Situationen zur Anwendung gebracht werden kann (Prenzel & Mandl 1992). Expertise gilt in Anlehnung an Posner (1988) hingegen als herausragende Leistung in einem bestimmten Gebiet, die sich dadurch auszeichnet, dass sie replifizierbar ist und nicht zufällig auftritt.
Die Experten-Novizen-Forschung hat die Leistungen von Experten und Anfängern in verschiedenen Domänen, wie z.B. der Physik, des Programmierens oder der Betriebswirtschaftslehre miteinander verglichen (Chi, Feltovich & Glaser 1981; Chi, Glaser & Rees 1982; Mandl, Gruber & Renkl 1992). Im Ergebnis wurde aufgezeigt, dass Leistungsunterschiede sich nicht ausschließlich durch übergreifende Fähigkeiten, wie z.B. Intelligenz oder logisches Denken, erklären lassen. Entscheidend scheinen vielmehr die kontextspezifischen Wissensbasen zu sein, ihr Umfang, ihre Strukturierung und flexible Vernetzung (Dörig 2006). Die Problemlöseforschung verwies ebenfalls auf wissens- und kontextbasierte Problemlösestrategien. Erfolgreiches Problemlösen bezieht sich immer auf spezifische Problemstellungen. Schon kleinere Variationen einer Problemstellung bereiten Probanden z.T. unlösbare Schwierigkeiten (Dörner 1979, 2007). Dörner begründet diese Transferschwierigkeiten mit falschen mentalen Repräsentationen von komplexen Problemräumen. Vor diesem Hintergrund halten Lass & Lüer (1990, S. 300) fest:

„Experte sein im Problemlösen bedeutet demnach immer, Experte für einen Problembereich mit spezifischen Anforderungen zu sein."

Dargelegte Forschungsergebnisse lenken den Blick auf für die Problemlösung notwendige Wissensbasen und damit auf die Bedeutung des Vorwissens. Dieses lässt sich in begriffliches Wissen und metakognitives Wissen differenzieren. Begriffliches Wissen umschließt das gesamte Faktenwissen sowie das Sprachwissen über Syntax und Textverstehen, das die Grundlage für inhaltliche Lernprozesse darstellt. Metakognitives Wissen besteht aus den drei Elementen: strategisches Wissen, Wissen über Ziele und Pläne und aufgabenspezifisches Wissen (Dörig 2006, S. 57).
Strategisches Wissen bezieht sich auf Prozesse des Lernens und Denkens. Zu prototypischen Denk- und Lernstrategien gehören Problemlösestrategien, Strategien des kritischen Denkens und Entscheidungsstrategien. Sie ermöglichen eine angemessene Repräsentation des Problemraumes und lassen sich adaptiv anwenden.

„Lernstrategien unterstützen den Prozess des Lernens, indem sie die Auswahl, die Aufnahme, das Speichern, den Abruf und das Anwenden von Wissen unterstützen. Lern- und Denkstrategien sind nicht universell anwendbar" (Dörig 2006, S. 57).

Lern- und Denkstrategien können jedoch nur dann adäquat eingesetzt werden, wenn sie mit realistischen Selbsteinschätzungen korrespondieren. Unrealistische Einschätzungen etwa in Form von negativen Selbstbewertungen verhindern hingegen ihren effektiven Einsatz.

Wissen über Ziele und Pläne schafft Transparenz im Lösungshandeln, setzt aber voraus, dass der Lernende über Lernziele informiert ist und sein Verhalten diesbezüglich adaptieren kann. Erst dann besteht die Möglichkeit, die Bedeutung einer Aufgabe einzuschätzen und individuelle Formen der Selbstüberprüfung auszuwählen.
Selbst eingeschätzte Ausprägungen von Aufgabentypen z.B. dahingehend, wie das kognitive Anspruchsniveau gelagert ist, spiegeln sich schließlich im aufgabenspezifischen Wissen. Auf der Basis dieses Wissens werden die Anforderungen einer Lernaufgabe in Richtung auf Schwierigkeits- und Vertrautheitsgrad subjektiv interpretiert. Die Interpretation kann entscheidende Folgen für den Lernprozess haben, indem es z.B. zum Lernabbruch oder auch zur Motivationssteigerung kommt.
Die Integration von bereichsspezifischem und metakognitivem Wissen begünstigt die Ausbildung von Transferwissen. Die Reichweite des Transfers muss aber als eher begrenzt angesehen werden (vgl. Kap. 2.5.1). Es kommt zu wechselseitigen Beeinflussungen von bereichsspezifischem und metakognitivem Wissen, wobei die Relationen als hochkomplex anzusehen sind. Sie lassen sich vorsichtig im Rückgriff auf die Ergebnisse der Lehr-Lernforschung folgendermaßen beschreiben (Friedrich & Mandl 1992, S. 13 f.).

- Die Menge des bereichsspezifischen Wissens wirkt sich positiv auf kognitive Problemlösungen aus.
- Metakognitives Wissen ist ein wesentliches Merkmal des Erwerbs und der Anwendung von bereichsspezifischem Wissen.
- Eine Erhöhung des bereichsspezifischen Wissens führt zu einer Erhöhung des metakognitiven Wissens.
- Angemessene Strategien der Kontrolle und Regulation wirken sich positiv auf kognitive Lernprozesse aus.

Die Relationen zwischen bereichsspezifischem und metakognitivem Wissen sollen nicht als lineare Wirkgrößen verstanden werden, sie lassen sich aber empirisch nachweisen. Die Art der Qualität dieses Wissens umfasst drei Bereiche (Dörig 2006, S. 58):

- Deklaratives Wissen (Wissen, über)
- Prozedurales Wissen (Wissen, wie)
- Konditionales Wissen (Wissen, wann)

„Die Erkenntnis, dass Wissen der Schlüssel zu Können ist, löste in der Kognitionspsychologie eine Art kopernikanische Wende aus. Auf die Frage, was jüngere Kinder von älteren, Experten von Novizen und gute von schlechten Schülern unterscheidet, würde die Antwort heute lauten: Qualität und Quantität von Wissen. Es ist das Konglomerat von analytischem Begriffswissen, fallbasiertem Wissen, Regelwissen und automatisiertem Handlungswissen, das Menschen kompetent macht“ (Stern 2006, S. 141).

In informationstheoretischen Ansätzen der Kognitionspsychologie stellt der Zusammenhang der verschiedenen Wissenstypen (bereichsspezifisch und

metakognitiv) die Grundlage für die Entwicklung vom Novizen zum Experten dar. Schlüssige Aussagen über den Entwicklungsverlauf lassen sich am besten anhand wohl definierter Domänen, wie dem bereits an früherer Stelle angeführten Schachspiel oder Programmieren, begründen.
Die von Anderson formulierte ACT-Theorie des Fertigkeitserwerbs kann hier als prominentes Beispiel angeführt werden (Anderson 1982, 1983, 1987). Sie erklärt den prozesshaften Erwerb von Fertigkeiten in drei Stufen. Auf der ersten (deklarativen) Stufe geht es um die domänenspezifische Wissensenkodierung mittels allgemeiner Strategien. Auf der zweiten Stufe (Kompilation) wird das deklarative Wissen in prozedurales Wissen umgewandelt, indem auf der Basis von Problemlösehandeln Produktionsregeln gewonnen werden, die Lösungsmöglichkeiten für den Problemraum bieten. Die dritte Stufe („Tuning“) kann als metatheoretische Evaluation und Adaption der Produktionsregeln und Produkte beschrieben werden. Die diesbezüglichen Wissensrepräsentationen ordnet Anderson drei verschiedenen Gedächtnistypen zu: dem Arbeitsgedächtnis, dem assoziativen Langzeitspeicher (deklaratives Gedächtnis) und dem frei programmierbaren Produktionssystem (prozedurales Gedächtnis). In der von ihm entwickelten ACT-Theorie baut das prozedurale Wissen systematisch auf dem deklarativen Wissen auf und wird kontextspezifisch adaptiv erworben bzw. prozeduralisiert. Die Theorie beschreibt den Geneseprozess prototypisch präzise. Problematisch erscheint allerdings die Anwendung auf schlecht[11] definierte Domänen, wie sie im Lehrerhandeln vorzufinden sind.

3.7.2 Expertiseforschung und berufliche Handlungskompetenz

Renkl & Nückles (2006) wenden den Ansatz der Expertiseforschung auf berufliche Handlungskompetenzen an und diagnostizieren z.T. vorfindbare „träge Kompetenzen“[12]. Berufliche Handlungskompetenz wird dabei nicht nur „als das effektive Erledigen von beruflichen Standardanforderungen“, sondern vor allem als „erfolgreiche Anpassung an neuartige Anforderungen“ thematisiert. Zudem ist sie nicht auf bereichsspezifische Wissensbestände beschränkt, sondern schließt auch nicht-kognitive Faktoren mit ein, so z.B. muss ein Arzt neben medizinischen Fachkenntnissen die Fähigkeit haben mit schwierigen Patienten umgehen zu können.
Renkl & Nückles diskutieren, inwieweit Kompetenzen übergreifend oder eher kontextgebunden gedacht werden können. Sie kommen zu dem Ergebnis, dass im Kompetenzbegriff eine gewisse „Generalität“ angelegt ist, betrachten Kompetenzen aber dennoch als vorrangig kontextgebunden. Beim Verlassen dieses Kontextes

[11] Schlecht definierte Domänen zeichnen sich durch folgende Merkmale aus: (1) Es existieren keine Regeln oder Prinzipien, die alle Fälle umspannen. (2) Hierarchische Verbindungen innerhalb der Wissensbasis können von Fall zu Fall auch umgedreht werden. (3) Gleiche Merkmale können in unterschiedlichen Kontexten Unterschiedliches bedeuten. (4) Es gibt multiple hierarchische Organisationsprinzipien von Merkmalskomplexen (Gruber & Mandl 1996, S. 592).
[12] Die Autoren verbinden die Begriffe „träge“ und „Kompetenz“ miteinander, um aufzuzeigen, dass es Entwicklungsprozesse gibt, die Kompetenzausprägungen auch blockieren. Die Verbindung mutet aber paradox an, da man bei „trägen Kompetenzen“ besser nicht von Kompetenzen sprechen sollte.

werden die Kompetenzen dann „träge“ im Sinne von nicht anwendbar. Dafür nennen die Autoren drei Begründungen (Renkl & Nückles 2006, S. 181):

„a) Das Erlernte passt nicht zum Anwendungskontext; b) Obwohl das Erlernte in passender Weise vorhanden ist, können andere bereichsspezifische Voraussetzungen fehlen, um es zur sinnvollen Anwendung zu bringen; c) Es fehlt Metawissen um die Anwendung des Erlernten.“

Diese drei Gründe können also Expertiseausprägungen in beruflichen Handlungskontexten entgegenstehen. Nach Meinung der Autoren bezieht sich die Kontextgebundenheit auf alle Kompetenzklassen, auch auf moralische Werthaltungen.
Berufliche Expertise wird von Renkl & Nückles als Ergebnis von langjähriger Ausbildung und Enkulturation in die Expertengemeinschaft dargestellt, in deren Folge der Experte fachliche Wahrnehmungs- und Denkschemata erwirbt, die ihn befähigen, komplexe fachliche Probleme schnell zu erkennen und effizient zu lösen. Die Verwendung von Fachbegriffen dient dabei nicht nur der Kommunikation, sondern strukturiert die Perspektive auf einen Gegenstand. Ein Unterschied zwischen Experten und Laien kann also nicht zuletzt an der Fachsprache festgemacht werden, was die Kommunikation zwischen diesen Gruppen erschwert.
Auch Gruber, Harteis & Rehrl (2006) stellen die Bedeutung von „Erfahrung“ als Grundlage für die Entwicklung von Handlungskompetenzen heraus. Berufliche Tätigkeiten werden dabei als ein Möglichkeitsraum zur Entwicklung von Expertise konzipiert. Im Rückgriff auf die Studien von Gruber & Rehrl (2005) und Stark et al. (1995) entwerfen die Autoren ein Modell beruflicher Handlungskompetenz, das vier Komponenten beinhaltet:

- Umfangreiches Wissen und gutes Gedächtnis
- Problemlösefähigkeit und Entscheidungsstärke
- Verfügbarkeit von Routinen
- Einbettung in Expertengemeinschaften

Dieses Modell orientiert sich allerdings am Wissen von Experten und setzt bestimmte Expertiseausprägungen bereits voraus. Leutner (1992) konnte mit Bezug auf das Lösen komplexer computerbasierter Simulationen kaum Zusammenhänge zwischen den verschiedenen Komponenten von Handlungskompetenz bei Probanden geringer Expertiseausprägung feststellen. Die Ergebnisse decken sich mit der weiter oben vorgestellten Studie von Frey (2006). Sie deuten auf die Notwendigkeit, bei der Förderung von Handlungskompetenz in beruflichen Kontexten die Integration von Komponenten anzustreben.

Wenn Gruber, Harteis & Rehrl (2006, S. 197) als voraussetzende Bedingung für die Entwicklung von Expertise „Erfahrung“ (vgl. Kap. 6.1) nennen, ist diese nicht in einem umgangssprachlichen Sinne quantitativ definiert, sondern qualitativ als „episodische Kenntnis über den Umgang mit Wissen, also darüber, wie, wann, und in welcher Situation welches Wissen am Erfolg versprechendsten zur Anwendung kommt“. Erfahrungen stellen demnach Katalysatoren für die Reorganisation und

Integration von Wissensstrukturen dar. Gruber (1999) stellt folgende Merkmale für den Erfahrungsbegriff heraus:

- Erfahrung hängt mit dem Besitz umfangreichen domänenspezifischen Wissens zusammen.
- Erfahrung hängt vor allem mit Wissen zusammen, das durch episodische, selbst erlebte Ereignisse entsteht und das von den Individuen aufgrund dieser Erlebensprozesse konstruiert wird.
- Erfahrung entsteht durch episodische Erlebnisse nur dann, wenn diese vom Individuum subjektiv als (selbst-)relevant und selbstbezogen eingeschätzt werden.
- Erfahrung kann gefördert werden, wenn ein Möglichkeitsraum besteht, Episoden zu erleben und wenn die individuelle Relevanz des Lerngegenstandes vermittelt wird, was durch das Bereitstellen komplexer Lernumgebungen geschehen kann.

Die Ermöglichung situierter Problemstellungen, als von Individuen subjektiv als relevant erlebbare Handlungsherausforderungen, kann damit als Konstituente für die Expertiseentwicklung in beruflichen Kontexten bezeichnet werden.
Ein weiteres Moment bringen Preckel & Holling (2006) schließlich in die Diskussion, indem sie die Rolle der Intelligenz als Prädiktor für die Entwicklung beruflicher Expertise herausstellen. Sie vertreten ein Verständnis von Intelligenz in Bezugnahme auf Wechsler (1956), welches ihrer Meinung nach Anknüpfungspunkte zum Kompetenzbegriff nahe legt. Danach sind affektive und motivationale Merkmale im Intelligenzkonstrukt mit eingeschlossen. Die Vorhersagefähigkeit des Konstruktes für berufliche Expertise nimmt mit steigender Komplexität und höheren kognitiven Anforderungen in beruflichen Handlungsfeldern zu (vgl. Preckel & Holling 2006, S. 175).

3.7.3 Prototypische Modellierung von Expertise

Dreyfuss und Dreyfuss (1986) haben ein hierarchisches Stufenmodell entwickelt, das den Weg vom Novizen zum Experten prototypisch in fünf Stufen beschreibt:

- Novize
- Fortgeschrittener Anfänger
- Kompetenter
- Erfahrener (Gewandter)
- Experte

Wesentliche Merkmale des Modells beziehen sich auf die ganzheitliche Wahrnehmung von Problem- und Anforderungssituationen und die dem Individuum zur Verfügung stehenden Ressourcen zur situationsadäquaten Bewältigung von Problemen und Anforderungen. Über die fünf Stufen werden die Quellen, aus denen sich die Ressourcen speisen, thematisiert. Danach entwickeln sich Expertiseressourcen ausgehend von der Befolgung kontextfreier Regeln im

Novizenstadium über die Vermischung von theoretischem Wissen und Erfahrungswissen im Kompetenzstadium, bis es zum vollkommen situationsadäquaten Expertenhandeln kommt. Der Lernende durchläuft eine Entwicklung vom Außenstehenden zum aktiv Beteiligten und Gestaltenden. Der Weg zum Expertenhandeln zeichnet sich durch einen Zugewinn an Flexibilität und Spontaneität aus, wobei durch die Gewinnung von Routinen das psychische Belastungserleben sukzessive kleiner wird. Diese Verringerung empfundener Belastungen macht den Weg frei für einen unbeschwerten, flexiblen und spontanen Umgang mit Problemstellungen, die verantwortungsvoll und im Einklang mit professionstheoretischen Zielsetzungen umgesetzt werden können. Lernprozesse gelten dabei als kontinuierlicher Vorgang.
Zusammenfassend generiert sich Expertise über eine zunehmend holistische Wahrnehmung, die Anreicherung von theoretischem Wissen mit Erfahrungswissen und die Ausbildung von Routinen, die psychisch entlastend wirken und damit die Möglichkeiten schaffen, Verhalten flexibel und adaptierend für Problemlösungen zu gestalten. Das Modell will allgemein die Entwicklung von Expertise beschreiben und lässt sich wegen der Betonung einer ganzheitlichen Wahrnehmung sinnvoll auf Lehrerhandeln gemäß einer Professionalisierung anwenden (Koch-Priewe 2002). Problematisch erscheint dabei der einseitig positiv dargestellte Nutzen von Routinen, die zwar durch ihre Entlastungsfunktion einerseits Flexibilität erzeugen, andererseits aber auch die Gefahr in sich bergen, rigide zu wirken und eine situative Adaption an jeweilige Erfordernisse zu verhindern.
Routinen verweisen auf ein implizites Wissen von Experten, welches z.T. nicht selbst expliziert werden kann. Es gerät auch unter den Stichworten „intuitives Handeln“ bei Dreyfuss und Dreyfuss (1986) oder „tacit knowledge“ bei Polanyi (1985) in den Blick. Routinen, verstanden als automatisierte Handlungen, können das Ergebnis verschiedener Geneseprozesse sein (Koch-Priewe 2002). Entstehen sie in der unbewussten Anpassung oder Nachahmung eines Modells, so sind sie nur schwerlich verbalisierbar und veränderbar. Damit nehmen sie unbeweglichen und starren Charakter an und führen zu stereotypen Handlungsmustern. Koch-Priewe (ebd.) sieht darin einen gewichtigen Einwand gegen die Einführung einer „Meisterlehre“, d.h. der konzentrierten und nachahmenden Orientierung von Auszubildenden am Handeln von erfahrenen Mitarbeitern. Jedoch können Routinen auch das Ergebnis von bewussten und reflexiven Lernprozessen in Praxissituationen sein. In diesem Fall bleiben sie flexibel, veränderbar und anschlussfähig an sich verändernde Kontexte oder neue Erkenntnisse. Die Entwicklung von Expertise im Spannungsfeld von Flexibilität und Rigidität verweist auf notwendig werdende externe Beratungsleistungen.

3.8 Verfahren zur Kompetenzmessung

Kompetenzen können nur auf der Grundlage adäquater Definitionen und Modelle diagnostiziert werden. Die damit verbundenen Schwierigkeiten sind oben diskutiert worden. Es lässt sich ein Konsens darstellen, nach dem Kompetenzen nur auf einem Performanzniveau als latente Variablen einer Untersuchung zugänglich gemacht werden können. Damit ergeben sich eine Vielzahl von Methoden, die für sich

beanspruchen Kompetenzen bzw. Performanzen feststellen zu können. Einen Überblick über gängige Verfahren in verschiedenen Fragekontexten geben Erpenbeck & Rosenstiel (2003). Dabei betonen sie:

„Kompetenzforschung kann und muss sich aller Mess-, Charakterisierungs-, und Beschreibungsverfahren bedienen, die von der Persönlichkeitspsychologie und Soziologie, der Arbeits- und Handlungspsychologie, der Performanzanalyse, der Qualifikationsforschung sowie von Sozialpsychologie, Kommunikationspsychologie, Sprachwissenschaft, Erziehungswissenschaft, Pädagogik usw. zur Verfügung gestellt werden“ (Erpenbeck & Rosenstiel 2003, S. XXII).

Die verschiedenen Verfahren berücksichtigen notwendigerweise die Differenzierung von Kompetenzklassen und -gruppen und lassen sich bezüglich einer eher objektiven oder subjektiven Beobachtung von Kompetenz unterscheiden. Methodisch resultiert daraus die Präferenz von qualitativen oder quantitativen Methoden der Kompetenzforschung, je nachdem ob eher formative bzw. eher summative Ziele verfolgt werden. Insgesamt können die Verfahren Kompetenzen

- quantitativ messen (z.B. Tests),
- qualitativ charakterisieren (z.B. Begriffsanalysen),
- komparativ beschreiben (z.B. Kompetenzbiographien) (ebd., S. XXVI).

Somit genügen nicht alle Verfahren den psychometrischen Gütekriterien von Objektivität, Validität und Reliabilität. In der Praxis haben sich oftmals die nicht methodisch exaktesten Verfahren durchgesetzt. Erpenbeck & Rosenstiel plädieren deshalb dafür, ein Verständnis von sozialer Validität, im Sinne von sozialer Akzeptanz, für die Beurteilung von Kompetenzmessverfahren zu berücksichtigen. Konkret werden in verschiedenen Untersuchungen methodisch folgende Erhebungsverfahren eingesetzt: interview- und fragebogenbezogene Selbst-, Fremd- und Gruppenbeurteilungsverfahren, Unterrichtsproben und Assessment-Center, Entwicklungsaufgaben, Entwicklungsportfolios, soziale Intelligenztests, Bildertests und projektive Verfahren (Frey 2006).

3.8.1 Kompetenzmessung durch Selbstbeurteilungen

Häufig werden Kompetenzen in Form von Selbstbeurteilungen erhoben. Dieser Forschungspraxis liegt ein forschungsökonomisches Argument zugrunde. Selbstbeurteilungsverfahren lassen sich häufig schnell und effizient einsetzen. Dabei kommen in der Regel Fragebogen- und Interviewverfahren zur Anwendung. Selbsteinschätzungen können zudem mit Fremdbeurteilungen ergänzt und validiert werden. Darüber hinaus wird von vielen Autoren die Auffassung vertreten, dass die betroffenen Personen ihre Kompetenzausprägungen selbst am besten beurteilen können (vgl. Amelang & Bartussek 2001). Die Selbstbeurteilung von Probanden basiert oftmals auf einer autobiographischen Perspektive, was z.B. bei der längsschnittlichen Beurteilung von Kompetenzentwicklung Vorteile mit sich bringt. Demgegenüber besteht der Nachteil dieses Erhebungsverfahrens in der möglichen Orientierung am Kriterium der „sozialen Erwünschtheit“. Speziell bei

Kompetenzmessungen, die unmittelbare Folgen für den Probanden haben (z.B. hinsichtlich der beruflichen Karriere), scheint der Einsatz von Selbstbeurteilungsverfahren unangemessen. Zudem stellen sie hohe Anforderungen an die Reflexionsfähigkeit der Probanden. Es kann nicht pauschal vorausgesetzt werden, dass Befragte eine objektive Retrospektion zu leisten in der Lage sind bzw. ein notwendiges Konstruktverständnis vorliegt. Besonders für den Kompetenzbegriff mit seinen vielen Konnotationen, d.h. für die Bewertung von Kompetenzdimensionen ist dies evident. Die Selbsteinschätzung in beruflichen Handlungsfeldern kann auch durch externe Konstrukte, wie z.B. Zufriedenheit in der Berufsausübung, verzerrt sein. Selbsteinschätzungen fehlt zudem eine vergleichende Perspektive, was ebenfalls individuelle Verzerrungen zur Folge haben kann.

3.8.2 Kompetenzmessung durch Fremdeinschätzungen

Aufgrund der Unschärfegefahr im Zusammenhang mit Selbsteinschätzungen werden in beruflichen Handlungsfeldern häufig Fremdeinschätzungsverfahren eingesetzt, weil man ihnen eine vergleichsweise hohe Objektivität zumisst. Fremdbeurteilungen gelten insofern als Expertenbeurteilungen, die in der konkreten beruflichen Tätigkeitsausübung oder in Tätigkeits-Simulationen durchgeführt werden. Ein zusätzlicher Vorteil solcher Assessment-Center liegt darin, dass sie Interaktionen[13] mit den einzuschätzenden Probanden bieten, was die Validität von Untersuchungen erhöhen kann.
Expertenbeurteilungen setzen allerdings voraus, dass Fremdbeurteiler auch tatsächlich zu Experten in dem einzuschätzenden Gebiet zählen, womit impliziert wird, dass der Experte gegenüber dem einzuschätzenden Probanden über einen Kompetenzvorschuss verfügt. In der Praxis ist dies nicht immer gewährleistet und unterliegt häufig keiner Kontrolle. Fremdbeurteilungen stellen zudem hohe Anforderungen an die diagnostischen Fähigkeiten der Beurteiler. Prinzipiell bieten sie die Möglichkeit komplexe Verhaltensweisen in Tätigkeitsausübungen zu beobachten, einzuschätzen und davon ausgehend auf Kompetenzen zu schließen. Damit ist die Hoffnung auf eine hohe prognostische Validität verbunden (Frey 2006). Die Anwesenheit von Beurteilern kann allerdings auch Invasivitätseffekte mit sich bringen, die die Situation und damit die Beurteilung beeinflussen. Wie Selbsteinschätzungen beinhalten auch Fremdbeurteilungen das zu möglichen Verzerrungseffekten führende Problem sozialer Erwünschtheit und Akquieszenz. Solche Effekte betreffen sowohl das gezeigte Verhalten der Probanden als auch die Einschätzungen der Beobachter.

3.8.3 Kompetenzmessung durch Portfolioansätze

Portfolioansätze stellen in bestimmter Weise eine Verbindung von Selbsteinschätzungen und Fremdbeurteilungen dar. Der Fokus richtet sich bei diesen Ansätzen auf die Bearbeitung von Entwicklungsaufgaben (Gruschka 1995) bzw. das

[13] Die Möglichkeit der Interaktion mit einzuschätzenden Probanden stellt allerdings kein Wesensmerkmal für Fremdeinschätzungen dar.

Anlegen von Entwicklungsportfolios (Lissmann 2002). Solche Verfahren wollen über die Untersuchung von selbst entwickelten Produkten oder Projekten der Probanden Rückschlüsse auf deren Kompetenzen ziehen. Sie haben sich an den Hochschulen und in der betrieblichen Praxis etabliert. Speziell in komplexen beruflichen Handlungsfeldern ist die Portfolioarbeit zentral geworden. Die Probanden erhalten die Möglichkeit ihre Lernverläufe zu dokumentieren, zu präsentieren und damit eine Selbsteinschätzung vorzunehmen. Unter der Perspektive von Selbstorganisationsprozessen kann dieser Ansatz überzeugen. Die letztendliche Einschätzung von Produkten und Projekten wird allerdings als Fremdbeurteilungsverfahren durchgeführt. Angesichts der mit Portfolioansätzen verbundenen Freiheitsgrade stellt sich die Frage nach verbindlichen Kriterien, die die Güte von Portfolios und daraus abzuleitenden Aussagen über Kompetenzen einschätzbar machen. Insofern genügen sie häufig nicht den Testgütekriterien und erschweren eine Vergleichbarkeit der Ergebnisse. Der zu betreibende Aufwand ist zudem auf allen Seiten hoch und kann kaum für größere Populationen betrieben werden.

3.8.4 Kompetenzmessung durch Testverfahren

Testverfahren, die den Gütekriterien von Objektivität, Validität und Reliabilität genügen, stellen eine elaborierte Möglichkeit der Kompetenzmessung dar. Die Vorteile von Testverfahren hinsichtlich ihrer allgemeingültigen Aussagekraft liegen in der hohen Standardisierung, der Einhaltung von Gütekriterien und der genormten Interpretation der Ergebnisse.
Es liegt eine Vielzahl von Testverfahren vor, die z.B. kognitive, emotionale und soziale Intelligenz messen (Erpenbeck & Rosenstiel 2003). Allerdings besteht kein Konsens darüber, inwieweit diese Konstrukte tatsächlich Kompetenzen abbilden, da sie normorientiert und nicht kriteriumsorientiert entwickelt worden sind. Die Gütekriterien der klassischen Testtheorie (Objektivität, Validität und Reliabilität) führen u.a. zu einer Konstruktspezialisierung, die dem Kompetenzbegriff nicht unbedingt gerecht wird. Die klassische Testtheorie, im Kern eine Messfehlertheorie, stößt hier an Grenzen. Um Rückschlüsse auf Kompetenzen ziehen zu können, müssen somit ganze Testbatterien hinzugezogen werden, was forschungsökonomisch häufig nicht zu bewerkstelligen ist. Ergebnisse von Testverfahren können als Indikatoren theoretischer psychologischer Konstrukte aufgefasst werden. Was insgesamt ihre Relevanz für Kompetenzuntersuchungen angeht, lassen sich konträre Positionen ausmachen. Einerseits würden die Ergebnisse von Intelligenz- und Persönlichkeitstests darüber Aufschluss geben, welches Potential jemand besitzt, um Kompetenzen ausbilden zu können (Lang-von Wins 2003). Andererseits stellt aber etwa Frey (2006, S. 132) dieser Auffassung entgegen:

„Intelligenztests scheiden bei der Beurteilung von Kompetenz häufig deshalb aus, da keine Klassifizierungen vorgenommen werden, sondern sich der Fokus auf das Ausmaß der individuellen oder kollektiven Anwendung von Kompetenz unterschiedlicher Individuen und Personengruppen bezieht. Gehaltvoll werden die über unterschiedliche Testverfahren gewonnenen Kompetenzwerte, die

einen angenäherten Ist-Zustand beschreiben sollen, dann, wenn sie mit einem Soll-Zustand, also mit festgelegten zu entwickelnden Kriterien verglichen werden können."

Für die Entwicklung von Testverfahren zur Messung von Kompetenzen eignet sich eher die Probabilistische Testtheorie bzw. die Item-Response-Theorie (IRT), weil die darauf aufbauenden Testverfahren Aussagen über individuelle, kriteriumsorientierte Leistungen liefern können, die wiederum Schlussfolgerungen über latente Merkmalsgruppen und damit Kompetenzen erlauben. Die Konstruktion solcher Testverfahren bedingt es, dass die Lösungswahrscheinlichkeit eines Items in Abhängigkeit des Schwierigkeitsgrades desselben und der latenten Fähigkeiten des Probanden ermittelt wird. Probabilistische Testverfahren weisen also quantitativen Leistungswerten qualitative Personenmerkmale zu (Rost 2005a, b). Einsatz finden dabei z.B. eindimensionale Skalen, sog. Rasch-Skalen, die Domänen oder Kompetenzstufen abbilden. Die auszumachenden Abstufungen in ihrer spezifischen Reihenfolge lassen Aussagen über lösungsrelevante kognitive Prozesse zu. Um Kompetenzklassen umfassend abbilden zu können, müssen allerdings mehrere Rasch-Skalen eingesetzt werden, womit sich der forschungsmethodische Aufwand entscheidend erhöht.
Obwohl die Probabilistische Testtheorie bzw. Item-Response-Theorie nicht neu ist, hat sie sich bislang nicht auf breiter Linie gegenüber der klassischen Testtheorie durchsetzen können. Der zu betreibende Aufwand für die Konstruktion entsprechender Testverfahren ist allerdings auch sehr hoch. Deshalb liegen zurzeit nur wenige Verfahren in dieser Richtung vor, was sich vermutlich angesichts der allgemein hohen Bedeutsamkeit der Kompetenzthematik zukünftig ändern wird.

3.9 Fazit

In diesem Kapitel sollten Kompetenzmodellierungen und operationale Kompetenzdefinitionen der psychologischen und erziehungswissenschaftlichen Forschung dargestellt und in ihrer empirischen Anschlussfähigkeit diskutiert werden. Beabsichtig war damit eine grundlegende theoretische und empirisch fundierte Absicherung des Kompetenzbegriffes. Es konnte gezeigt werden, dass das Konstrukt „Kompetenz" z.T. sehr unterschiedlich definiert und modelliert wird. Gleichwohl gibt es Annäherungen für Konvergenzen in den theoretischen Herangehensweisen und daraus abgeleiteten empirischen Ansätzen. Die von Weinert formulierte Definition (2001, S. 27f.) bzw. das vorgestellte Klassifikationsmodell von Frey (2006) stehen für derartige Konvergenzen und seien es auch nur pragmatische Konventionen der „science community" aus Plausibilitätsgründen.
In der hier dargelegten Weise erscheint der Kompetenzbegriff theoretisch und empirisch, wenn auch nicht unbedingt unproblematisch, handhabbar. Das diese Arbeit eigentlich interessierende Feld der Lehrerbildung und damit verbunden die Definition und Entwicklung professioneller unterrichtlicher Handlungskompetenz von Lehrkräften wurde zugunsten einer theoretischen Grundlegung bislang bewusst weitgehend ausgeblendet und soll im nächsten Kapitel genauer diskutiert werden.

4 Entwicklung professioneller didaktischer Handlungskompetenz

Bromme (1997) hat in seiner Studie über Kompetenzen und ihre Funktionen im Lehrerhandeln darauf hingewiesen, dass es unter Lehrern und auch in der diesbezüglichen Forschung eine Tendenz gibt, die Möglichkeiten zur Identifizierung von erfolgreichem Lehrerhandeln und eine damit verbundene empirische Analyse von Wirkfaktoren eher pessimistisch einzuschätzen. Dieser Pessimismus lässt sich mit der Komplexität unterrichtlicher Lehr-Lernprozesse erklären. Grundsätzlich kann es so auch als unwahrscheinlich angesehen werden, dass es zukünftig eine Theorie geben wird, die genau abbildet, was notwendige (nicht nur hinreichende) Merkmale und Bedingungen guten Unterrichts sind und wie dementsprechend eine Ausbildung von Lehrern zu konzipieren wäre. Gleichwohl gibt es mittlerweile eine Vielzahl von internationalen empirischen Untersuchungen über Wirkungszusammenhänge in der pädagogischen Praxis und der erziehungswissenschaftlichen Ausbildung, die genauer im nächsten Kapitel dargestellt werden sollen. An dieser Stelle geht es zunächst um die generelle Anschlussfähigkeit der Kompetenzdebatte an pädagogische, konkreter unterrichtliche Handlungsanforderungen.
Über die vorangehend dargestellte Forschung sind Begriffe der Kompetenz und Expertise auch in die Erziehungswissenschaft eingewandert, wo sie z.T. kontrovers diskutiert werden (vgl. Kap. 4.3.1). Aus einer bildungs- und strukturtheoretischen Perspektive erscheinen die mit dem Kompetenzbegriff verbundenen Konzepte als eine Verkürzung und Einschränkung pädagogischer Handlungsmöglichkeiten und der sie betreffenden Diskurse (Herzog 2005). Die nachfolgenden Abschnitte arbeiten daher Unterschiede, aber auch Gemeinsamkeiten der hier zur Rede stehenden Forschungs- und Theoriekulturen heraus und beziehen sie wechselseitig aufeinander. Die Integration von Kompetenzkonzepten in den erziehungswissenschaftlichen Diskurs führt zur Möglichkeit, Kompetenzen mit Bezug auf pädagogisches Handeln zu messen und damit zu der Frage, inwieweit die Implementierung von Standards in die Lehrerbildung positive oder negative Potentiale eröffnet (vgl. Kap. 5).

4.1 Paradigmen empirischer Forschung in didaktischen Handlungsfeldern des Lehrerberufes

Für die Beschreibung von forschungsleitenden Perspektiven der empirischen Untersuchung von Lehr-Lernprozessen hat sich der Begriff des „Paradigma“ durchgesetzt (Bromme 1997; Ingenkamp 1970; Gage 1963). Paradigmen können als Verdichtung verschiedener Hypothesen über empirische Wirkungszusammenhänge bezeichnet werden, andersherum speisen sie sich aus den Ergebnissen empirischer Forschung. Damit unterliegen sie einem zeithistorischen Wandel und können rekonstruiert werden, wobei teilweise sich überschneidende Ausprägungen an den Übergängen von noch geltenden, etablierten und schon neu entstehenden Paradigmen auszumachen sind.

4.1.1 Paradigma der Lehrerpersönlichkeit

Empirische Forschungsergebnisse zu didaktischen Handlungsfeldern des Lehrerberufes lassen sich bis in die 1950er und 1960er Jahre zurückverfolgen. Die Forschungsbestrebungen orientierten sich zunächst am sog. Paradigma der Lehrerpersönlichkeit, womit impliziert wurde, man müsse zum Lehrer geboren werden. Gegen diese Aussage führte bereits Spranger (1958) Argumente an und behauptete, dass es keinen Beruf gäbe, zu dem man weniger geboren sein müsse als zu dem des Erziehers – womit der Lehrerberuf eingeschlossen war (Hascher 2004). Der zumeist unreflektierte Begriff der Lehrerpersönlichkeit hat sich aber bis heute in Praxiskontexten gehalten. Forschungsansätze auf der Basis dieses Paradigmas wollten Personenmerkmalen im Hinblick auf ihre erzieherische Wirkung nachgehen. Die Wirkungen wurden aber nicht kriteriumsorientiert untersucht, was entweder zu trivialen Ergebnisaussagen führte oder zumindest nicht zur überzeugenden und eindeutigen Klärung von Zusammenhängen komplexer Lehr-Lernprozesse beitrug.
Der Begriff der Persönlichkeit, so kann rückblickend gesagt werden, führte in eine falsche Richtung, weil er nicht auf der Ebene von Verhaltensmerkmalen der Lehrpersonen, sondern global angegangen wurde. Hinzu kam eine unzureichende Methodik. Aktuell stellt Terhart (2006) klar, dass die Suche nach der „guten Lehrerpersönlichkeit“ keine zielführende Perspektive darstellt, weil die Persönlichkeit des Lehrers zwar einen Einfluss auf Schüler haben kann, dieser Einfluss aber bei verschiedenen Schülern unterschiedlich ausfällt.
Das Paradigma kann damit als Grundlage empirischer Forschung als gescheitert betrachtet werden und stößt weitgehend auf Ablehnung. Bromme (1997) macht allerdings darauf aufmerksam, dass es eine Weiterentwicklung des Paradigmas gibt, wonach nicht mehr die Persönlichkeit von Lehrpersonen im Zentrum steht, sondern deren Personeneigenschaften.

> „Ein Beispiel ist die Suche nach Attributionsstilen und Stilen der Leistungsbewertung (z.B. soziale vs. individuelle Bezugsnormen) von Lehrern, die diese Stile aus einer (modernen) differentiellen Perspektive, aber eben doch als Personeneigenschaft des Lehrers betrachtet“ (Bromme 1997, S. 183).

So gewendet hat die Frage nach der unterrichtenden Person nach wie vor aktuelle Bedeutsamkeit. Helmke (2003) weist in diesem Zusammenhang darauf hin, dass nur wenige Daten über Merkmale der Lehrperson und deren Einfluss auf Faktoren der Unterrichtsqualität existieren. Auch Mayr (2003) vertritt die Ansicht einer notwendigen Berücksichtigung personaler Eigenschaften für die Beschreibung von Lehrkompetenz. Er bezieht sich auf den seiner Meinung nach erfolgreichen Einsatz persönlichkeitsdiagnostischer Verfahren als Instrumente der Personalentwicklung. Gleichzeitig konnte er in einer Längsschnittstudie zeigen (Mayr 2005), dass Persönlichkeitsvariablen, erhoben nach dem persönlichkeitspsychologischen Konstrukt der „Big Five“ (Bartussek 1996), eine nicht zu vernachlässigende Varianzaufklärung für die Selbstwahrnehmung unterrichtlicher Kompetenzen, die Berufszufriedenheit und die Bewältigung von Belastungsfaktoren leisten (vgl. Urban

1984, 2005; Baer et al. 2006). Untersuchungsergebnisse von Schaarschmidt & Fischer (2001) weisen in eine ähnliche Richtung.

4.1.2 Prozess-Produkt-Paradigma

Im Zentrum des seit den 1960ern Verwendung findenden Prozess-Produkt-Ansatzes stand zunächst die Frage, wie sich konkrete Instruktionsformen von Lehrpersonen auf die Lernleistungen von Schülern auswirken. Man ging dabei von behavioristischen Vorstellungen aus, wonach sich ein direkter Effekt zwischen Lehrerhandeln und Schülerleistungen messen lässt. Das Ziel dieser Forschungsperspektive bestand in der Aufdeckung von Zusammenhangsmaßen zwischen definierten Elementen des Unterrichts (z.B. Klarheit und Verständlichkeit) und erfassten Zielkriterien (z.B. Lernleistung). Der Ansatz lieferte eine Vielzahl von empirisch gut belegten Ergebnissen bezüglich der Merkmale des Unterrichtens und ihrer Auswirkungen auf der Schülerseite (Brophy & Good 1986; Helmke & Weinert 1997). Mit dem Forschungsprogramm wollte man wirksame und personenübergreifende Fertigkeiten identifizieren und didaktisch nutzbar machen. Dabei wurde festgestellt, dass unterschiedliches Verhalten der Lehrpersonen zu unterschiedlichen Leistungen bei den Schülern führt (Good, Biddle & Brophy 1975) und dass die impliziten und expliziten Lehrererwartungen einen Einfluss auf das Leistungsverhalten der Schüler haben (Rosenthal & Jacobson 1968).
Vertretern des Prozess-Produkt-Paradigmas ist jedoch vorgeworfen worden, unterrichtliche Prozesse einseitig auf Grund von Lehrerhandeln erklären zu wollen und dabei die Schülerperspektive sowie andere Variablen zu vernachlässigen und insofern Ergebnisse zu produzieren, die wesentliche Mediatorvariablen nicht berücksichtigen (vgl. Gage & Needels 1989). Die Kritikpunkte führten zu einer programmatischen Erweiterung des Paradigmas. Die Einbeziehung von komplexen Wirkbeziehungen zwischen Lehrer- und Schülerverhalten sowie die integrierende Berücksichtigung interpretativer Prozesse durch die Schüler (Doyle 1986) können als forschungsstrategische Konsequenzen dieser Erweiterung beispielhaft angeführt werden. Darüber hinaus erfuhr die Perspektive auch eine Ergänzung um außerunterrichtliche Merkmale wie etwa Schulform oder Schulklima. Mit den eingearbeiteten Erweiterungen unterliegt das Prozess-Produkt-Paradigma einer nach wie vor bestehenden Akzeptanz (Helmke 2003, S. 30).

4.1.3 Das Experten-Paradigma

Im Experten-Paradigma spiegelt sich die erhebliche Kontextabhängigkeit erfolgreichen und erfolglosen Lehrerhandelns, wie sie auch schon in den Erweiterungen des Prozess-Produkt-Ansatzes deutlich wurde. Die Entwicklung des Konstruktivismus „vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur“ (Reusser 2006) hat die Forschungsperspektive verändert. In dieser Sichtweise übt der Lehrer nur mehr eine begrenzte direkte Wirksamkeit auf das Schülerlernen aus und sein Aufgabenspektrum wird umformuliert. Die Lehrer-Schüler-Beziehung kann hinsichtlich der jeweils autonomen Vermittlungs- bzw.

Aneignungsaktivitäten als reziproke Relation beschrieben werden. Gleichwohl steht der (erfolgreiche) Experte in der Verantwortung komplexe Lernumgebungen zu entwickeln, die Lerngelegenheiten ermöglichen.
Mit dem Begriff des „Experten“ schließt ein Verständnis des Lehrerhandelns an den Diskurs der Expertiseforschung an. Bromme (1997) konstatiert, dass in den Unterrichtsmodellen von Carroll (1985) und Slavin (1989) eine Orientierung am Expertenansatz bereits angelegt ist. Die einflussreiche Studie von Berliner (1992) konnte spezifische Unterschiede zwischen Experten und Novizen im Lehrerhandeln ausmachen, die sich z.B. an einer ganzheitlicheren Wahrnehmung der Experten festmachen lassen (vgl. Kounin 2006). Ähnlich hält Bromme (1992, S. 30) für Lehrpersonen fest: „Der entscheidende Unterschied zwischen Experten und Nicht-Experten liegt im Inhalt und der Struktur des Wissens“. Die inhaltliche Aussage ist im Kern nicht neu. Neu ist aber, dass hier das Lehrerwissen als Expertenwissen deklariert wird.

Weiteren Aufschluss für Lehr-Lernsituationen bietet das Konzept des „reflective practitioner“ (Schön 1983) als ein dem Expertenparadigma nahe stehender Ansatz. Zentral ist hier die Ermöglichung und Generierung von Reflexionswissen in Praxissituationen; dieses Wissen dient dabei nicht der Problemlösung, sondern der Problemdefinition und dem reflexiven Perspektivenwechsel (vgl. Bromme 1997, S. 188). Bromme (1992, S. 70) betont in diesem Zusammenhang die konstruktive Aktivität im Expertenhandeln. Schöns Ansatz wurde von Altrichter (2000) aufgenommen und weiterentwickelt. Er sieht darin vor allem die Möglichkeit einer intensiven und differenzierten Verbindung von Wissen, Denken und Handeln und stellt die Bedeutung von „reflexiven Praktika“ in Ausbildungskonzepten der Lehrerbildung heraus. Anknüpfungspunkte zum „reflective practitioner“ bietet auch der Ansatz des „forschenden Lernens“ (Wildt 1996; Bräuer 2006).

Mit Modellen der „subjektiven Theorien“ von Lehrpersonen werden die kognitiven Konstruktionsprozesse, die lehrbezogenes Expertenhandeln konstituieren oder verhindern, näher untersucht. Diesbezüglich liegen umfangreiche empirische Befunde vor (vgl. Wahl 1991, 2005) (vgl. Kap. 4.3.3).

„Wie beim Persönlichkeitsansatz wird jetzt zwar wieder nach der ‚guten’ oder ‚erfolgreichen’ Lehrperson gesucht – nur geht es jetzt nicht mehr um Charaktereigenschaften oder Führungsstile, sondern um berufsbezogenes Wissen und Können, um fachliche und fachdidaktische Expertise, um subjektive und intuitive Theorien zum Lehren und Lernen“ (Helmke 2003, S. 30).

Die Hinwendung zum Expertenbegriff kann allerdings nicht nur als Folge theoretischer Entwicklungen und empirischer Forschungsergebnisse erklärt werden, sondern ist auch Ausdruck gesellschaftlichen Wandels, der vor der Schule nicht halt macht und die Arbeitssituation von Lehrern beeinflusst. Durch ausdifferenzierte und universalisierte Erfahrungsmöglichkeiten und vor allem deren medialen Vermittlung hat die Schule als Ausbildungsinstanz tendenziell an prioritärem Gewicht verloren und ihr Alleinstellungsmerkmal eingebüßt (vgl. Bromme 1997, S. 188). Außerdem wirken die sog. neuen Medien als moderne Instruktionstechnologien auf die

Gestaltung von Unterricht ein, womit neue Anforderungen an das Lehrerhandeln gestellt werden.

4.2 Das Angebot-Nutzungs-Modell unterrichtlicher Wirkungen

Das erweiterte Prozess-Produkt-Paradigma und das Expertenparadigma schließen sich nicht gegenseitig aus, sondern können komplementär aufeinander bezogen werden. Entsprechende Konkretisierungen finden sich im Bereich der schulpädagogischen Forschung im „Angebot-Nutzungs-Modell“ schulischer Wirksamkeit (Fend 1998, 2002; Helmke & Weinert 1997; Helmke 2003, S. 104).
Dieses Modell stellt die forschungstheoretische Grundlage der eigenen Arbeit dar. Es basiert auf der Annahme, dass Angebote von Lehrpersonen in Form von Lernumgebungen seitens der Schüler individuell genutzt werden, wobei sich jeweilige Zusammenhänge zwischen lehrender und lernender Seite als strukturelle Kopplungen beschreiben lassen. Merkmale und Verhaltensweisen von Lehrpersonen innerhalb des Prozessgefüges „Unterricht“ verstehen sich so nicht mehr als isolierte Determinanten; gleichwohl unterliegt damit Lehrverhalten hinsichtlich bestehender Anforderungen, Erwartungen und Beschreibungsmöglichkeiten keiner Beliebigkeit (Helmke 2006). Vielmehr wird so eine mehrperspektivische Analyse des Unterrichts möglich, die unterschiedliche Merkmale sowie Merkmalseffekte berücksichtigt. Unterricht gibt einen vieldimensionalen Möglichkeitsraum für die Genese von Lernprozessen ab.
Konkret werden im Angebot-Nutzungs-Modell sechs Erklärungs- bzw. Beschreibungsblöcke unterschieden, die in ihrer wechselseitigen Bezogenheit unterrichtliche Wirkungen aufdecken (vgl. Abb. 4-1):

- Lehrerpersönlichkeit (Lehrperson und -expertise)
- Unterricht (bzw. Unterrichtsqualität)
- Klassenkontext und fachlicher Kontext
- Individuelle Eingangsvoraussetzungen (Mediationsprozesse)
- Individuelle Eingangsvoraussetzungen (Lernaktivitäten)
- Wirkungen (Ertrag)

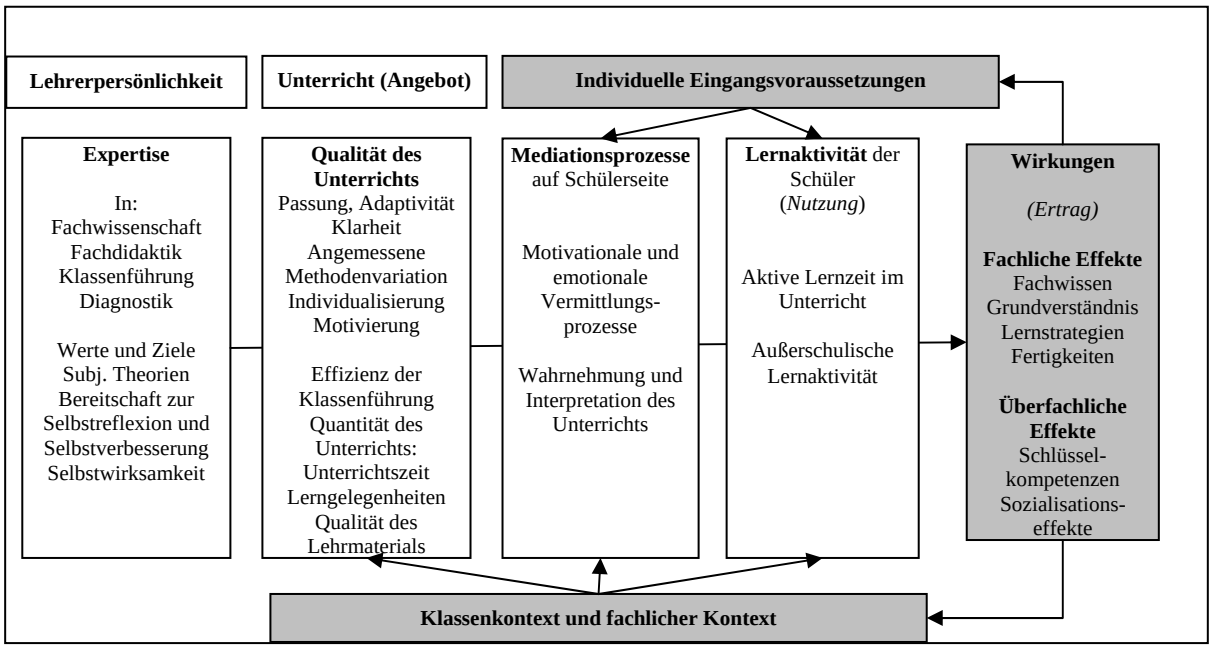

Abbildung 4-1: Angebot-Nutzungs-Modell der Wirkungsweise des Unterrichts (Helmke 2003, S. 42)

Das Modell bietet einen differenzierten Analyserahmen und ist empirisch anschlussfähig. Dabei berücksichtigt es – was ein Erklärungsmuster für seine allgemein zu beobachtende aktuelle Durchsetzungskraft darstellt – die didaktischen Implikationen, die sich aus der Konstruktivismusdebatte ergeben haben.
In diesem Kontext warnt Weinert jedoch davor, aus der konstruktivistischen Perspektive dysfunktionale Schlüsse für die Gestaltung des Unterrichts zu ziehen etwa derart, dass man die Rolle von Lehrpersonen im Unterrichtsgeschehen marginalisiert. Er unterstreicht die Tatsache, dass ein lehrergesteuerter Unterricht nicht per se in einem Widerspruch zu aktivem Lernhandeln steht.

„Ein extrem anregender, systematisch organisierter und durch den Lehrer kontrollierter Wissenserwerb ist sehr wohl mit intrinsischer Motiviertheit, konzentrierter Aufmerksamkeit und mentaler Lernaktivität vereinbar“ (Weinert 2000, S. 44ff.).

In dieser Perspektive wird die Untersuchung von didaktischen Handlungskompetenzen bzw. Einflussfaktoren ihrer Entwicklung umso notwendiger. Auch Reusser (2006, S. 157f.) verweist auf „Fehlschlüsse“, die sich aus einem naiven Verständnis des Konstruktivismus ergeben können. Er führt den sog. „epistemologischen Fehlschluss“ in das Feld, nach dem die Kernthese des Konstruktivismus, alles Erkennen beruhe auf aktiver individueller und sozialer Sinnkonstruktion, häufig dahingehend einer irrtümlichen Deutung unterliegt, dass eine intersubjektive Verständigung im pädagogischen Interaktionsprozess grundlegend in Frage gestellt wird. Diese Fehldeutung führt er auf eine forschungsmethodologisch nicht beachtete und gleichwohl existente Unterscheidung zwischen dem universellen (epistemischen) und dem individuellen (empirischen) Subjekt zurück. Danach gibt es Wissensbestände, die einen kulturell stabilen und

universellen Charakter für sich beanspruchen und im Unterschied zu jeweiligen individuellen Konstruktionen bzw. Rekonstruktionen einen Wert „an sich“ darstellen. Auf der individuellen Ebene sind Wahrnehmungs- und Orientierungsleistungen sehr wohl an konkrete, auch pädagogisch inszenierte Interaktionskonstellationen gebunden. Für die Unterrichtspraxis falsch gedeutete Rückschlüsse aus der Konstruktivismusdebatte sieht Reusser zudem in sog. „choreographischen“ und „instruktionsdidaktischen“ Fehlschlüssen. Hierunter versteht er die unmittelbare Koppelung von anthropologischen und erkenntnistheoretischen Ebenen des Konstruktivismustheorems in ihrer Übertragung auf konkrete Unterrichtsformen, die demgemäß präferiert werden müssten (Choreographie) bzw. die die Rolle der Lehrperson in Lehr-Lernprozessen (Instruktion) in bestimmter Weise fassen. Die von Reusser vorgenommenen Relationierungen des Konstruktivismusbegriffes schränken die Reichweite der Schlussfolgerungen, die aus der Konstruktivismusdebatte für eine Erneuerung der didaktischen Kultur gezogen werden können, einerseits ein, andererseits erhöhen sie die Anschlussfähigkeit an eine bereits bestehende pädagogische Praxis in einem evolutionären Sinne.

Eine Variation des Angebot-Nutzungs-Modells liegt von Baumert und Kunter (2006) vor. Sie verschärfen die im Ausgangsmodell angelegte Kontingenz von Lehrerhandeln und Mediationsprozessen der Schüler, indem sie Lerngelegenheiten immer als das Ergebnis von sozialen Ko-Konstruktionen beschreiben, an denen jeweils sowohl die Lehrperson als auch Schüler beteiligt sind. Damit wird eine zusätzliche Unsicherheit im Lehrerhandeln markiert. Die Autoren entwickeln auf Basis dieser Annahme ein „Opportunitäts-Nutzungsmodell mit doppelter Kontingenz“ (Baumert und Kunter 2006, S. 477). Das im Angebot-Nutzungs-Modell angelegte Kontingenztheorem wird damit stärker herausgestellt (vgl. Kap. 6.3).

4.3 Der Kompetenzbegriff in der Lehrerbildung

Die 1999 von der Kultusministerkonferenz eingesetzte Kommission „Lehrerbildung“ versteht den Kompetenzbegriff in ihrem Abschlussbericht folgendermaßen:

„Mit dem Begriff ‚Kompetenz’ bezeichnet die Kommission das Verfügen über Wissensbestände, Handlungsroutinen und Reflexionsformen, die aus der Sicht einschlägiger Professionen und wissenschaftlicher Disziplinen zweck- und situationsangemessenes Handeln gestatten“ (Terhart 2000, S. 54).

Die Grundlagen für den Erwerb oder Aufbau von Kompetenzen sieht die Kommission wie folgt: „Sie (die Lehrerkompetenz, O.K.) beruht auf wissenschaftlich fundiertem Wissen, auf situativ flexibel anwendbaren Routinen und auf einem besonderen Berufsethos, das handlungsleitende Wertmaßstäbe repräsentiert“ (ebd., S. 55).
Die von der Kommission vorgeschlagene Kompetenzdefinition wirkt einerseits schlüssig, andererseits verkürzend im Sinne eines kleinsten gemeinsamen Nenners, der wenig Analysepotential für die Genese und Modellierung von Kompetenzen in der Lehrerbildung bietet. Insgesamt hat sie appellativen Charakter, indem Lehrerkompetenzen in – aus Sicht der Kommission – wünschenswerten

Dimensionen beschrieben werden, die aber als Ergebnis des gewählten Duktus, kaum untersuchbar erscheinen. Diese eher appellativ erscheinende Setzung wünschenswerter Kompetenzen ist vermutlich auch ein Ausdruck langjähriger Diskursprozesse über die Ziele der Lehrerbildung, in denen immer auch eine gewisse Skepsis darüber artikuliert wurde, ob valide Lehrerbeurteilungen möglich oder gar wünschenswert seien, während umgekehrt Prüfungen und Fremdbeurteilungen einen wichtigen Teil der Berufsrealität von Lehrkräften ausmachen (Terhart 2007).

4.3.1 Referenzsysteme zur Generierung von Wissen als Grundlage für die Entwicklung von Lehrerkompetenzen

Die Frage nach den Wissensbasen, über die professionell tätige Lehrpersonen verfügen sollten, richtet sich nach den Anforderungen, die pädagogische Handlungsfelder an Betroffene stellen. Sie wird insofern in der Literatur unter einer funktionalen Perspektive diskutiert. Folgerichtig verweist die Diskussion auf einen Zusammenhang zwischen Wissen und Handeln bzw. Können als Kern der Professionalität von Lehrpersonen (Bromme 1997, S. 194f.). Jedoch bezeichnet Oser (1997a, S. 27) Wissen und Handeln als „zwei Geschwister im Streit". So gibt es in der Erziehungswissenschaft und der Psychologie zwar einen breiten Konsens darüber, dass professionelles Handeln nur auf der Basis elaborierten Wissens erfolgen kann. Über den konkreten Zusammenhang von Wissen und Können im Hinblick auf damit verbundene pädagogische Zielsetzungen, mentale Repräsentationen und die Genese professionellen Wissens und Könnens bestehen aber Uneinigkeiten. Sie verschaffen sich in verschiedenen und nebeneinander laufenden Diskussionssträngen Ausdruck. Nachfolgend sollen sie genauer betrachtet werden.

Der erste (und ältere) Diskussionsstrang verweist auf das Feld der Allgemeinen Didaktik,[14] in deren Rahmen eine Vielzahl von Modellen, Schemata und Regeln für den „guten Unterricht" ausgearbeitet worden sind, die sich wesentlich auf Reflexions- und Begründungszusammenhänge des Unterrichts beziehen. Insofern sind sie nicht produktorientiert (was soll das nachhaltige Ergebnis von „gutem Unterricht" sein?), sondern prozessorientiert (wie soll „guter Unterricht" gestaltet werden?) angelegt (Helmke 2003, S. 28f.). Die Allgemeine Didaktik stellt noch immer die Grundlage für die Lehrerbildung an deutschen Hochschulen dar, womit generell angenommen wird, dass sie eine wichtige Funktion für Stundenvorbereitung und -durchführung hat, was an dieser Stelle auch nicht grundsätzlich abgestritten, aber mit Blick auf die zweifelhafte empirische Fundiertheit und gesellschaftsadäquate Aussagekraft relationiert werden soll. Vorausschickend ist festzustellen, dass sich die Entwicklung didaktischer Theorien aufgrund des ausdifferenzierten Diskussionsstandes kaum noch umfassend beschreiben lässt, wobei aber eine generelle Dominanz von bildungstheoretischen, lehr-lerntheoretischen sowie kommunikations- und

[14] Hervorzuheben ist hier das im Bereich der Allgemeinen Didaktik prominent gewordene Modell der „Didaktischen Analyse", welches von Klafki (1963) im Anschluss an seine Theorie der „kategorialen Bildung" entwickelt wurde.

interaktionstheoretischen Ansätzen ausgemacht werden kann (Terhart 2005a, b). Nach Staub (2006, S. 169) gibt es zugleich Anzeichen dafür, dass die „in Lehrbüchern tradierte Allgemeine Didaktik eine Phase der Stagnation erreicht hat". Bezogen auf die schulische Ausbildungsfunktion der Allgemeinen Didaktik geht es um „die Theoretisierung und operative Gestaltung von Lehren und Lernen im Kontext von Ausbildung für den pädagogischen Beruf des Lehrers" (Terhart 2005c, S. 102).

Was ein Verständnis über die Allgemeine Didaktik betrifft, steht sie für „die Berufswissenschaft von Lehrerinnen und Lehrern" (Jank & Meyer 2002, S. 10). Genauer kann sie als theoretische Modellierung für die wissenschaftliche Analyse von Unterrichtsprozessen bezogen auf die Unterrichtsplanung und damit verbunden die Festlegung auf das konkrete didaktische Handeln verstanden werden (Kolbe 2001, S. 202). Heursen (1989, S. 307) definiert Didaktik in einem abstrakteren Sinne als „die wissenschaftliche Reflexion des Lehrens und Lernens". Zentrale Reflexions- und Handlungsfelder betreffen dabei die Inhalte, Ziele, Methoden und Medien, die theoretisch aufeinander bezogen sind, grundlegend aber als Einzelelemente mit daran geknüpften und jeweils zu bewältigenden didaktischen Entscheidungsanforderungen gefasst werden.[15] Die Allgemeine Didaktik stellt damit wissenschaftstheoretische Modelle zur Verfügung, die die epistemische Struktur des Unterrichts modellieren, gleichzeitig aber auch eine normative Orientierungsfunktion für die Entwicklung professioneller didaktischer Kompetenzen darstellen und also Wirksamkeit auf das Unterrichtshandeln beanspruchen. Auf einer theoretischen Ebene mag der beanspruchte integrative Geltungsbereich dieser Theorien und Modelle überzeugen, weil damit der Zusammenhang von reflexiven wissenschaftlichen Kategorien und der Wissens- und Könnensbasis von angehenden Pädagogen hergestellt ist. Kolbe (2001, S. 202) merkt an, dass diese Verbindung, als Ergebnis einer „noch geistes-wissenschaftlichen Vorstellung einer Einheit von Theorie und Praxis", gewissermaßen „unter der Hand" zustande gekommen ist. Es ist nun Aufgabe der empirischen Bildungsforschung, die damit verbundenen Probleme präzisierend zu analysieren (s.u.).
Die angeführten komplexen Didaktikentwürfe lassen ihre nur schwer zu vollziehende praxisbezogene Umsetzung erahnen. Meyer (selbst einer der führenden Vertreter der Allgemeinen Didaktik) konnte bereits 1980 Rezeptionsprobleme der Allgemeinen Didaktik bei Referendaren aufzeigen. Die von ihnen erstellten Stundenvorbereitungen erwiesen sich als wenig handlungsanleitend und die dort angegebenen Planungselemente erschienen Meyer als „Scheinbegründungen". Aktuell weist Oelkers (2004b, S. 1) darauf hin, dass man insgesamt wenig über das reale Wissen und Können von praktizierenden Lehrkräften weiß und die diesbezügliche Diskussion vor allem um die Frage der adäquaten Ausbildung kreist. Hier soll die These aufgestellt werden, dass die Prozessorientierung der Allgemeinen Didaktik in der Vergangenheit den Blick dafür versperrt hat, in welchem Zusammenhang sie mit dem realen Wissen und Können von Lehrpersonen in Interaktionsprozessen mit Schülern steht, womit sich die insgesamt geringe empirische Datenbasis bezüglich allgemeindidaktischer Fragestellungen erklären lässt. So kommt die professionssoziologisch fundierte

[15] Die vier Handlungsfelder basieren auf dem Modell der Berliner Didaktik (Heimann, Otto & Schulz 1965).

Wissensverwendungsforschung mit Blick auf das Lehrerhandeln zu dem Schluss, dass es eine dauerhafte Differenz zwischen wissenschaftlichem Wissen und beruflichem Handlungswissen gibt (Dewe & Radtke 1991). Die Transformation von Wissenschaftswissen in Praxissituationen sei eine Illusion, weil Wissenschaft und Praxis zwei getrennte Bereiche darstellen, die in Anlehnung an die Systemtheorie von Luhmann (1990) als autopoietische Systeme lediglich strukturelle Koppelungen aufweisen. Demzufolge wird das Theorie-Praxis-Problem nicht als Schwierigkeit der Hochschule definiert; denn diese würde in der Behauptung praxisrelevante Inhalte zu vermitteln, falsche Erwartungen wecken (Radtke 1996; Neuweg 2000). Dewe (1997, S. 242) geht sogar so weit zu konstatieren, das Ziel des Studiums sei allein der „reflexive Umgang mit Theorie". Der argumentativen Linie einer Theorie-Praxis-Kluft folgend stellt Oelkers (2004b, S. 1) fest:

„Ausbildung, nochmals, ist immer Idealisierung, sie zeigt nur an bestimmten Stellen what works und auch an diesen Stellen wird nicht genau antizipiert, was nach der Ausbildung den Alltag ausmacht. Daher ist zu vermuten, dass der Aufbau professioneller Kompetenz von Lehrkräften in den Berufssituationen geschieht und nur begrenzt von der Ausbildung beeinflusst werden kann. Zudem wirkt die Ausbildung immer selektiv, also nie als Ganze, sondern immer nur in relevanten Teilen."

In der strukturell angelegten Trennung von Theorie und Praxis kann ein Grund dafür ausgemacht werden, warum sich die Erziehungswissenschaft lange Zeit wenig für die unmittelbaren Handlungsanforderungen an ihre Absolventen interessiert hat. Kritisch sei hier angemerkt, dass die professionssoziologische Konzeptualisierung einer dauerhaften Trennung von Theorie und Praxis für das Hochschulfeld Lehrerausbildung mit einem klaren Berufsbild, welches wenige Alternativentwürfe bereithält, insgesamt unbefriedigend erscheint. Dies gilt besonders dann, wenn die genannten Theoriepositionen Ergebnisse einer Überprüfung von wissenschaftlichen Bildungskonzepten zu wenig in den Blick nehmen und damit nicht zur Verbesserung der Praxisfelder beitragen, für die sie aber ausbilden wollen.

Damit ist der zweite Diskussionsstrang, der sich mit der Untersuchung professionellen Wissens beschäftigt, bereits implizit angedeutet, nämlich die empirische Bildungsforschung. Tenorth (1997, S. 971) konstatiert, dass es einen „Dualismus" von Bildungstheorie und Bildungsforschung gibt, die er als zwei eigenständige Diskurse bezeichnet. Die Eigenständigkeit bezieht sich auf nicht nur heterogene, sondern zudem heteronome theoretische Zugänge. Kritisch äußert sich v. Prondczynsky (i.V.) derart, dass sich zum einen bildungstheoretisches und -kritisches Denken über weite Strecken ohne vorhandene Referenzoptionen auf Bildungsforschung entfalten konnte und sich zum anderen Bildungsforschung seit den 60er Jahren des 20. Jahrhunderts ohne systematischen Bezug auf bildungstheoretische Optionen legitimierte.
Versöhnlicher charakterisiert Terhart (2005) das Verhältnis zwischen der Allgemeinen Didaktik und der empirischen Bildungsforschung mit der Metapher „fremder Schwestern". Er will mit dieser Metapher zum Ausdruck bringen, dass sich sowohl Ähnlichkeiten als auch Unterschiede zwischen den beiden Theorie- und Forschungsbereichen ausmachen lassen. Ähnlichkeiten beziehen sich auf zu untersuchende Felder und z.T. auch auf resultierende Ergebnisse bzw. – wie noch

zu zeigen sein wird – auf ergebnisbezogene Schlussfolgerungen. Unterschiede zwischen den Diskussionssträngen verweisen auf jeweils statuierte theoretische Referenzsysteme und grundlegende methodische Fragen.
In seinem Überblick über die erziehungswissenschaftliche Bildungsforschung unterscheidet Zedler (2002, S. 21ff.) voraussetzend wiederum zwischen „klassischen Fragestellungen handlungstheoretischer Pädagogik“ und der „erziehungswissenschaftlichen Bildungsforschung“. Die zentrale Fragestellung der handlungstheoretischen Pädagogik lautet: „was kann und was soll auf welchem Wege durch erzieherisches und unterrichtliches Handeln bewirkt werden und was will die ältere mit der jüngeren Generation“, während die erziehungswissenschaftliche Bildungsforschung danach fragt, was „die ältere mit der jüngeren Generation tatsächlich“ macht (ebd.). In der ersten Fragerichtung drückt sich ein so in der zweiten Frage nicht unmittelbar zu findendes normatives Moment aus. Jedoch ist darauf hinzuweisen, dass auch die empirische Bildungsforschung nicht auf normative Setzungen verzichten kann, sei es bezogen auf die operationale Definition eines bestimmten Subjekt- oder Weltausschnittes oder auf die Interpretation qualitativer und quantitativer Daten. Zedler führt weiterhin drei auszumachende Grundsatzpositionen der Bildungsforschung an, die sich nach einem Erkenntnisinteresse

- über ein Korpus an Subdisziplinen und Personen,
- über einen Gegenstandsbereich oder
- über ein Methodenverständnis

differenzieren lassen.

Es sind heute Tendenzen eines integrativen erziehungswissenschaftlichen Selbstverständnisses vor dem Hintergrund einerseits der traditionsbasierten bildungstheoretischen Erkenntnisse und andererseits der neueren Bildungsforschung zu beobachten. Tippelt (2002, S. 11) als Herausgeber des einflussreichen „Handbuch Bildungsforschung“ stellt jedoch vorsichtig fest, dass bisher „die Integration und Annäherung von Bildungsforschung und Allgemeiner Pädagogik nur partiell gelungen“ ist. Benner (2002, S. 87) diskutiert das Verhältnis von Bildungstheorie und Empirie im Bezug auf die Ergebnisse der Pisa-Studie und kommt zu dem Ergebnis, dass „sich Bildungstheorie und Empirie nicht in ein Kontinuum aufheben“ lassen.
Offensiver negiert Wigger (2004, S. 490) die Dualismus-These von Tenorth im Verweis auf die bildungstheoretisch orientierte Biographieforschung. Zudem hält er es für denk- und wünschenswert, dass sich weitere Synergien ergeben und sieht damit die Möglichkeit „einer ambitionierten Bildungstheorie, die sich durch empirische Forschungen ergänzt, und einer methodisch ausgefeilten Bildungsforschung, die ihre implizite Theorie entfaltet, um die Akzeptanz, Bewertung und Realisierbarkeit der jeweils thematisierten Ziele und Konzepte von Bildung“ zu erzielen (ebd.). Helmke (2002, S. 265) stellt fest, dass „… Abstraktionsniveau und Duktus didaktischer Texte oft auf einem Niveau (angesiedelt sind, O.K.), das den Bezug zu empirischen Studien nicht unbedingt nahe legt“. Gleichwohl hebt er, verweisend auf die einschlägigen angloamerikanischen Handbücher und Enzyklopädien zum Thema

„Teaching“, die prinzipielle Wichtigkeit einer bildungstheoretischen Fundierung hervor.[16] Bohl (2004) betrachtet die Empirische Unterrichtsforschung und die Allgemeine Didaktik „nicht als derart wesensfremd, wie es aufgrund ihres grundsätzlich unterschiedlichen Zugangs, hier empirisch – dort normativ, erscheinen mag“. Des Weiteren sieht er Annäherungen beim Lernbegriff und Gemeinsamkeiten bezüglich des Unterrichts als Gegenstand empirischer oder theoretischer Bemühungen sowie der prinzipiellen Notwendigkeit einer Komplexitätsreduktion.
Geschlossene Modelle und Theorien der Allgemeinen Didaktik werden im weiteren Verlauf der Studie nicht mehr explizit diskutiert,[17] weil ihre Anschlussfähigkeit für die Entwicklung empirischer Verfahren als problematisch einzuschätzen ist. Nachfolgend sollen in diesem Begründungszusammenhang vielmehr Topologien und Typologien professionellen Lehrerwissens vorgestellt werden, die eine angemessene Grundlage für empirische Untersuchungen darstellen können bzw. als solche bereits Verwendung fanden.

4.3.2 Topologien und Typologien professionellen Lehrerwissens

Die bekannteste Topologie professionellen Lehrerwissens stammt von Shulman (1986). Sie erfährt nach wie vor Akzeptanz und hat sich auf breiter Linie durchgesetzt. Shulman unterscheidet zunächst folgende topologische Größen:

- General pedagogical knowledge (allgemeines pädagogisches Wissen)
- Subjekt-matter content knowledge (Fachwissen)
- Pedagogical content knowledge (fachdidaktisches Wissen)
- Curriculum knowledge (Wissen über das Fachcurriculum)

In einer späteren Veröffentlichung hat der Autor noch zwei weitere Wissensbereiche hinzugefügt (vgl. Shulman 1987):

- Knowledge of learners (lernpsychologisches Wissen)
- Knowledge of educational context (Organisationswissen)

[16] Differenzen eines Verständnisses der Allgemeinen Didaktik, in ihrer spezifisch deutschen Lesart und den Annäherungen an Lehr-Lernprozesse im angloamerikanischen Raum, lassen sich neben historisch geprägten Unterschieden auch durch verschiedene Sprachtraditionen erklären. Wie soll man etwa die Formulierung „Powerful Teacher Education“ (Darling-Hammond 2006) in die deutsche Sprache übersetzen? „Machtvolle Lehrerbildung“ würde etwa von der KMK oder anderen definitionsmächtigen Institutionen kaum Verwendung finden. Nahe liegender wären Übersetzungen im Sinne einer „wirksamen“ oder „nachhaltigen Lehrerbildung“, die aber nicht die gleiche Präsenz entfalten, wie das von Darling-Hammond verwendete Wort „Powerful“.

[17] Grundsätzlich halte ich es für wünschenswert Theorien und Modelle der Allgemeinen Didaktik – in einer pragmatischen Annäherung – auf die empirische Unterrichtsforschung zu beziehen. In dieser Studie beschäftige ich mich aber methodisch ausschließlich mit Beobachtungsverfahren. Die Allgemeine Didaktik als Referenzsystem würde es bedingen, zusätzlich z.B. mit Hilfe von Interviewverfahren die subjektiven Theorien der Studierenden zu rekonstruieren, um Denkprozesse im Planungshandeln sichtbar zu machen und nicht nur das Unterrichtshandeln aufzuzeigen. Im Rahmen des Gesamtprojektes wurde versucht dieser Problematik mit Vignettentests gerecht zu werden (vgl. Kap. 1).

Damit liefert Shulman nicht nur eine Topologie, sondern eine Typologie professionellen Lehrerwissens. Sein Ansatz wurde von Bromme (1992, 1997) überarbeitet und für den deutschsprachigen Raum erschlossen. Bromme unterscheidet die Wissensbereiche:

- Fachliches Wissen
- Curriculares Wissen
- Philosophie des Schulfaches
- Pädagogisches Wissen
- Fachspezifisch-pädagogisches Wissen

Fachliches Wissen beschreibt Bromme als wissenschaftsdisziplinäres Wissen. Er grenzt es vom curricularen Wissen ab, weil sich Unterrichtsinhalte nicht einfach als wissenschaftspropädeutisches Wissen darstellen. Die Unterrichtsinhalte sind z.B. nicht „einfach“ Mathematik, sondern sie bilden einen eigenen Kanon mit eigener Dignität. Zwischen dem fachlichen Wissen und dem curricularen Wissen stellt Bromme Spannungsverhältnisse heraus, die sich etwa daraus ergeben, dass in der Schule die Umsetzung wissenschaftlichen Wissens nicht in der disziplinär verlangten „Strenge“ möglich ist. Auffassungen über Nützlichkeit und Anwendungsfelder des Faches bezeichnet er als „Philosophie des Schulfaches“. Sie ist auch impliziter Unterrichtsinhalt und geht als bewertende Perspektive in den Unterricht ein. Der Autor räumt allerdings ein, dass hier eine strikte Trennung zu den anderen Bereichen nicht immer möglich ist. Pädagogisches Wissen beschreibt Bromme als relativ unabhängig von den Unterrichtsfächern. Als Beispiele führt er Bereiche des Klassenmanagements, der Diagnostik oder pädagogischer Erklärungsmuster hinsichtlich der Bedeutung des Elternhauses für schülerbezogenes Unterrichtsverhalten an. Diesbezüglich hebt er besonders das pädagogische Ethos von Lehrern hervor, das seiner Meinung nach bisher zu wenig empirische Beachtung erfahren hat. Didaktische Zugänge für die Darstellung und Vermittlung von Inhalten lassen sich jedoch nicht allein aus sachlogischen Strukturen des Unterrichtsstoffes deduzieren. Benötigt wird ein integratives Wissen, das psychologisch-pädagogische Kenntnisse sowie eigene Erfahrungen der Lehrperson auf das curriculare Wissen bezieht. Einen solchen Wissensbereich subsumiert Bromme unter die Formel fachspezifisch-pädagogisches Wissen. Insgesamt beschränkt sich der Autor so auffälliger Weise nicht auf ein fachdidaktisches Wissen. Mit Bezug auf den fachspezifisch-pädagogischen Wissensbereich führt er an, dass die Fachdidaktiken zwar strukturelle Muster für ein erforderliches integratives Wissen an die Hand geben, die letztendliche Explifizierung aber durch den Lehrer selbst und vor dem Hintergrund seiner Erfahrungen geschieht.

Die vorgestellte Topologie mit ihren Adaptionen verdeutlicht, dass die einzelnen Wissensbereiche sich aus verschiedenen Quellen speisen und stark aufeinander bezogen sind. In dieser komplexen, lebens- und berufsbiographischen Verwobenheit sieht Bromme das Besondere am professionellen Wissen von Lehrern, das sich von dem rein kodifizierten Wissen der Fachdisziplinen abhebt. So betont er nachdrücklich den Stellenwert der Fachdisziplin innerhalb seiner Topologie, die durch

pädagogische, psychologische und philosophische wissenstypologische Verfügungsmuster allerdings erst für den Unterricht aufgeschlossen wird.

„Professionelles Wissen bezeichnet die einmal bewusst gelernten Fakten, Theorien und Regeln, sowie die Erfahrungen und Einstellungen des Lehrers. Der Begriff umfasst also auch Wertvorstellungen, nicht nur deskriptives und erklärendes Wissen“ (Bromme 1992, S. 10).

Grundlegend für die Beschreibung von professionellen Wissenstypen des Lehrerberufes ist die Unterscheidung in:

- Allgemein pädagogisches Wissens
- Fachwissen
- Fachdidaktisches Wissen

Diese Unterscheidung wird in aktuellen Veröffentlichungen aufgegriffen (vgl. z.B. Helmke 2003; Lipowsky 2006) und spiegelt auch die Ausbildungspraxis an deutschen Hochschulen. Sie wird für die eigene Studie als prinzipiell plausibles Beschreibungsraster übernommen.
Baumert & Kunter (2006, S. 482) schlagen vor, die drei Wissenstypen als Kernelemente proximal für die Durchführung von Unterricht anzusehen und um die Facetten Organisationswissen (Unterricht findet im Kontext organisatorischer Prozesse des Schullebens statt) und Beratungswissen (professionelle Lehrer sind darauf angewiesen mit Laien, z.B. Eltern, zu kommunizieren) zu ergänzen. Diese Erweiterungsdimensionen finden in der eigenen Arbeit, die sich mit individuellen Entwicklungsfortschritten in proximalen Prozessen des didaktischen Unterrichthandelns beschäftigt, keine Berücksichtigung.

4.3.3 Kognitionspsychologische Erforschung des Lehrerwissens und -handelns

Die kognitionspsychologische Lehrerwissensforschung nahm ihren Anfang im angloamerikanischen Raum der 1970er und 1980er Jahre und untersucht kognitive Entwicklungsprozesse von Lehrpersonen in ihrer Bedeutung für den Unterricht. Vorausgesetzte Annahme lautet, dass sich Wissenstypen identifizieren lassen, die sowohl untersuch- als auch unterscheidbar sind und deren Verwendung oder Nichtverwendung im jeweiligen Handeln aufscheint. Wurde anfangs noch – ganz im Sinne eines frühen Stadiums des Prozess-Produkt-Paradigmas – der Versuch unternommen nomothetisches Regelwissen aufzuzeigen, das kausale Attributionen zwischen dem Lehrerhandeln und Prozessmerkmalen auf der Schülerseite erlaubt (vgl. Kap. 4.1.2), geht es aktuell darum Wissenstypen zu explifizieren, die, wenn auch nicht als unmittelbar handlungssteuernd, so doch zumindest als handlungsanleitend betrachtet und damit als Determinanten zur Überbrückung der Kluft zwischen Wissen und Handeln bzw. für die Entwicklung von Expertise angesehen werden können (Gruber & Renkl 2000; Baumert & Kunter 2006). Dabei sind verschiedene Wissenstypen, basierend auf den sie differenzierenden Funktionen, zu unterscheiden.

Gruber und Renkl (2000) nehmen z.B. eine Unterteilung in situationales Wissen, konzeptuelles Wissen, prozedurales Wissen und strategisches Wissen vor und unterscheiden jeweils zusätzlich Wissensmerkmale wie Verarbeitungstiefe, innere Struktur, Automatisierungsgrad, Modalität und Allgemeinheitsgrad. Eine solche wissenstypologische und merkmalsbezogene Aufgliederung erlaubt eine differenzierte Sichtweise auf die funktionale Struktur des zu erfassenden Wissens, stellt aber auch hohe methodische Anforderungen an die Forscher.
Anders unterscheidet Fenstermacher (1994) spezifisch für die Anforderungen des Lehrerhandelns und für diesbezügliche empirische Forschungsansätze die Typen theoretisch-formales Wissen und praktisches Wissen und Können („formal vs. practical knowledge[18]“) (vgl. Baumert & Kunter 2006, S. 483). Dem Typus des theoretisch-formalen Wissens kann der überwiegende Anteil des im wissenschaftlich-theoretischen Diskurs erzeugten Wissens zugeordnet werden. Hierunter fallen der größte Teil des fachwissenschaftlichen Wissens und große Anteile des fachdidaktischen sowie des allgemeinpädagogischen Wissens. Didaktisch sozial-kommunikatives Handeln im Unterricht ist in der Regel wiederum dem Typus des praktischen Wissens und Könnens zuzuordnen. Hier lassen sich zudem geringe Anteile des fachwissenschaftlichen Wissens sowie Anteile des fachdidaktischen und allgemeinpädagogischen Wissens verorten, so es erfahrungsbasiert, in spezifische Kontexte eingebettet und auf konkrete Problemlösungen bezogen ist. Die von Fenstermacher vorgeschlagene Unterteilung greift zum einen das Spannungsverhältnis zwischen Wissen und Handeln auf, zum anderen aber die komplexe Relation zwischen dem Geneseprozess (Ort und Herkunft) der Wissenstypen und der expertenhaften Repräsentation und der Anwendung des Wissens in spezifischen Kontexten.

Kognitive Entwicklungsprozesse können zusammenfassend als der Aufbau von mentalen Repräsentationen der Realität sowie ihrer deklarativen Strukturen und, darauf basierend, als die Entwicklung von metakognitiven prozeduralen Strukturen, die ein Operieren mit bzw. innerhalb dieser Repräsentationen erlauben, modelliert werden. Damit sind Denkprozesse beschrieben, die auf einer abstrahierenden Rekonstruktion von Erfahrung beruhen und als kognitive Prozesse des Wahrnehmens und des begrifflichen Operierens mit symbolischen Konstrukten und sprachlichen Repräsentationen charakterisiert werden können (Kolbe 2004, S. 207). Die abstrahierende Rekonstruktion der Elemente des Systems „Unterricht“ und ihrer Relationen führt zu komplexen kognitiven Schemata und damit zu einem Expertenwissen, in dem Wahrnehmung, Interpretation von Situationen und Handlung ineinander fließen.[19] Abstraktionsprozesse müssen dabei zweierlei leisten:

[18] Der englische Terminus „knowledge“ weist im Verhältnis zum deutschen Begriff „Wissen“, der in einem allgemeinen Sinne wesentlich für eine Verfügung über theoretische Inhalte steht, eine stärkere Handlungsnähe auf.

[19] Hier wird ein prototypischer Aufbau der Entwicklung von Experten-Schemata beschrieben, wohl wissend, dass dies in der Praxis nicht immer gelingt. Die Faktoren, die diese Expertiseentwicklung verhindern, werden in der Literatur aber unterschiedlich beschrieben. Zentral sind Konzepte, die sich auf die Verbindung von Theorie, Praxis, Übung und Reflexion als die Konstituenten von Expertise beziehen und in einem „zu viel“ oder „zu wenig“ der genannten Elemente den Grund für eine mangelnde Expertiseentwicklung ausmachen (Mayr 2006b).

- die Integration möglichst vieler Elemente des Systems und der damit verbundenen netzwerkartigen Relationen zwischen den Elementen, die das System erst generieren,
- eine kategorial-hierarchische Typisierung der Elemente und ihrer Relationen.

Der Abstraktionsgrad und die Komplexität des Unterrichtssystems bieten Erklärungen für das bisher noch nicht angeführte Phänomen des „impliziten Wissens“. Es bezeichnet das Wissen, das dem Akteur in der Situation des Handelns nicht bewusst ist und als unmittelbare Handlungsanleitung in der konkreten Situation auch nicht bewusst repräsentiert werden kann, weil die Person ansonsten nicht handlungsfähig wäre. Die Unbewusstheit des Wissens bezieht sich aber nur auf das situationale Handeln, besonders beim „Handeln unter Druck“ (Wahl, 1991), nicht auf die Wissensform an sich. Kolbe (2004, S. 211) stellt die These auf, dass Experten auf der Basis impliziten Wissens[20] agieren und dass dieses sich im Handeln zeigende Wissen nur partiell expliziert und rekonstruiert werden kann.
Hiermit sind Grenzen der empirischen Forschung aufgezeigt, die der Komplexität menschlicher Informationsverarbeitung geschuldet sind. Problematisch erscheint aber aus Sicht der psychologischen und erziehungswissenschaftlichen empirischen Bildungsforschung, dass die Begrifflichkeiten „wissenschaftliches Wissen“ und „implizites Wissen“ als unumstößliches Gegensatzpaar[21] gelten. Damit wird eine notwendige Grenzziehung für die Entwicklung empirischer Forschungsverfahren suggeriert, die einer Untersuchung impliziter Wissensformen in dem Zusammenhang ihrer auch wissenschaftlichen Genese und der Verwendung daraus gewonnener Ergebnisse für die weiterführende wissenschaftliche Theoriebildung abträglich ist. Eine solche Positionierung kann nicht im Interesse der Bildungsforschung sein.
Demgemäß formulierte es Perrig (1990) auch als Herausforderung für die Kognitionspsychologie, implizites Wissen untersuchbar zu machen und in der Begriffsbestimmung zwischen einer operationalen und einer theoretischen Beschreibungsebene zu unterscheiden (Perrig 1996, S. 212 f.). Auf einer operationalen Ebene lässt sich das Phänomen des impliziten Wissens untersuchen und in Experimenten gezielt induzieren (Marcel 1983; Perrig & Hofer 1989). Unklar bleibt dabei aber, welche theoretischen Schlussfolgerungen daraus gezogen werden können, ob z.B. implizites Wissen ein Charakteristikum für das Handeln von Experten darstellt oder ob sich in dem Phänomen nicht vielmehr ein grundlegendes Steuerungssystem der menschlichen Informationsverarbeitung und -repräsentation ausdrückt, das aufgrund der zugrunde gelegten Komplexität des Verhältnisses zwischen Gedächtnis, Verhalten und Bewusstsein erklärt, warum handlungssteuerndes Wissen z.T. implizit bleibt (Perrig 1996, S. 227). In dieser letztgenannten lernpsychologischen Perspektive kann man die These aufstellen, dass menschliche Lernprozesse Überschusspotentiale generieren, die sich nicht nur

[20] Perrig (2006, S. 61) kritisiert, dass zumeist der Begriff Wissen anstatt Gedächtnis Verwendung findet. Er führt aus (ebd.): „… wenn beispielsweise ‚Wissen’ die Bedeutung hat von ‚sprachlich formulierbarer Erfahrung’, so ist der Begriff ‚implizites Wissen’ eine unpassende Bezeichnung für die eigentlich gemeinte unbewusste Repräsentation, die eben in dem Moment ein ‚Nicht-Wissen’ ist und somit eine andere Erfahrungsgrundlage darstellt“.
[21] Polemisch könnte man einwenden, dass ein Verständnis vom Gegensatzpaar nur dann eine Rechtfertigung erhält, wenn man wissenschaftliches Wissen als einen Wissenstypus konzeptualisiert, der per se nicht anschlussfähig an das Akteurshandeln in der Praxis ist.

auf die bewussten Erfordernisse einer aktuellen Situation beziehen oder darauf zurückgeführt werden können und die von Reber, Walkenfeld & Hernstadt (1991) als „System unbewusster Kognitionen“ bezeichnet werden. Lernprozesse sind offenbar nicht ausschließlich an bewusste Steuerungen gebunden. Das im Expertenhandeln aufscheinende implizite Wissen wäre damit das Ergebnis einer domänenspezifischen sukzessiven Prozeduralisierung deklarativen Wissens, womit über die Entwicklung kognitiver Schemata Transferprozesse durch strukturelle Analogiebildungen eine Erklärung für das Expertenhandeln liefern (vgl. Kap. 3.7.1).
Die kognitiven Schemata erlauben die kategoriale Wahrnehmung und Interpretation verschiedener Systemelemente und sind mit Handlungsoptionen verknüpft. Hiermit lässt sich erklären, warum Experten in unübersichtlichen und komplexen Handlungszusammenhängen – als empirischer Normalfall des Unterrichtens – schnell, flüssig und situationsangemessen handeln können, ohne kognitive Operationen bewusst zu kontrollieren. Kognitive Schemata sind zwar das Ergebnis von über einzelne Situationen hinausreichenden Generalisierungsprozessen, entstehen aber aus einer praktischen Verwendungsperspektive des situativen Problemlösens. Bromme (1992) macht darauf aufmerksam, dass sich in der typisierenden Verbindung von Situations- und Aktivitätsmustern eine problem- und zielspezifische Integration widerspiegelt.[22]

Auch in neueren empirischen Studien wurde den handlungsleitenden Kognitionen verstärkte Aufmerksamkeit geschenkt. Im Rahmen der TIMMS[23]-Videostudie (Baumert et al. 1997; Seidel 2003a, b; Prenzel & Duit 2002) wurden die Unterrichtsmuster der aufgenommenen Unterrichtsstunden analysiert. Als Ergebnis dieser Analysen wurde die Erkenntnis formuliert, dass Unterricht in Deutschland vergleichsweise gleichförmig abläuft. Diese weit reichende Gleichförmigkeit kann daraus hergeleitet werden, dass bei Lehrkräften offenbar sehr ähnliche kognitive Strukturen und Repräsentationen zu finden sind.
Analog zum Begriff des „Schemas“ wird der Begriff des „Scripts“ seit der durch die internationalen Schulleistungsstudien angestoßenen Diskussion häufig genutzt, um bestimmte Formen der mentalen Repräsentation im Unterrichtsablauf zu benennen. Seidel (2003a, S. 42) beschreibt Skripts als routinisierte Handlungsabläufe und stereotype Vorstellungen, die kulturell geprägt sind und damit funktional als „Drehbücher“ für bestimmte Situationen bezeichnet werden können. Blömeke, Müller & Eichler (2005) stellen den Begriff der „Scripts“ zudem in Beziehung zu den „subjektiven Theorien“ von Lehrpersonen und ordnen beide Begriffe im Rückgriff auf Ehlich & Rehbein (1977, 1979) dem abstrakten Begriff der Handlungsmuster zu. Subjektive Theorien definieren sie als „gegenstandsbezogene Kognition“ im Gegensatz zum Begriff des „Scripts“, das sie als „situationsbezogene handlungsleitende Kognition“ bezeichnen. Damit stellen subjektive Theorien den umfassenderen mentalen Repräsentationsrahmen dar, wobei gleichwohl das

[22] Problematisch ist diese Integrationsperspektive allerdings bezüglich ihrer Störanfälligkeit. So weisen integrativ repräsentierte Muster für die Wahrnehmung von Situationen und daraus abgeleiteter Aktivitäten auf zwei Ebenen Fehlerpotentiale auf, die sich kumulativ auf das Unterrichtsgeschehen auswirken können. Zudem kann man die Hypothese aufstellen, dass die Integration Rigiditäten aufbaut, die nur schwer korrigierbar sind, weil sie einer individuell plausiblen Handlungslogik folgen.
[23] Third International Mathematics and Science Study.

Verhältnis von subjektiven Theorien[24] und Skripts nach Meinung der Autorinnen bislang weithin ungeklärt ist. Trotzdem stellen sie ein heuristisches Modell auf, nach dem Scripts sowie subjektive Theorien als handlungsleitende Kognitionen einen Einfluss auf das Handeln in konkreten Situationen ausüben und umgekehrt auch durch situationale Elemente beeinflusst werden. Subjektive Theorien entwerfen sie als komplexe Überzeugungen des Individuums zu den grundlegenden Fragen des Lehrens und Lernens, während Scripts als situationsspezifische didaktische Routinen[25] konzeptualisiert werden.
Der Begriff der „subjektiven Theorien" wurde von Groeben et al. (1988) in einem noch umfassenderen Verständnis konzipiert und von Wahl (1991, 2005) sowie Wahl, Weinert & Huber (2001) auf das didaktische Handeln von Lehrpersonen im Unterricht angewandt.

„Subjektive Theorien[26] stellen … die komplexeste Form der für Handlungen zentralen Merkmale von Intentionalität, über Reflexivität, sprachliche Kommunikationsfähigkeit bis hin zur potentiellen Rationalität dar" (Wahl 1991, S. 54).

In dieser Forschungsperspektive ist die Identifikation von „Handlungsauffassungen" und „Situationsauffassungen", d.h. die Zuordnung von Klassen nach Ähnlichkeitsmustern bzw. die Rekonstruktion der Verbindungen zwischen diesen Klassen, zentral.[27] Situationsauffassungen haben dabei eine orientierende Funktion für die Handlungsauffassungen. Wahl konnte aufzeigen, dass Verbindungen zwischen den Klassen häufig sehr rigide ausfallen und deshalb Veränderungspotentiale nur schwer erschlossen werden können. Subjektive Theorien avancieren damit zu den wichtigsten Moderatorgrößen, die es ermöglichen oder verhindern, dass wissenschaftliches Wissen wirksam in Handeln eingeht. Generell werden subjektiven Theorien in diesem Zusammenhang die gleichen Funktionen der Erklärung, Prognose und Technologie zugebilligt wie auch objektiven (wissenschaftlichen) Theorien.
Zwischen den Begriffen „Schema", „Skript" und „subjektiven Theorien" konnten Überschneidungen und Abweichungen aufgezeigt werden. Gemeinsam ist allen Konzepten, dass die Ausgangsannahme einer kognitionsbezogenen

[24] Wahl (2005) stellt fest, dass es subjektive Theorien verschiedener Reichweite und damit unterscheidbarer Bereichsspezifität gibt. Diese Differenzierung lässt möglicherweise einen Verzicht auf den „Script"-Begriff zu, was den Vorteil böte, innerhalb einer Begriffsdefinition operieren zu können.
[25] Wahl (1991, S. 34) bezeichnet solche Routinen als „(Teil-) Lösungen von Problemen …, die bei früheren Informationsverarbeitungsprozessen gefunden wurden und die jetzt nur abgerufen werden brauchen".
[26] Mit der Begrifflichkeit „Subjektive Theorien" wird zum einen der Gegensatz zu intersubjektiven wissenschaftlichen Theorien herausgestellt, zum anderen aber eben die Betonung auf die theoretische Dimension gelegt, wobei der Begriff „Theorie" implizieren soll, dass die zugrunde gelegten Kognitionen in einem relationalen Verhältnis zueinander stehen und damit rekonstruierbare Schlussfolgerungen erlauben (Groeben et al. 1988, S. 18). Epistemisches Wissen fällt deshalb nicht per se unter diese Kategorie, auch wenn es eine Form verdichteter Kognitionen darstellt. Routinen und Automatismen als Kognitionsstrukturen, die nicht explizit vom Individuum rekonstruiert werden können, zählen ebenfalls nicht zu subjektiven Theorien (Wahl, 1991, S. 55; Groeben et al. 1988).
[27] Die sehr aufwendige Methodik zur Rekonstruktion subjektiver Theorien soll hier nicht vertiefend dargestellt werden. Herausragendes Merkmal ist aber, dass die Rekonstruktion gemeinsam mit der untersuchten Person in einem Prozess der kommunikativen Validierung erfolgt.

Handlungsrelevanz besteht. Die Begriffe „Schema“ oder „Script“ heben auf Transferprozesse vom Wissen zum Handeln ab. Diese verlaufen nicht linear, sondern werden als komplexe mentale Repräsentationsprozesse von unterschiedlicher Reichweite modelliert. Damit verbundene Theorien und Modelle stehen für die Überzeugung, dass vom Handeln auf zugrunde liegende kognitive Prozesse rückgeschlossen werden kann. Im Gegensatz dazu verlässt das Forschungskonzept „subjektive Theorien“ die Vorstellung, dass sich handlungsanleitende Kognitionen ausschließlich durch eine externe Untersuchung des Handelns identifizieren lassen und postuliert einen äußerst eigenständigen Umgang der Praktiker mit wissenschaftlichem Wissen, woraus sich die Notwendigkeit einer aufeinander bezogenen Rekonstruktion von theoretischer und praxisorientierter Positionierung ergibt.
Insgesamt erweist sich der Schemabegriff, auf den sowohl die Skripttheorie als auch die Definition subjektiver Theorien rekurriert, in der Literatur als der umfassendste. Wie der Schemabegriff ist auch der Skriptbegriff aus der angloamerikanischen Forschung in den deutschen Diskurs eingewandert, hat theoretisch gegenüber dem Schemabegriff allerdings keinen Mehrwert. Subjektive Theorien wiederum sind in der von Groeben und Wahl entwickelten Form sehr komplex und umfassend modelliert worden. Sie legen entsprechend spezifische und aufwendige Untersuchungsmethoden nahe, um einem dahinter liegenden theoretischen Verständnis gerecht zu werden. Blömeke, Müller & Eichler (2005) zeigen jedoch auf, dass sich der Begriff auch in einer pragmatischeren Annäherung verwenden lässt, ohne seine heuristische Funktion zu verlieren. Auf Grundlage der dargelegten Abgrenzungen zum Skriptbegriff und zu den subjektiven Theorien in den Lesarten von Groeben und Wahl wird in dieser Studie der Schemabegriff präferiert.

4.3.4 Generierung von Kompetenzmodellen in der Lehrerbildung

Die oben dargestellten Klassifikationen von Kompetenzen zur Generierung von Kompetenzmodellen sind auf die Lehrerbildung übertragbar (Roth 1971; Frey 2006). Diese Übertragung bietet den Vorteil einer Vergleichsmöglichkeit mit anderen Berufsfeldern und den Anschluss an diesbezügliche interdisziplinäre Diskurse. Als Theoriefolie erscheint es daher sinnvoll, auch für die Lehrerbildung eine Orientierung an entsprechenden Modellen vorzunehmen. Jedoch ist das Berufsfeld von Lehrerinnen und Lehrern von Spezifika geprägt, die in der Klassifikation von Kompetenzen als Ausgangspunkt für die Generierung eines Modells auch kategorial und terminologisch zu berücksichtigen sind. Arnim (2000, S. 302-315) macht einen Vorschlag, wie ein solches Modell aussehen könnte und orientiert sich dabei an inhaltlichen Dimensionen des professionellen Wissens, das Lehrer benötigen. Er unterscheidet fünf Kompetenzdimensionen:

- Selbstverantwortungskompetenz
- Allgemeine pädagogische Handlungskompetenz
- Spezielle pädagogische Handlungskompetenz
- Fachliche Kompetenz
- Organisationskompetenz

Unter Selbstverantwortungskompetenz versteht Arnim Kompetenzdimensionen, die aus der Reflexion des eigenen Denkens und Handelns erwachsen und damit z.B. die Überprüfung von Projektionen und Übertragungen erlauben. Weiterhin unterscheidet er die allgemeine von der speziellen pädagogischen Handlungskompetenz, weil erstere nicht allein auf die Bewältigung von schulischen Anforderungen gerichtet ist. Er meint damit eine basale pädagogische Kompetenz, die es erlaubt ein Gespür für pädagogische Prozesse zu entwickeln und diese kompetent zu begleiten. Spezielle pädagogische Handlungskompetenz zielt auf die im engeren Sinne schulischen Anforderungen ab, wie z.B. Lernumgebungen gestalten, Methodik und den Umgang mit Medien. Fachliche Kompetenzen konstituieren sich im Umgang mit Sach- und Funktionszusammenhängen auf der Basis fachwissenschaftlicher Grundlagen, während Organisationskompetenz ein adäquates Handeln in der Institution Schule meint und Bereiche schulischer und außerschulischer Kooperation einschließt.
Arnim stellt heraus, dass die verschiedenen Kompetenzbereiche nicht unverbunden nebeneinander stehen und sich hingegen wechselseitig beeinflussen. Im Gegensatz zu Frey (2006) steht nicht die Fachkompetenz im Zentrum seines Modells. Vielmehr plädiert er dafür, den Kompetenzbereichen einen gleichberechtigten Stellenwert zuzugestehen. Das Modell bildet einerseits überzeugend die Inhaltsbereiche und Anforderungen des schulischen Handlungsfeldes ab und erweist sich als vereinbar mit den Diskursen zur Professionalisierung im Lehrerberuf. Andererseits sind die Formulierungen nicht weitgehend genug operationalisiert und damit zu vage, um als kategorialer Orientierungsrahmen der empirischen Kompetenzforschung zu dienen.

4.3.5 Kompetenzentwicklung durch Professionalisierung?

Im Vergleich zu anderen als Professionen bezeichneten Berufsbereichen der Theologie, Medizin und Jurisprudenz wurde der Lehrerberuf erst relativ spät an den deutschen Universitäten verankert; insofern ist ein Begründungsmuster dafür geliefert, dass die Lehramtsstudiengänge inneruniversitär nicht selten als Studiengänge zweiter Wahl erscheinen und auch der Lehrerberuf aus einer gesellschaftlichen Perspektive nicht die Akzeptanz erfährt, wie es bei anderen Professionen der Fall ist. Der Professionsbegriff liefert in seiner wissenschaftlichen Verwendungsweise eine Unterscheidungsfolie dafür, was einen Beruf als Profession von anderen Berufen unterscheidet. In einer ersten Annäherung und in Anlehnung an die soziologischen struktur- und systemtheoretischen Konzepte Talcott Parsons (1964, S. 160-178) kann man folgende Merkmalsbereiche als konstituierend für eine Profession anführen:

- Wissenschaftliches Wissen und eine darauf aufbauende Fachterminologie als Basis für die (freiberufliche) Berufsausübung;
- Autonome Formulierung von Standards für die Berufsausübung und Ausbildung sowie die diesbezügliche Selbstkontrolle in Form von Berufsverbänden und Interessenvertretungen;
- Niederschrift eines expliziten Normenkatalogs für das berufsethische Handeln, womit die Verbindung zu zentralen Werten der Gesellschaft deutlich wird;

- Tätigkeitsbereich, der prinzipiell für alle Mitglieder der Gesellschaft bedeutsam ist und insofern eine gemeinnützige Funktion aufweist;
- Besonderheit der Klientenorientierung als Grundlage für ein gelingendes berufliches Handeln.

Die hier aufgelisteten Merkmalsbereiche verweisen auf die historische Entwicklung des Professionsbegriffes. Eine trennscharfe Unterscheidung dessen, was eine Profession von einer nicht professionalisierten Berufstätigkeit abgrenzt, lässt sich heute sicher nicht mehr an der Frage der akademischen Ausbildung festmachen.[28] Aus einer berufssoziologischen Perspektive sei auf die Begriffe „Bildungsexpansion" und „Akademisierung der Gesellschaft" verwiesen. In ähnlicher Weise stellen die anderen genannten Bereiche nicht in einem trennscharfen Sinne Alleinstellungsmerkmale zur Beschreibung einer Profession dar, weshalb sie aber trotzdem nicht an Bedeutung verlieren. Combe & Helsper (2002, S. 30) sehen etwa in der Klientenzentrierung den zentralen Bezugspunkt für ein differenziertes Professionsverständnis, von dem aus eine sinnlogische Organisation des Begriffes gelingen könne.

Professionalisierungstheorien sind nach einer grundlagentheoretischen Debatte seit den 1980ern erneut in den letzten Jahren und in stärker handlungsfeldorientierter Perspektive sowohl in der Pädagogik als auch in der Soziologie breit diskutiert worden. Dabei wurde aus der Sicht beider Disziplinen untersucht, inwieweit das Handeln von Lehrern in der Institution Schule dem allgemeinen Typus professionellen Handelns entspricht bzw. inwieweit hier Spezifika auszumachen sind (Dewe, Ferchhoff & Radtke 1992; Bastian et al. 2000)[29]. Voraussetzung für die Formulierung einer Professionstheorie wäre dann die Rekonstruktion der Logik erzieherischen Handelns (Dewe 1998, S. 70).
Wernet (2003) legt eine empirische Studie vor, in der er die soziologische Rollentheorie Parsons und deren Weiterentwicklung durch Oevermann als Folie für die Analyse pädagogischer Interaktionen, die in Form transkribierter Unterrichtssequenzen vorlagen, nutzt. Die von der Professionalisierungstheorie geforderte Vermittlung zwischen rollenförmigem Handeln des Lehrers und seiner Fähigkeit zur Begegnung unter Einbringung der „ganzen Person" negiert Wernet. Vielmehr sollen sich Lehrer konsequent auf ein Verhalten und Handeln in der „Lehrerrolle" beschränken und sich in Situationen, in denen das Arbeitsbündnis zwischen Lehrer und Schülern zu scheitern droht, permissiven Strategien bedienen. Der Ansatz von Wernet wurde aus soziologischer Perspektive von Twardella (2004) kritisch gewürdigt, in dem er in einer erneuten Analyse der von Wernet vorgelegten Interaktionssequenzen aufzeigt, dass Lehrerverhalten und -handeln notwendigerweise nicht nur rollenkonform angelegt sein kann. Dies gilt besonders für

[28] Gleichwohl schwingt in der Professionalisierungsdiskussion eine Verbindung von „Professionalisierung" und „Verwissenschaftlichung" mit.
[29] Dewe (1998, S. 72) macht auf einen entscheidenden Unterschied der Pädagogik zur Medizin oder Justiz aufmerksam. Geht es bei den letztgenannten Professionen um eine respektive Wiederherstellung von als erstrebenswert definierten Zuständen, nämlich Gesundheit und Gerechtigkeit, liegt der Schwerpunkt im pädagogischen Handeln auf dem prospektiven Zugriff der Erzeugung von Wert- und Normorientierungen, die dem Adressaten nicht unmittelbar aufgehen oder die Verbesserung einer bestimmten Problemlage bringen.

Situationen, in denen Schüler selbst nicht rollenkonform handeln, sodass die Situationen nur noch durch „die Bereitschaft des Lehrers, persönlich für den Sinn der Aufgabe einzustehen“, „gerettet“ werden können und permissive Strategien unter Umständen dann nicht mehr ausreichen. Die Diskussion vermittelt einen Eindruck darüber, wie sich Aushandlungsprozesse zwischen eher soziologischen bzw. eher pädagogischen Annäherungen an das Professionalisierungstheorem gestalten. Der dabei gewonnene diskursive Erkenntnisfortschritt in der Analyse pädagogischer Interaktionen erscheint einerseits als Bereicherung, andererseits ist nicht immer durchsichtig, inwieweit einzelne Ergebnisse, gemessen am Abstraktionsgrad und dem Geltungsniveau, für das praktische Lehrerhandeln aufgeschlossen werden können.

Für die Lehrerprofession zeigen Bastian & Helsper (2000) auf, dass die typischen pädagogischen Tätigkeiten, nämlich „Erziehen“, „Beraten“ oder „etwas Vermitteln“, sich als Anforderungssituationen beschreiben lassen, die eine hohe Affinität zu laienhaften Praxisvollzügen aufweisen, womit auch Nichtprofessionelle sich als kompetent und zuständig für die Beantwortung von Fragen der Lehrerbildung empfinden würden. Demnach verfügen Lehrer zu wenig über eine Fachsprache und also ein symbolisches Deutungswissen, welches sich aus wissenschaftlichen Erklärungen, Theorien und empirisch gesicherten Wissensbeständen speist, um sich damit gegenüber Alltagstheorien und dem Laienwissen abgrenzen zu können.

Lehramtsstudierende messen während ihrer universitären Ausbildung dem erziehungswissenschaftlichen Wissen nur geringe Relevanz für ihre spätere Berufstätigkeit bei (ebd.). Bauer (2000) schließt, dass hierin eine Ursache dafür zu finden ist, weshalb die im Lehrerberuf agierenden „Praktiker“ dem erziehungswissenschaftlichen Wissen eine relativ niedrige Akzeptanz entgegenbringen. Der universitären Ausbildung gelingt es offensichtlich nicht ein fundiertes, konturiertes und gesichertes „professionelles Selbst“ auszubilden. Die angeführten Strukturdefizite – zum einen auf die universitäre Lehrerbildung, zum anderen auf die spezifischen Professionalisierungswege der Lehramtsstudien und Lehrerbildung im historischen Verlauf bezogen – führen dazu, dass angehende Lehrer ein mangelndes pädagogisches Selbstverständnis entwickeln, welches nur unzureichend durch ein pädagogisches Reflexionswissen gesichert ist und damit Verunsicherungen durch äußere Impulse ausgesetzt ist.

Das Lehrerhandeln unterliegt zudem grundlegenden Antinomien, welche sich aus den Widersprüchlichkeiten der Schulorganisation erklären lassen. Helsper (1996, 2000) verweist diesbezüglich auf unlösbare Paradoxien, die für das Lehrerhandeln aus der Schulpflicht und hoheitsstaatlichen Kontrollformen der Selektion erwachsen. Strukturell angelegt ist damit die Entstehung von systematisch scheiternden Bildungsbiographien, sodass die Entstehung eines professionellen Habitus besonderen Spannungskräften und Entgrenzungen unterliegt. Combe & Kolbe (2004, S. 834) sprechen deshalb von einer systematischen Unsicherheit pädagogischen Handelns, wobei die pädagogischen Handlungssituationen sich mit Begriffen wie Komplexität, Intransparenz und Riskanz beschreiben lassen. Die Begriffe sind einer Analyse pädagogischer Praxis aus strukturtheoretischer Perspektive zuträglich, konstruieren aber ein Anforderungsniveau und Problemstellungen, welche aus individueller Perspektive auf der Handlungsebene kaum noch lösbar erscheinen. Gleichzeitig stellen sie eine Legitimationsfolie dafür dar, Lehrerhandeln dem Typus

professionellen Handelns zuzuordnen. Combe & Kolbe schließen folgerichtig, dass die im Lehrerhandeln aufscheinenden Tätigkeitsbereiche vor diesem Hintergrund nicht standardisierbar sind und einen hohen Grad an reflexiven und hermeneutischen Kompetenzen verlangen. Die hier an Lehrpersonen herangetragene hohe Verantwortlichkeit erfordert es, Handeln vor dem Hintergrund wissenschaftlicher Standards zu reflektieren. Es werden folgende Voraussetzungen benannt (ebd., S. 853):

„All das setzt Lehrer voraus, die sich zweierlei zu Eigen gemacht haben: wissenschaftlich angeleitete theoretische und fallverstehende Reflexivität einerseits und praktisches Können andererseits. Unstrittig sind für den Aufbau dieser Handlungsbasis zwei Elemente erforderlich: die Bildung durch Wissenschaft und die eigenständig-konstruktive Verwendung der Wissensbestände in Abarbeitung an der Praxis."

Mit dieser Aussage markieren die Autoren einen Konsens, der auch interdisziplinär betrachtet Zustimmung beanspruchen kann. Welche Konsequenzen daraus aber für die Lehrerbildung gezogen werden können bzw. müssen, ist strittig.

Die schon oben angedeutete Gestaltung der Klientenbeziehung als Grundlage für professionelles Handeln wurde wesentlich durch Oevermann (1996) herausgearbeitet. Seine zentralen Grundtheoreme leitet er dabei, analog zu Parsons, aus der therapeutischen Praxis ab. Die Bearbeitung lebenspraktischer Krisen, mit dem Ziel der Ermöglichung von individueller Autonomie, stellt ein konstituierendes Element von Professionen dar. Bezogen auf professionell gestaltete pädagogische Interaktionen treten in dieser Perspektive Momente der Wissensvermittlung bzw. der Institutionalisierung von Lehr-Lernprozessen in den Hintergrund. Die Beziehung zwischen einem professionell Handelnden und einer näher zu spezifizierenden Klientel wird vielmehr im Modus der „stellvertretenden Deutung" angelegt.
Der „Deutungsmusteransatz" stellt Tätigkeitsbereiche einer Profession heraus, die als Rekonstruktion latenter Sinnstrukturen von Ausdrucksgestalten des Klienten verstanden werden können und insofern eine mäeutische Funktion haben. In Anlehnung an therapeutische Settings erhält der Begriff des „Arbeitsbündnisses" zentrale Bedeutung. Rekonstruktionen latenter Sinnstrukturen im Rahmen stellvertretender Deutungsprozesse verlangen nicht die Relativierung von personalen Deutungsmustern vor dem Hintergrund eines wissenschaftlich gewonnenen und reflektierten Wissens, sondern viel eher die Relationierung dieser Strukturen als Ergebnis einer methodisch geleiteten Perspektivübernahme.

„In diesen Studien (den strukturtheoretischen Analysen, O.K.) wird Professionalisierung verstanden als das Hervorbringen einer Handlungsstruktur, die es ermöglicht, in der Alltagspraxis auftretende Handlungsprobleme (von Patienten oder Klienten) aus der Distanz stellvertretend wissenschaftlich reflektiert zu bearbeiten. Das Kriterium des wissenschaftlichen Expertentums reicht dabei allein für eine Bestimmung dieses Handlungstyps nicht aus! Professionalität basiert auf einem Handlungsmechanismus, den Parsons ‚unbewusste Psychotherapie' genannt und den Oevermann zu der Figur der Dialektik von universalisierter Regelanwendung bei strikten Fallbezug verdichtet" hat (Dewe 1998, S. 70).

Der strukturtheoretische Ansatz von Oevermann ist in der Erziehungswissenschaft breit rezipiert worden und nach wie vor aktuell. Für pädagogische Settings, die durch face-to-face-Interaktionen geprägt sind, wie in großen Teilen sozialpädagogischer und zumindest nennenswerten Teilen erwachsenenpädagogischer Handlungsfelder, bietet der Ansatz von Oevermann ein Erkenntnis- und Handlungsmodell, das theoretisch überzeugen kann, allerdings in soziologischen Theorietraditionen verhaftet bleibt. Wendet man das Theoriemodell auf didaktische Handlungsfelder der Schulpädagogik an, wie es Oevermann vorschlägt, so bleiben Zweifel, ob hier ein Handeln modelliert werden kann, das sich erkenntnistheoretisch sinnvoll dem therapeutischen Grundmodell zuordnen lässt.
Der strukturtheoretische Ansatz thematisiert schulpädagogisches Handeln, neben der Wissens- und Normvermittlung, vor allem im Sinne einer prophylaktisch-therapeutischen Funktion, indem es um Interaktionen mit Menschen geht, die als Kinder und Jugendliche noch kein stabiles Selbst entwickelt haben. Professionelles Lehrerhandeln wird damit zur Quasi-Therapie umdefiniert. Die Grundlage jeder therapeutischen Interaktion ist nun jedoch ein freiwilliges Arbeitsbündnis, das sich aus der Motivation des Klienten bzw. einem Leidensdruck speist. Die Freiwilligkeit wird damit zur Konstituente eines solchen Arbeitsbündnisses, welche aber mit der gesetzlichen Schulpflicht im deutschen Schulsystem nicht gegeben ist. Hierin sehe ich ein gewichtiges Argument gegen die Etikettierung schulpädagogischen Handelns als quasi-therapeutisches Handeln. Oevermann reagiert auf diese Einschränkung, indem er die Neugierde und den Wissensdrang per se als äquivalente Funktionen zum „Leidensdruck“ beschreibt. In konsequenter Fortführung seiner strukturtheoretischen Logik fordert er denn auch die Abschaffung der Schulpflicht, um die Autonomie des Arbeitsbündnisses zu gewährleisten und eine Professionalisierung des Lehrerhandelns vorantreiben zu können.
Der von Oevermann entwickelte und von Helsper fortgeführte Ansatz einer strukturtheoretischen Analyse pädagogischer Interaktionen stellt einen theoretisch einleuchtenden Zugang dar. Jedoch wird in dieser Sichtweise die pädagogische Profession als Semiprofession beschrieben und als stark professionalisierungsbedürftig abgetan. Dabei erfährt zu wenig Betonung, dass es sich bei dem theoretischen Ansatz um eine spezifische Herangehensweise für die Analyse pädagogischer Interaktionen handelt, die eben die System- und Strukturebene in den Blick nimmt. Unter diesem Fokus bleiben pädagogische Interaktionen grundlegend defizitär, wird der Lehrerberuf zum „unmöglichen Beruf“.
Tenorth (2006, S. 580) stellt die These auf, dass die beschriebenen Codierungen für die Charakterisierung der pädagogischen Profession weitgehend unbrauchbar sind, weil sie zu wenig Sensibilität für das schulpädagogische Feld aufweisen und insofern gelingende pädagogische Praxis tendenziell abwerten. Damit werden entsprechende Theorieanstrengungen zur Belastung für eine Profession, deren Lage nach Tenorth ohnehin schon schwierig genug ist und nicht zusätzlich noch durch eingeforderte Anpassungen an perspektivenfremde Konstruktionen erschwert werden sollte. In Anlehnung an Holzkamp verwendet er den Begriff der „Letaltheorie“ und überträgt ihn auf die dargestellten Professionalisierungstheorien. „Sterblich“ bezüglich ihres Geltungsanspruches würden Theorien dann, wenn sie Alltagsphänomene nicht mehr plausibel beschreiben und befriedigend erklären können.

Darüber hinaus erscheint es wenig zielführend, eine Theorie unter dem Anspruch zu entwickeln, alle pädagogischen Praxen auf dasselbe Analyseraster zu beziehen, weil damit funktionsspezifische Unterschiede negiert werden und auch die jeweiligen Settings unberücksichtigt bleiben. Tenorth (2006, S. 585) führt aus:

„Die Einordnung in den Kontext der ‚sozialisatorischen Interaktion' und den Funktionskreis der Therapie, wie das in der Oevermann-Nachfolge geschieht, ist deshalb falsch, irreführend und eine schulisch nicht lösbare Generalisierung der ‚Erziehungs'-Erwartung gegenüber Lehrer und Schule. Diese Orientierung an Therapie und der ihr impliziten Normativität erzeugt auch erst die Schwierigkeiten, die anschließend in komplizierten Überlegungen im Zeichen von ‚Professionstheorie' bearbeitet werden müssen."

Aktuell bemühen sich Baumert & Kunter (2006) um einen Ausgleich zwischen einer strukturtheoretischen Deutung des Lehrerhandelns und Ansätzen, die – internationalen Diskussionen folgend – eine Professionalisierung der Lehrerbildung auf der Basis von kompetenz- und standardorientierten Modellen intendieren. Sie stellen fest, dass die Entwicklungen in beiden Diskussionssträngen in divergenten, sich praktisch nicht berührenden Bahnen verlaufen. In ihrer Kritik an der strukturtheoretischen Deutung weisen sie darauf hin, dass die Sozialbeziehung zwischen Lehrkraft und Schülern durch die staatliche Institution Schule wesentlich vorstrukturiert ist. Im Mittelpunkt steht immer der Unterricht, der so zum Kerngeschäft des Lehrerhandelns wird. Die daraus resultierende Rollenbeziehung ist sachlich und universalistisch ausgeprägt und genau deshalb nicht psychotherapeutischer oder psychoanalytischer Natur. Aus der Zuschreibung, dass Lehrerhandeln als komplexe Tätigkeit eines herausgehobenen Berufes im Kern nicht standardisierbar und von Erfolgsunsicherheit gezeichnet ist, leiten Baumert & Kunter allerdings kein generelles Technologiedefizit ab. Hingegen zielt ihr Interesse darauf, im Sinne einer „Neujustierung der theoretischen Perspektive" Rückschlüsse auf den Typus und die Struktur des pädagogischen Handlungswissens zu ziehen.

4.4 Fazit

Im Kapitel wurde dargestellt, wie sich normative und empirische Herangehensweisen zur Beschreibung des Lehrerhandelns im historischen Verlauf verändert haben. Im Zuge dieser Veränderung können Paradigmenwechsel ausgemacht werden, die sich wesentlich darauf ausgewirkt haben, welche Bereiche des Lehrerhandelns bzw. der Lehrerbildung hermeneutisch oder empirisch untersucht wurden. Gleichzeitig lässt sich eine Diskussion der Leitbegriffe aufzeigen. Dabei wird der Bildungsbegriff tendenziell durch den Kompetenzbegriff abgelöst, woraus sich neue Schwerpunktsetzungen für das Lehrerhandeln und die Lehrerbildung ergeben (vgl. Kap. 2 / 3).
Innerhalb der Erziehungswissenschaft lassen sich mit der Allgemeinen Didaktik und der empirischen Lehr-Lernforschung zwei Theorieströnge ausmachen, die z.T. von sehr stark divergierenden Grundsatzpositionen ausgehen und jeweils die Deutungshoheit für die Lehrerbildung beanspruchen. Während die Allgemeine Didaktik von soziologischen Theoriepositionen beeinflusst ist und deshalb die Frage

von Begründungszusammenhängen des didaktischen Handelns fokussiert und in diesem Kontext auch den Professionsbegriff thematisiert, sind Fragestellungen und Methoden der empirischen Lehr-Lernforschung über die Psychologie eingewandert. Mit dieser unterschiedlichen Bezugspunktorientierung lassen sich die Divergenzen der Theoriestränge erklären, die sich z.B. auch in einer stärker gewichteten Prozess- oder Produktorientierung ausmachen lassen und Fragen der Methodik (hermeneutisch vs. empirisch; qualitativ vs. quantitativ) betreffen.
Beide Stränge wurden hier breit diskutiert, weil ich mit Baumert & Kunter (2006) die Auffassung vertrete, dass erst in einem angemessenen Ausgleich der Positionen eine „Neujustierung der theoretischen Perspektive“ der Lehrerbildung gelingen kann. Wenn dies auch nicht Aufgabe der vorliegenden Arbeit ist, so möchte ich hervorheben, dass in der Instrumentenentwicklung (vgl. Kap. 10) bzw. in der Interpretation der Ergebnisse (vgl. Kap. 11) versucht wurde dieser Gratwanderung gerecht zu werden.
In beiden Theoriesträngen lassen sich Maximalpositionen erkennen, die auf problematische Entwicklungslinien hindeuten. So darf man mit einiger Berechtigung fragen, inwieweit die in der akademischen Ausbildung für den erziehungswissenschaftlichen Teil dominierende Allgemeine Didaktik in der Praxis des Lehrerhandelns tatsächlich den Status einer Berufswissenschaft beanspruchen kann oder auch überhaupt beanspruchen will, wenn sie nur wenig bereit ist, sich die Ergebnisse ihrer Bemühungen genauer anzuschauen und evtl. aus dieser Anschauung auch notwendige Revisionen abzuleiten. Umgekehrt ist jede empirische Studie, die sich mit dem Feld Lehrerbildung auseinandersetzt und das konkrete Handeln z.B. von angehenden Lehrern untersuchen will, zwingend darauf angewiesen, sich mit Eckpfeilern erziehungswissenschaftlicher Diskurse zu befassen, um adäquate Konstrukte zu entwickeln und die Ergebnisse vor diesem Theoriehintergrund einordnen zu können. Andernfalls besteht die Gefahr, dass Annahmen und Konstrukte an das Feld Lehrerbildung angelegt werden, die unter Umständen einseitig z.B. politische oder ökonomische Interessen bedienen und nicht das Ziel einer Verbesserung professioneller Kompetenzen anstreben.
Ein Beispiel für solche Gefährdungslagen und Potentiale gibt die Diskussion um die Einführung von Standards innerhalb der Lehrerbildung ab, die im nächsten Kapitel diskutiert wird.

5 Standardorientierte Kompetenzmodelle der Lehrerbildung

Die Diskussion um Kompetenzmodelle in der Lehrerbildung und der unterrichtlichen Praxis zeigt sich prinzipiell anschlussfähig an den Diskurs über allgemeine Kompetenz- und Expertiseentwicklung (vgl. Kap. 2.3). Die Modelle von Dreyfuss & Dreyfuss (1986) und Frey (2006) sind auf die Spezifika didaktischer Handlungsfelder angewandt worden und können dort auch Erklärungsmuster für die Entwicklung von Kompetenzen liefern (Koch-Priewe 2002; Frey 2004, 2006). Sie teilen sich allerdings den Nachteil, dass sie auf der Basis anderer Berufe entwickelt wurden und so den Professionscharakter des Lehrerhandelns zu wenig abbilden bzw. erst im Nachhinein dieser Zielvorstellung gerecht zu werden versuchen. Die von der KMK (2004)[30] und Girmes (2006a, b) vorgestellten Kompetenzmodelle erscheinen in begrifflicher und kategorialer Sicht als zu wenig anschlussfähig, um eine orientierungsleitende Funktion für empirische Forschungsvorhaben leisten zu können.
Für das deutschsprachige Feld lassen sich zwei Kompetenzmodelle explifizieren, die genuin für das Feld professioneller Handlungskompetenzen im Lehrerberuf entwickelt worden sind und der empirischen Überprüfung dienen können, nämlich zum einen das Standardmodell von Oser (1997a, 1997b; Oser & Oelkers 2001) und zum anderen das Kompetenzmodell von Terhart (2000, 2002, 2006). Beide Modelle speisen sich aus unterschiedlichen metatheoretischen Annahmen und verwenden verschiedene Konstrukte und Termini. Dieser Umstand ist nicht zuletzt der Ursache geschuldet, dass das eine Modell in der Schweiz (Oser) und das andere in Deutschland (Terhart) seine Wurzeln hat. Die folgenden Ausführungen setzen jedoch noch früher an, so bei der Entwicklung von Standardkonzepten in den USA, weil sie als Ausgangspunkt für die Diskussion von Standardmodellen im deutschsprachigen Raum angesehen werden kann. Sowohl die Theoriediskussion als auch die praktische Implementation von Standards in die Lehrerbildung sind dort weiter vorangeschritten.

5.1 Standardmodelle in den USA

Die Entwicklung und Implementation von Standards hat in den Vereinigten Staaten von Amerika bereits eine längere Tradition. Ein grundlegender Vorschlag zur Implementation von Standards in die Curricula der Lehrerbildung stammt von Ravitch (1995). Metatheoretisch unterscheidet die Autorin drei Merkmale von Standards, die auch funktional auf unterschiedliche Dimensionen verweisen, und zwar:

[30] Blömeke (2006b, S. 29) hält zu Recht fest, dass durch Standardformulierungen eine Orientierung auf kumulatives und nachhaltiges Lernen unterstützt werden sollte – ein Anspruch, der sich allerdings in der Ausformulierung der KMK-Standards nicht wieder findet. Vielmehr wird ein sehr umfassender Katalog von Kennens-Zielen für die eher theoretische Ausbildungsphase an der Hochschule gesetzt, der vom quantitativ vergleichsweise geringen Anteil der Erziehungswissenschaft innerhalb der Lehrerbildung kaum abgedeckt werden kann. Überraschend ist in diesem Zusammenhang auch die Beschränkung von kognitiven Zielsetzungen auf ein bloßes „Kennen“. Damit werden die Hochschulen gleichzeitig quantitativ überfordert und qualitativ unterfordert. Kritisch weist Blömeke darauf hin, dass die Expertengruppe zur Entwicklung der Standards nicht nur aus Erziehungswissenschaftlern bestand, womit sich die fehlende Berücksichtigung von Fachbezügen erklären ließe. Die Entwicklung von Standards für eine fächerübergreifende Didaktik kann ihrer Auffassung nach jedoch nur über allgemeindidaktische und damit erziehungswissenschaftliche Standarddefinitionen erfolgen.

- content standards
- performance standards
- opportunity-to-learn standards

„Content standards" beziehen sich auf die definitive Festlegung der Inhalte in den Curricula und sollen Transparenz auf der Lehrer- sowie Lernerseite schaffen. Vorauszusetzen ist dafür nach Ravitch, dass die Inhalte zweifelsfrei und verbindlich festgelegt werden und eine zielperspektivische Begründung erfahren.

„Content standards (or curriculum standards) describe what teachers are supposed to teach and students are expected to learn. They provide clear, specific descriptions of the skills and knowledge that should be taught to students" (Ravitch 1995, S. 12).

Dem gegenüber sollen „performance standards" die Unterscheidung von Leistungsniveaus ermöglichen, die als Ergebnis der Beschäftigung mit curricularen Lerninhalten angestrebt werden. Sie stellen eine Grundlage der Notengebung dar.

„Performance standards define degrees of mastery or level of attainment. They answer the question: `How good is good enough?´ Performance standards describe what kind of performance represents inadequate, acceptable, or outstanding accomplishment" (Ravitch 1995, S. 13).

Schließlich lenkt Ravitch den Blick auf einen Bereich, der in erziehungswissenschaftlichen Diskursen teilweise wenig Beachtung findet, weil er im staatlichen Schulsystem wesentlich einer politischen Steuerung unterliegt, nämlich die Frage der im unterschiedlichen Umfang zur Verfügung stehen Ressourcen. Sie beziehen sich nicht nur auf materielle, sondern als „oportunity-to-learn standards" auch auf ideelle Dimensionen und sollen damit unterstützend auf die anderen beiden Standarddimensionen einwirken.

„Opportunity-to-learn standards define the availability of programs, staff and other resources that schools, districts, and states provide so that students are able to meet challenging content and performance standards" (Ravitch 1995, S. 13).

Das Ravitch-Modell bezieht Standards für die Definition der Inhalte, die Erreichung von Leistungsniveaus und die dafür erforderlichen Ressourcen aufeinander. In der schematischen Darstellungslogik kann das Modell eine hohe Plausibilität beanspruchen. Über die Betonung der Ressourcen berücksichtigt es eine ökonomische Perspektive. Obwohl es genuin für die Veränderung der Lehrerbildung angelegt wurde, führt die Autorin hier eine Begrifflichkeit ein, die einen ausdrücklichen Kontrapunkt zu pädagogischen Diskursen bilden soll.[31]

Neben dieser eher grundsätzlichen Annäherung an die Implementation von Standards besteht eine anders nuancierte Standarddiskussion, die darum kreist, wie

[31] Für diese These steht der Titel des Buches „Left Back: A Century of Failed School Reform" (2000), der als eine Abrechnung mit traditionellen pädagogischen, speziell reformpädagogischen Diskursen, gewertet werden kann.

man konkret Ziele und Inhalte der Lehrerbildung benennen und standardisieren kann, um eine Verbesserung zu erreichen.
Bezogen auf die einzelnen Bundesstaaten ergeben sich hier unterschiedliche Ansätze, die nicht in ihrer Breite dargestellt werden sollen. Insgesamt lassen sich jedoch drei Standardgruppen und jeweils für sie stehende Institutionen unterscheiden, die verschiedene Interessen hinsichtlich der Beeinflussung des amerikanischen Bildungssystems verkörpern, aber alle bestimmend für die dort geführte Diskussion sind (Tulodziecki & Grafe 2006):

- Standards des „National Council for Accreditation of Teacher Education" für die „Accreditation of Schools, Colleges and Departments of Education" (NCATE 2002); sie stellen in erster Linie Institutionenstandards dar
- Standards des „Interstate New Teacher Assessment and Support Consortium" für die Zulassung und Zertifizierung (auch fachfremder) Personen, die in der Schule unterrichten wollen (INTASC 1992)
- Standards des „National Board for Professional Teaching Standards" (NBPTS 2003) für die Lizenzierung und Zertifizierung erfahrener Lehrpersonen

Die sogenannten INTASC-Standards sind zu Beginn der 1990er Jahre aus einem Projekt hervorgegangen, welches vom Council of Chief State School Officers, dem Verband der höchsten Bildungsadministratoren aus den 50 Bundesstaaten der USA, gefördert wurde. Als Ergebnis des Projektes wurden zehn Standards vorgelegt, die jeweils noch durch die Benennung von zugehörigen Wissenselementen, Fähigkeiten und Dispositionen gekennzeichnet waren. Die INTASC-Standards explifizieren die wesentlichen Wissens- und Fähigkeitselemente, über die jeder Lehrer verfügen sollte. Sie beziehen sich auf folgende Bereiche:

- Fachwissen und Lebensweltbedeutung
- Entwicklungsprozesse und Lerntheorien
- Lernziele und Differenzierung
- Entwurf und Umsetzung von Unterrichtsstrategien
- Motivation und Gestaltung von Lernumgebungen
- Kommunikation und Medien
- Unterrichtsplanung und Unterrichtsdurchführung
- Förderung und Beurteilung
- Reflexion und Evaluation
- Interpersonale Beziehungen und schulische Umwelt

Die aufgeführten Standards sind das Ergebnis von konsensualen Entscheidungen zusammengestellter Expertengruppen aus der Wissenschaft und Praxis und spiegeln keine Daten aus empirischen Forschungsprojekten wider (Zeichner 2006, S. 100f.). Wegleitend für die Expertengruppen waren fünf zentrale Annahmen über gelingendes Unterrichten, die wiederum vom National Board for Professional Teacher Standards (NBPTS) übernommen wurden und als Leitziele der Lehrerbildung interpretiert werden können:

- Teachers are committed to students and their learning

- Teachers know the subjects they teach and how to teach those subjects to students
- Teachers are responsible for managing and monitoring student learning
- Teachers think systematically about their practice and learn from experience
- Teachers are members of learning communities

Sowohl die Leitziele als auch die Standards sind einem schülerorientierten Unterricht verpflichtet und können als Professionalisierungsansätze der Lehrerbildung gewertet werden. Ihre Anhänger, als prominentes Mitglied etwa Darling-Hammond, wollen die Lehrerbildung aufwendiger gestalten, tendenziell verlängern und betonen die erforderliche Berücksichtigung genuin pädagogischer sowie fachdidaktischer Ausbildungselemente. Die generelle Notwendigkeit von Standards wird dabei allerdings nicht in Zweifel gezogen.

„In organized professions, the major lever for professionwide transfer of knowledge and continuous improvement of preparation is the development and enforcement of professional standards through (1) professional accreditation of preparation programs; (2) state licensing, which grants permission to practice; and (3) certification (like board certification in medicine, psychology, and now, teaching), which is a professional recognition of high-level competence" (Darling-Hammond, Wise & Klein 1999; zit. n. Darling Hammond 2006, S. 314).

In den Vereinigten Staaten lässt sich ein zweiter Diskussionsstrang zur Veränderung der Lehrerbildung ausmachen, der sich aus Sicht der Profession als Deregulierungsanliegen versteht. Im Deregulierungsdiskurs wird die Frage der Ausbildungskonzeption der Lehrerbildung bzw. der Zertifikate, die eine Person aufweisen muss, um an staatlichen Schulen unterrichten zu dürfen, dahingehend thematisiert, dass eine fachwissenschaftliche Ausbildung genügendes Fundament bietet bzw. dass sogar der gesamte Lehrerarbeitsmarkt dereguliert werden sollte. Schulen hätten dann die Möglichkeit, Personen einzustellen, die sie für geeignet halten und könnten diese bei festgestellter Nichteignung wieder entlassen. Das Kriterium für die Eignung als Lehrer wäre die Überprüfung von Schülerleistungen anhand standardisierter Testverfahren. Gemäß dieser Linie sind einige Bundesstaaten bereits heute dazu übergegangen die Zulassung zum Schuldienst an standardisierte fachwissenschaftliche und pädagogische schriftliche Testverfahren sowie ein einwandfreies polizeiliches Führungszeugnis zu knüpfen.
Die Errichtung ambitionierter Standardkonzepte kann also auch als bildungspolitische Bemühung verstanden werden das Verhältnis zwischen der Lehrerprofession und dem Staat dahingehend zu beeinflussen, die Steuerung des Bildungssystems in erster Linie über Standards zu vollziehen, die von Experten der Profession selbst formuliert worden sind, wodurch (bundes)staatliche Einflussmöglichkeiten nur noch indirekt gegeben wären (Darling-Hammond 2006, S. 314). In dieser Sichtweise werden die Professionsstandards gewissermaßen zu einem „Bollwerk", das die Profession vor zu direkten Zugriffen staatlicherseits schützt. Mintrop (2006, S. 101) macht jedoch deutlich, dass diese argumentative Strategie in ihrer Umsetzung nicht unbedingt aufgeht, weil der Erfolg von Professionalisierungsbemühungen abhängig ist von der politischen Kooperation einzelner Bundesstaaten oder der wohlwollenden Unterstützung der

Bundesregierung. So hat das von der Bush-Regierung 2001 verabschiedete Gesetz „No Child Left Behind“ umfangreiche Mittel für kompensatorische Bildungsmaßnahmen bereitgestellt, knüpft allerdings die Inanspruchnahme dieser Mittel an den Einsatz vorgegebener Testverfahren, die ihren Schwerpunkt auf der standardisierten Messung von Schülerleistungen in wenigen Fächern haben. Die an dem Programm teilnehmenden Schulen verpflichten sich zu einem Berichtssystem, in dem die Leistungsdaten im mindestens jährlichen Turnus dem Staat zugänglich gemacht werden. Guter Unterricht wird damit auf die Erreichung guter Testwerte enggeführt. Bei Nichterreichung entsprechender Werte droht einzelnen Schulen nach fünf Jahren die Schließung oder private Weiterführung bzw. betroffenen Lehrern die Kündigung.
Die auszumachenden diskursiven Spannungspole zwischen Professionalisierern und Deregulierern verdeutlichen, dass in den USA die Diskussion um Standards in der Lehrbildung nicht mehr so grundsätzlich (wie zur Zeit noch in Deutschland) geführt wird, sondern dass es verstärkt darum geht, auf welcher Ebene Standards eingeführt werden, was ihre Funktionen sind und welche Folgen man aus den Ergebnissen ziehen kann. Die in der Deregulierungsdebatte aufscheinenden Konsequenzen der Standardthematik markieren einerseits einen „Nebenkriegsschauplatz“, weil hier lediglich die Marginalisierung pädagogischer Kompetenzen fortgeführt wird, die von jeher bestimmend für die Bedeutung der Pädagogik in einer fachfremden Wahrnehmung gewesen ist, andererseits betonen sie ein professionsbezogenes Gefährdungspotential der Interpretation standardbezogener Ergebnisse. Trotzdem spiegeln die Entwicklungen meines Erachtens keine grundsätzliche Kritik an Standardkonzepten, sondern unterstreichen nur, dass sich Standards in vielfältiger Weise instrumentalisieren lassen.

5.2 Standardbasierte Untersuchung der Wirksamkeit der Lehrerbildung in der Schweiz (Oser)

Als Ergebnis der Zusammenarbeit der schweizerischen Hochschulen Fribourg (unter der Projektleitung von Oser) und Bern (unter der Projektleitung von Oelkers) im Zeitraum von 1994 bis 2000 stellen Oser & Oelkers 2001 eine Studie über die umfassende empirische Untersuchung der Wirksamkeit der schweizerischen Lehrerbildung vor (Oser & Oelkers 2001). Obwohl die Studie mehreren wichtigen Fragestellungen nachgegangen ist, so z.B. über die Bedeutung der Motivation in der Lehrerbildung, wurde sie am bekanntesten durch diejenige Teilstudie, die sich mit einer standardbasierten empirischen Untersuchung der Wirksamkeit der Lehrerbildung in der Schweiz befasst hat. Das zugrunde gelegte Standardkonzept wurde bereits 1997 veröffentlicht (Oser 1997a, b) und hat schon damals für Diskussionen gesorgt.
Oser & Oelkers (2001, S. 12) wollen mit ihrer Studie Entwicklungschancen für die Lehrerbildung aufzeigen, die keine bildungstheoretisch-idealistische Positionen bedienen, sondern kompetenzorientiert auf das Theorem der „Tüchtigkeit“ verweisen. Damit vertreten sie ein Verständnis, nach dem das Hauptproblem der Lehrerbildung in der Schweiz in der fehlenden Ausbildung von professionellen Kompetenzen zu

sehen ist, weniger hingegen in der mangelnden Ermöglichung von Reflexionsmöglichkeiten.
Die von Oser formulierten Standardgruppen sind nicht im Alleingang von ihm zusammengestellt worden, noch gehen sie aus einer theoretisch fundierten Konzeption hervor. Er stellt sogar ausdrücklich fest, dass Standards in der von ihm geprägten Bedeutungsart nicht alleine aus der Theorie abzuleiten sind (Oser 2006). Vielmehr präsentiert er sie als Synopse von handlungsorientierten Kompetenzbereichen, die in einem dreistufigen Prozess der Befragung von Didaktikern, Fachdidaktikern und Personen aus dem schulischen Praxisfeld als besonders wichtig für eine erfolgreiche Unterrichtsgestaltung herausgestellt wurden. Konkret sind sie das Ergebnis von Expertengesprächen und -befragungen. Oser bezeichnet sein Vorgehen als „Quasi-Delphi-Studie“, weil die Experten ihre Einschätzungen auf einer Skala, mit den Dimensionen „absolut notwendig“ bzw. „nicht notwendig“, vornahmen. Er verweist darauf, dass Standards bzw. Standardgruppen als unterschiedlich bedeutsam eingeschätzt wurden, woraus allerdings nicht gefolgert werden könne, dass einzelne Standardgruppen für sich genommen unwichtig wären (Oser 1997a, S. 30).
In der Übernahme dieser Gewichtungen gelang ihm eine Komplexitätsreduktion auf zwölf Standardgruppen mit insgesamt 88[32] Einzelstandards. Der Bereich der Fachwissenschaft bzw. Fachdidaktik blieb dabei weitgehend ausgeklammert. Oser konnte die folgenden 12 Standardgruppen identifizieren (Oser 1997a, S. 31; 2001, S. 230):

1. Lehrer-Schüler-Beziehungen
2. Schülerunterstützendes Handeln und Diagnose
3. Bewältigung von Disziplinproblemen und Schülerrisiken
4. Aufbau und Förderung von sozialem Verhalten
5. Lernstrategien vermitteln und Lernprozesse begleiten
6. Gestaltung und Methoden des Unterrichts
7. Leistungsmessung
8. Medien
9. Zusammenarbeit in der Schule
10. Schule und Öffentlichkeit
11. Selbstorganisationskompetenz der Lehrkraft
12. Allgemeindidaktische und fachdidaktische Kompetenzen

Vergleicht man die Oserschen Publikationen von 1997 und 2001 miteinander, so kann man feststellen, dass die Standardgruppen nur terminologisch leicht verändert worden sind und auch die Einzelstandards nur geringfügige inhaltliche Veränderungen aufweisen. Oser stellt fest, dass die Standards auf einer Stufe der Handlungskompetenz formuliert worden sind, weshalb sowohl kommunikative als auch inhaltliche und systematisch organisatorische Anteile mit berücksichtigt werden

[32] Irritierend scheint, dass in der Erstveröffentlichung der Standardgruppen zwar von 88 Einzelstandards die Rede ist (Oser 1997a), im Text aber insgesamt 90 aufgeführt werden. In der neueren Version (Oser 2001, S. 230-242) findet man insgesamt 89 Einzelstandards, worauf unter Betonung von 88 wiederum nur kurz hingewiesen wird (ebd., S. 242). Insofern überrascht es, dass sich die Anzahl 88, sowohl in den Primär- als auch in den Sekundärquellen, durchgesetzt hat.

mussten. Hiermit erklärt sich die doch vergleichsweise große und damit unübersichtliche Anzahl der Einzelstandards. Dabei verweisen die Standardgruppen auf unterschiedliche Referenzpunkte von schulischen Handlungsfeldern, im Kern steht jedoch der eigentliche Unterricht, auf den sich die ersten acht Standardgruppen und die letzte Standardgruppe beziehen. Innerhalb dieser neun unterrichtsnahen Standardgruppen kann noch eine Unterscheidung zwischen entwicklungspsychologisch-sozialkognitivem Unterstützungshandeln (Standardgruppen 1, 2 und 4), Klassenmanagement (Standardgruppe 3), allgemeindidaktischen Handlungsentwürfen (Standardgruppen 5-8) sowie der Fachdidaktik (Standardgruppe 12) getroffen werden. Die Standardgruppen 9-11 beinhalten eher unterrichtsfernere, organisationsbezogene bzw. übergreifende Kompetenzprofile in Form von Kooperationsformen innerhalb der Schule sowie der Kommunikation mit Personenkreisen außerhalb der Institution. Die Standardgruppe 11 „Selbstorganisationskompetenz“ zielt auf übergreifende Kompetenzprofile, die sich auf alle anderen Standardgruppen auswirken können.
Die von Oser formulierten Standards sind somit anschlussfähig an psychologisch-lerntheoretische Konzepte, allgemeindidaktische und fachdidaktische Theoriebildung sowie organisationstheoretische Überlegungen. In diesem hohen Transferpotential kann ein Grund dafür gesehen werden, dass sein Modell in der Schweiz eine herausgehobene Bedeutung erlangt hat, die auch nach Deutschland ausstrahlt. Gleichzeitig erweckt die Bezugnahme auf verschiedenste Theoriekonzepte den Anschein einer gewissen Beliebigkeit, spiegelt aber auch das Fehlen einer Metatheorie, aus der sich alle schulpädagogischen Handlungsdimensionen – inklusive ihrer organisatorischen Implikationen – ableiten lassen. Die Beschreibung von handlungsorientierten Kompetenzprofilen ist offenbar auf Bezugswissenschaften angewiesen, ein Sachverhalt, der sich auch auf die Erziehungswissenschaft insgesamt ausdehnen lässt. Oser spricht selbst von einem unauflösbaren Legitimationsdefizit (Oser 2001, S. 242), weshalb sein Standardkonzept keinen Anspruch auf Vollständigkeit erheben kann, sondern sich eher auf der Basis negativer Selektion legitimiert; danach müsste man gegebenenfalls begründen können, warum ein einmal formulierter Standard keine Bedeutung mehr für die Unterrichtsrealität haben soll.
Die Frage der Anzahl von Standards lässt sich nach Oser nicht endgültig sinnvoll beantworten. Jedoch kann die Anzahl auch nicht als unendlich beschrieben werden, weil sie maximal so groß ist, wie es notwendige Handlungssituationen im Unterricht gibt. Ausdrücklich hebt Oser hervor, dass Standards keinen Schlüsselkompetenzen entsprechen, sondern stets an konkret definierte Situationen und Inhalte gekoppelt sind und somit nicht „dem Mythos der Transferierbarkeit“ (Oser 2006) erliegen.
Die Entwicklung einer Synopse von Kompetenzprofilen, die von schulischen Experten als besonders bedeutungsvoll eingeschätzt werden, ist als Leistung Osers unbedingt anzuerkennen. Darüber hinaus ist es ihm gelungen den Begriff des „Standards“ in der Schweiz zu etablieren, indem er spezifische theoretische Modelldefinitionen an den Begriff knüpft, die als Gütekriterien für die Professionalisierung der Lehrerbildung gewertet werden können. Oser (1997a, S. 28) verwendet den Standardbegriff in zweierlei Hinsicht:

„a) für eine hochprofessionelle Kompetenz, und b) für deren optimale Erreichung. Wir meinen also, Standards seien optimal ausgeführte bzw. optimal beherrschte und in vielen Situationen anwendbare Fähigkeiten und Fertigkeiten, die nur von Professionellen Verwendung finden können, aber nicht von Laien oder von Personen anderer Professionen."

Mit dieser Begriffsbildung werden zwei Definitionseinheiten miteinander verbunden, um das begriffliche Analysepotential zu erhöhen und eine professionstheoretische Differenzierung zu ermöglich. Im Rahmen der tiefer gehenden Begriffsbildung nennt Oser vier zu erfüllende Gütekriterien, die in ihrer Gesamtheit die Rede vom „Standard" erst rechtfertigen (Oser 1997a, S. 29):

a) Das Kriterium der Theorie
b) Das Kriterium der Empirie
c) Das Kriterium der Qualität
d) Das Kriterium der Ausführbarkeit

Oser geht davon aus, dass von einem Gütekriterium der Theorie dann gesprochen werden kann, wenn es plausible theoretische Feststellungen zu einem Standardfeld gibt, die es ermöglichen konditionale Aussagen hinsichtlich der Anwendung eines Standards und diesbezüglicher Auswirkungen zu generieren. Theorien (Kriterium a) erlauben in dieser funktionalen Lesart Prognosen über standardorientiertes Handeln. Damit wird unmittelbar auf das zweite Kriterium der Empirie (b) verwiesen. Empirische Untersuchungen können theoretische Aussagensysteme falsifizieren oder auch bestätigen. Somit verwendet Oser Theorie- und Empiriebegriffe, die erst in ihrem wechselseitigen Bezug zueinander die Grundlage gehaltvoller Aussagen darstellen und konstruiert sie nicht als Gegensatzpaar. Auf dieser Basis lassen sich weiterhin Aussagen über Qualitätsstufen (Kriterium c) treffen. Bauen sie zwar auf Theorien und empirischen Ergebnissen auf, werden sie aber vor dem Hintergrund lebensweltlicher Einflüsse und Begründungen realisiert. Schließlich sind Standards so zu formulieren, dass sie vereinbar mit Handlungstraditionen im Feld erscheinen und eine Praxisrelevanz aufweisen. Derartige Traditionen unterliegen damit ebenfalls dem Einfluss durch Theorien und Expertentum, beinhalten aber auch eine spezifische Kultur der Praxis (Kriterium d), um hieran anschlussfähig zu sein.

Die genannten Kriterien stellt Oser zum einen als gleichwertig heraus, in dem Sinne, dass man eben nicht von einem Standard sprechen kann, wenn eines der Kriterien nicht erfüllt ist. Dabei sind alle Kriterien wechselseitig aufeinander bezogen und entfalten erst in dieser Bezogenheit ihr Potential. Zum anderen soll die alphabetische Reihenfolge implizieren, dass in der Abfolge jeweils ein größeres Ressourcenmoment hinzukommt, d.h. die nachfolgenden Kriterien in der bezeichneten Abstufung die vorangegangenen Kriterien jeweils umschließen. Diesen etwas abstrakten Sachverhalt kann man mit dem Verhältnis von Theorie und Praxis illustrieren, da theoretische Aussagen allein niemals ein standardbezogenes Handeln in der lebensweltlichen Praxis beschreiben können, womit Theorie gewissermaßen in der Praxis erst aufgeht. Professionelle Standards ergeben sich letztlich aus der Vernetzung der vier Kriterien (Oser 2001, S. 224f.):

„Standards sind keine skills, also Automatismen, die zur Gedächtnis- und Bewusstseinsentlastung führen. Standards sind komplexe, berufliche Kompetenzen, die zu theoriegeleitetem Handeln werden, dies weil ein Bezug zur Wissenschaft und Forschung einerseits besteht und weil andererseits eine analysierte und dadurch kritisch reflektierte Praxis diese Praxis erst ermöglicht."

Die vier standardbezogenen Gütekriterien stellen eine metatheoretische Orientierung zur Fundierung des Standardbegriffes dar. Darüber hinaus markieren sie aber Lernwege, die Studierende zur Aneignung von Standards beschreiten müssen. Für das Curriculum von Studiengängen des Lehramtes könnte das heißen, dass Studierende Theorien und Informationen über bestimmte Standards bzw. Standardgruppen erhalten, sie diesbezügliche empirische Studien lesen, über teilnehmende Beobachtung oder Videostudien einen Eindruck von Qualitätsunterschieden standardorientierter Handlungsformen vermittelt bekommen, um schließlich darauf bezogen das eigene Handeln in der Praxis zu erproben und gemeinsam mit Experten zu diskutieren und zu reflektieren. Dieses komplexe Verfahren stellt den angemessenen Weg dar, um Standards in der Lesart von Oser zu vermitteln. Damit ist aber ein Aufwand verbunden, der seitens der einzelnen Institution sicher nicht für alle 88 Standards und vermutlich auch kaum für die zwölf Standardgruppen betrieben werden kann. Skizzierte Lernwege würden es hingegen nahe legen sogenannte „Portfolioansätze" stärker in der Lehrerbildung zu fokussieren. Oser (1997a, S. 36f.) bezeichnet Portfolios als

„Sammlungen von reflektierten Erfahrungsstücken, Unterrichtsmaterialien, analysierten Video-, Tagebuch- und Unterrichtsprotokollaufzeichnungen, Berichten über Zusammenarbeit mit Expertenlehrern u.ä. Portfolios sind stets das Resultat eines längeren Theorie-, Übungs- und Praxisumgangs hinsichtlich einer professionellen Kompetenz. Sie sind ein Hinweis der Erreichung von Standards."

Schließlich verlangen auch portfoliofundierte Lernwege zur Aneignung von Standards, die Lernwege mit Hilfe spezifischer Erhebungsverfahren zu rekonstruieren. Die Oser-Studie (1997) stellt exemplarisch vor, wie das für Erhebungsformen von Selbsteinschätzungsverfahren aussehen kann.[33] Für andere, stärker objektivierende Test-, Beobachtungs- und Beurteilungsverfahren liegt in der Berücksichtigung der Gütekriterien eine große methodische Herausforderung, wobei zurzeit noch Spezifizierungen darüber ausstehen, wie sie z.B. für Beobachtungs- und Beurteilungsverfahren angegangen werden kann.

Oser ist es gelungen den Standardbegriff weiterzuentwickeln, theoretisch aufzuladen und von der alltagssprachlichen Verwendung im Sinne von „Vereinheitlichung" und „Vergleichbarkeit" abzuheben. Dabei hat er das Standardkonzept sukzessive ausgebaut und auf verschiedene pädagogische Felder angewandt (Oser 2004; Oser & Renold 2005; Oser, Curcio & Düggeli 2007). Oser & Renold (2005, S. 129) sprechen von einer „multilevel Dimensionierung" eines jeden Standards. Als konkrete Dimensionen nennen sie Abgrenzbarkeit, inhaltliche Zieladäquatheit und Effektivität bezogen auf das Lernen sowie ethische Reflexivität.

[33] Zu der Vorgehensweise und den Ergebnissen der Studie vgl. Kapitel 6.

In einer berufsethischen Perspektive diskutieren Oser & Renold (2005, S. 124) die Entwicklung von Kompetenzprofilen als individuelle Entwicklungsaufgabe. Genauer ergeben sich hier Kompetenzprofile aus der situativen Analyse von offenen pädagogischen Notwendigkeiten.

„Ein Kompetenzprofil, das unentbehrlich ist für qualifiziertes Lehrpersonenhandeln, muss also den Test einer offenen pädagogischen Notwendigkeit passieren, um den professionellen Anforderungen Stand zu halten. Dies alles sind Aspekte eines noch besser zu entwickelnden kompetenztheoretischen Ansatzes für das Berufsfeld ‚Lehrkräfte'."

Standards werden also unter dem Gesichtspunkt der Handlungsnotwendigkeit nicht nur beschreibend, sondern auch normativ gefasst. Die normative Setzung einer Aneignung von Standards verdeutlicht sich auch an der Aussage „Lehrerstudierende müssen selber die Standards erfüllen wollen" (Oser 2001, S. 332). Standards können in dieser Perspektive erst dann sinnvoll eingesetzt werden, wenn sie das Prinzip der Notwendigkeit berücksichtigen (Oser 2006). Damit sind spezifische Anforderungspotentiale gemeint, die sich in pädagogischen Situationen ergeben und als handlungsleitende Aspekte mitzudenken sind. In diesem Sinne verlangen spezifische pädagogische Situationen und Interaktionen jeweils passungsfähige Kompetenzprofile und übergreifend den flexiblen Einsatz von situtionsadäquaten Profilen (Oser, Curcio & Düggeli 2007, S. 23):

„Mit dem Konzept der Pädagogischen Notwendigkeit wird die Kompetenzmessung als Interaktion zwischen Person und Situation verstanden. Es geht also nicht mehr nur um die Sichtbarmachung der Kompetenzprofile einer Lehrperson, sondern um die Elemente der Situation, welche den Einsatz eines bestimmten Kompetenzprofils erfordern."

Der Begriff der „Pädagogischen Notwendigkeit" als Abarbeitungsfolie für die Entwicklung professioneller Kompetenzprofile ist bislang theoretisch wenig bestimmt. Es ist evident, dass sich verschiedene pädagogische Situationen und Interaktionsmuster, wie z.B. Disziplinprobleme oder diagnostische Fragestellungen, grundsätzlich voneinander unterscheiden lassen. Hiervon ausgehend ist auch die Entwicklung einer Typologie pädagogischer Situationen und implizierter notwendiger Handlungsmuster vorstellbar. Die Verankerung von Kompetenzprofilen in ihrer situativen Notwendigkeit stellt eine konsequente Fortführung des Standardansatzes dar. Allerdings sind diesbezüglich vermutlich nur Annäherungen möglich, bedenkt man die Komplexität unterrichtlicher Prozesse. Speziell in seinem didaktischen Potential ist dieser Ansatz aber vielversprechend, so in Verbindung mit situationsbezogenen Beispielen in Form videographierter Unterrichtsstunden, die Reflexionen auf der Studierendenseite anregen können.
Genau diese Ausrichtung verfolgt Oser u.a. in einem aktuellen Projekt mit dem Titel „Professionel Minds", in dem er den Standardansatz auf berufspädagogische Ausbildungen überträgt. In einem dort angesiedelten Teilprojekt verfolgt er das Ziel standardgemäßes Handeln filmisch darzustellen und die Verfilmungen für didaktische und diagnostische Zwecke fruchtbar zu machen (Oser & Renold 2005, S. 132). Oser, Curcio & Düggeli (2007, S. 15) identifizieren für das Feld der

Berufsbildung 45 Kompetenzprofile für Unterrichtende, die in vier Hauptgruppen mit jeweils zwei bis drei Untergruppen untergliedert sind.

Standardhauptgruppen	Standarduntergruppen
A Standards der Lehre	A1 Unterrichtsvorbereitung A2 Vermittlungsformen
B Standards der Lernbedingungen	B1 Lernbedingungen B2 Wert- und Konfliktmanagement
C Standards des Lernprozesses	C1 Diagnose C2 Begleitung C3 Evaluation
D Berufscoping und Zusammenarbeit	D1 Zusammenarbeit D2 Coping der Lehrperson

Tabelle 5-1: Übersicht der thematischen Standardhaupt- und -untergruppen

Die Autoren stellen heraus, dass Kompetenzen im alltäglichen Unterricht zumeist im Verbund eingesetzt werden und deshalb in der Regel nicht als scharf voneinander abgegrenzte Einzelkompetenzen beobachtet werden können. Diese Einschätzung teile ich ausdrücklich. Für die Entwicklung von hoch-inferenten Beobachtungsverfahren ist damit die Notwendigkeiten gegeben, einzuschätzende Merkmals- und Handlungsbereiche genau zu explifizieren (vgl. Kap. 10.5 und Anhang). Curcio & Düggeli sprechen in diesem Zusammenhang von Kompetenzbündeln oder Kompetenzprofilen. Den Standardbegriff verwenden sie hingegen nur mit Bezug auf Fragen der Qualität bzw. der Messbarkeit.[34]

5.3 Standards für die Lehrerbildung in Deutschland (Terhart)

Terhart (2002) hat eine Expertise vorgelegt, die im Auftrag der KMK die Möglichkeiten und Grenzen einer standardorientierten Evaluation der deutschen Lehrerbildung diskutiert. Auftragsgemäß wurden dabei die erste Phase der Lehrerbildung (Studium an einer Hochschule) und die zweite Phase (Referendariat am Studienseminar), nicht jedoch die dritte Phase (Fort- und Weiterbildung der Lehrer im Beruf) in den Blick genommen. Die Einführung von Standardkonzepten als kriterielle Grundlage für eine Evaluation der Lehrerbildung stellt seiner Meinung nach eine in gewisser Weise revolutionäre Herangehensweise dar, weil sich damit die grundlegende Betrachtungsweise für eine Verbesserung der Lehrerbildung ändert. Der Schwerpunkt liegt nicht mehr auf der Vorgabe detaillierter Inputs durch Gesetze, Lehrpläne, Erlasse etc., sondern orientiert auf die Erfassung der Outputs bzw. Outcomes der Lehrerbildung, also auf das, was Lehrerbildung tatsächlich erreicht und möglicherweise nur erreichen kann. Die Einführung von Standardkonzepten

[34] Eine derart systematische Abgrenzung zwischen Kompetenzen, Kompetenzprofilen und Standards hätte auch den vorangegangenen Publikationen von Oser eine insgesamt stärkere Klarheit verliehen.

stellt somit ein Steuerungssystem dar, das zum einen eine evaluative Diagnose des Ist-Zustandes liefert (summativ), auf dessen Basis zum anderen eine Weiterentwicklung des Systems (formativ) möglich erscheint. In dieser Perspektive genügt die Einführung von Standardkonzepten keinem Selbstzweck, sondern liefert einen viel versprechenden Möglichkeitsrahmen für die Verbesserung der Lehrerbildung.

„Die Erörterung von Standards in der Lehrerbildung wird – wie schon angedeutet – berücksichtigen müssen, dass eine möglichst hohe Qualität der Lehrerbildung nicht um ihrer selbst willen angestrebt wird, sondern darauf abgestellt ist, solche Kompetenzen und schließlich ein solches Lehrerhandeln anzubahnen, zu befördern und möglichst weitgehend sicherzustellen, welches dann auf Seiten der Schüler zu den allgemein gewünschten Wirkungen in kognitiver, sozial-moralischer und ästhetisch-expressiver Hinsicht führt“ (Terhart 2002, S. 11).

Terhart hebt damit die weiterführende Zielperspektive einer solchen Evaluation hervor, macht aber zugleich deutlich, dass in seinem Evaluationskonzept, auch wenn es prinzipiell wünschenswert wäre, Schülerleistungen zunächst keine Berücksichtigung finden. In einer ersten Annäherung geht es erst einmal darum grundlegende methodische Problemstellungen zu lösen. Das Verhältnis zwischen einem curricularen Angebot, den erfahrungsbasierten Lernprozessen von Studierenden des Lehramtes bzw. ausgebildeten Lehrern und der Entwicklung von berufsbezogenen Kompetenzen enthält nach Terhart viele Fragezeichen. Komplexer noch gestaltet sich die Fragestellung, wenn man Entwicklungsprozesse von Lehramtsstudierenden konditional auf Lernprozesse von Schülern bezieht. Generell muss die empirische Untersuchung von Wirkzusammenhängen, so Terhart, unter der Berücksichtigung großer Zeiträume erfolgen. Auch dann besteht noch die Gefahr, dass wichtige, aber nicht unbedingt messbare Effekte der Lehrerbildung vernachlässigt werden. Die Definition von Standards möchte der Autor auf verschiedene Ebenen der Lehrerbildung bezogen wissen, wobei er folgende Ebenen identifiziert:

- Standards für ausgebildete Personen (Absolventenstandards)
- Standards für Institutionen der Lehrerbildung
- Standards für das Steuerungssystem der Lehrerbildung

Eine Evaluation der ersten Phase der Lehrerbildung berücksichtigt die vier konstituierenden Grundelemente der hochschulischen Ausbildung, nämlich Fachwissenschaft, Fachdidaktik, Erziehungs- und Bildungswissenschaften sowie schulpraktische Studien. Für die drei erstgenannten Elemente formuliert Terhart jeweils zehn Standards, für die schulpraktischen Studien lediglich fünf. Die Standards sollen hier nicht ausführlich dargestellt werden, im Kern liefern sie aber das inhaltliche Gerüst für die Entwicklung eines Curriculums. Auf einer zweiten Dimension entwirft Terhart ein Modell von Absolventenstandards, welche die Ziele der ersten Phase der Lehrerbildung, bezogen auf die vier Grundelemente, beschreiben (Terhart 2002, S. 30):

- Wissensbasis für und über das spätere Berufsfeld

- Reflexionsfähigkeit über Sachthemen, aber auch die eigene Person in Verbindung mit den Anforderungen des beruflichen Feldes
- Kommunikationsfähigkeit über Inhalte, Strukturen und Probleme des unterrichtsfachlichen, pädagogisch-didaktischen und schulbezogenen Bereichs
- Urteilsfähigkeit angesichts pädagogischer Handlungsprobleme und Entscheidungsfragen

Terhart weist darauf hin, dass die vier Kompetenzbereiche als Personenstandards nur formal voneinander abgegrenzt werden können und als kontinuierliche Entwicklungsaufgabe im Sinne eines Stufenmodells zu modellieren sind.
Damit präsentiert er ein Kompetenzmodell, das in der ersten Dimension inhaltliche Standards für eine Topologie des Kerncurriculums für die Lehrerbildung liefert und auf einer zweiten Ebene die Taxonomie von Kompetenzfacetten beschreibt (vgl. Baumert & Kunter 2006). Die dritte Dimension zielt wiederum auf die Professionsentwicklung als biographischer Prozess, wobei Terhart dafür plädiert, Standards in der Lehrerbildung lebensphasenorientiert zu formulieren (Terhart 2002, S. 27).

„So wird denn auch darauf hingewiesen, dass erst nach der Berufseingangsphase und während des kontinuierlichen Weiterlernens im Beruf sich diese Standards immer stärker ausprägen bzw. auch immer deutlicher ausprägen sollten. Die Lehrerausbildung in erster und zweiter Phase ist dafür nur die Ausgangsbasis: ihr Ziel kann und sollte nicht der kompetente Lehrer, sondern der kompetente Berufsanfänger sein."

Die oben dargestellten Kompetenzfacetten werden in der zweiten Phase der Lehrerbildung als Ergebnis der Entwicklung und Erprobung eigener beruflicher Handlungs- und Reflexionsfähigkeit weiterentwickelt. Auf der Ebene der Personenstandards formuliert Terhart folgende Aufgabenbereiche (Terhart 2002, S. 30):

- Planung, Durchführung und Auswertung des eigenen Unterrichts
- Erarbeitung eines breiten, fachbezogenen und ebenso fächerverbindenden Methodenrepertoirs
- Erfassung spezifischer Lernbedürfnisse von Schülern
- Kooperation mit Kollegen zum Zweck der Unterrichts- und Schulentwicklung
- Kommunikation und Beratung von Eltern
- Zusammenarbeit mit außerschulischen Institutionen

Die von Terhart vorschlagenden Standards für die Evaluation der an der Lehrerbildung beteiligten Institutionen bzw. ihrer Steuerungssysteme stehen nicht im Fokus des eigenen Frageinteresses und erfahren hier keine genauere Aufmerksamkeit.
Herauszuheben an dem von Terhart entwickelten Expertisemodell ist aber noch, dass er eine Stufung von standardorientierten Evaluationsverfahren auf der Personenebene vornimmt, die aufsteigend jeweils aussagekräftigere Daten liefern (Terhart 2002, S. 37):

1. Selbsteinschätzungen
2. Testverfahren
3. Beobachtung und Beurteilung
4. Lernleistung / Erfahrung der Schüler

Selbsteinschätzungen liefern danach vergleichsweise unsichere Daten, während Lern- und Erfahrungsleistungen der Schüler als kriterielle Evaluationsdaten zwar am aussagekräftigsten wären, methodisch aber viele Schwierigkeiten in sich bergen (s.o.). Insofern hätte sich eine Evaluation der Lehrerbildung auf Testverfahren bzw. Beobachtungen und Beurteilungen zu konzentrieren. Solcherlei Testverfahren liegen in Deutschland bislang aber nur in rudimentärer Form vor, während Beobachtungs- und Beurteilungsverfahren wiederum äußerst aufwendig durchzuführen sind. Die in der Expertise vorgenommene Stufung von Evaluationsverfahren folgt damit der aktuellen methodischen Diskussion über empirische Erhebungsverfahren (vgl. Erpenbeck & Rosenstiel 2003; Blömeke 2004).
Das von Terhart entwickelte Kompetenzmodell hat den Vorteil, dass es speziell für die deutsche Lehrerbildung entwickelt wurde und die hiesigen Besonderheiten auf der curricularen Ebene berücksichtigt. Zudem klärt es das Verhältnis der in der ersten bzw. zweiten Phase auszubildenden Kompetenzbereiche und bezieht sie aufeinander. Kompetenzentwicklung wird in einer berufsbiographischen Perspektive modelliert, die empirisch anschlussfähig erscheint. Eine solche Konstellation spricht für eine längsschnittliche Evaluation personenbezogener Standardgruppen, worauf Terhart auch ausdrücklich hinweist (vgl. Terhart 2002, S. 49); jedoch bewertet er auch Querschnittsstudien als gute Evaluationsgrundlage.
Abschließend kann man sagen, dass Terhart ein Standard- und Kompetenzmodell vertritt, das nach wie vor Aktualität für sich beanspruchen kann und auf plausible Weise die Inhaltsbereiche und Phasen der deutschen Lehrerbildung abdeckt. Angesichts dieses schon 2002 vorgelegten differenzierten Modells bleibt es unverständlich, warum das 2004 von der KMK veröffentlichte Standardmodell dazu nur wenige Anknüpfungspunkte aufweist. Die Standards der KMK erschöpfen sich in einer Beschreibung deklarativer Wissensbestände, wobei das Verhältnis der ersten und zweiten Phase rein funktional mit „kennen“ und „können“ beschrieben wird. Es mag dem föderalen Bildungssystem in Deutschland geschuldet sein, dass das Standardmodell der KMK im Vergleich zum früher konzipierten Modell so „blass“ erscheint, unverständlich bleibt die frühere Befunde ignorierende Entwicklung dennoch.
Ähnlich wie auch Oser kann man Terhart schließlich den Vorwurf machen, dass die Auswahl seiner Standards nicht Ergebnisse einer theoretisch fundierten Modellbildung widerspiegelt, sondern sich pragmatisch an den Erfordernissen des Feldes „Lehrerbildung“ orientiert. Nach wie vor gibt es jedoch keine Anzeichen dafür, dass sich eine zugrunde liegende Theorie ankündigt, weshalb sich die Formulierung von plausiblen Standardgruppen, wie sie Oser und Terhart vorgenommen haben, momentan als einzig gangbarer Weg andeutet.

5.4 Vergleich der Standardmodelle von Oser und Terhart

In den Standardmodellen von Oser und Terhart spiegeln sich die jeweils länderspezifischen Ausbildungskonzeptionen wider. Die von Oser entwickelten pädagogisch-psychologischen Kompetenzprofile sind vergleichsweise stark auf den Unterricht fokussiert und zielen auf dessen Verbesserung. Lediglich die Standardgruppen „Zusammenarbeit in der Schule“ und „Schule und Öffentlichkeit“ zeigen Kompetenzprofile auf, deren Kernbereiche nicht in unterrichtlichen Prozessen aufgehen.[35] Dies kann in gewisser Weise als Manko gesehen werden, weil das Modell die Strukturanforderungen nur randständig „mitdenkt“, die jedoch für eine nachhaltige Standardimplementation unerlässlich zu berücksichtigen sind. Jenseits der Standardgruppen lassen sich denn auch keine zusätzlichen Ebenen und Stufen ausmachen, die z.B. eine Zuordnung zu Institutionen oder Tätigkeitsphasen (Hochschule/ Praktika/ Berufseingangsphase/ Berufpraxis) erlauben. Oser postuliert lediglich, dass die Arbeit an der Verbesserung standardorientierter Kompetenzen eine kontinuierliche Aufgabe für die gesamte Zeit der Ausbildung und Berufspraxis darstellt und insofern nie beendet ist. Oser & Oelkers (2001, S. 591) kommen in den Schlussfolgerungen aus den Ergebnissen Ihrer Studie zu der Überlegung:

„Vorstellbar sind Standardprofile, die für Novizen besonders geeignet sind; die Idee ist nicht, sämtliche denkbaren Standards des Berufsfeldes in der Grundausbildung zu erwerben. Es ist auch nicht so, dass die Summe aller Standards die perfekte Lehrkraft ausmachen würden.“

Im Gegensatz zu Oser entwirft Terhart ein Modell, dass genuin sowohl die Zweiphasigkeit[36] der Lehrerbildung in Deutschland als auch die sie betreibenden Institutionen reflektiert und für beide Phasen jeweils spezifische Standards formuliert. Die Kompetenzentwicklung wird darin als Stufenmodell längsschnittlich über die Phasen angelegt.
Die strukturellen Unterschiede in der Konzeption des schweizerischen und des deutschen Modells erschweren deren Vergleichbarkeit. Darüber hinaus orientiert Terhart in erster Linie auf pädagogische Kompetenzen, während Oser sich schon in seiner Formulierung von standardbezogenen Gütekriterien auf einen Theorie- und Empiriebegriff stützt, der eine große Affinität zur psychologischen Forschung aufweist. Diese starke Orientierung an der psychologischen Lehr-Lernforschung lässt sich u.a. mit der großen Bedeutung von Aebli für die Lehrerbildung in der Schweiz erklären. Die vom ihm entwickelten zwölf Grundformen des Lehrens bzw. seine „allgemeine Didaktik auf kognitionspsychologischer Grundlage“ (Aebli 1977) haben den didaktischen Diskurs in der Schweiz im Sinne der Integration von Forschungsergebnissen anschlussfähig an die empirisch-psychologische Lehr-

[35] Die Standardgruppe „Selbstorganisationskompetenz der Lehrkraft“ steht quer zu allen anderen Standardgruppen, berücksichtigt aber auch die Berufsanforderungen des Lehrerberufs außerhalb des Unterrichts.
[36] An sich ist die Lehrerbildung in Deutschland als dreiphasiges Modell angelegt. Dabei fällt im Literaturstudium auf, dass die sogenannte „dritte Phase“, nämlich die Berufpraxis, insgesamt recht stiefmütterlich behandelt wird. Trotzdem ist es überraschend, dass die KMK als Auftraggeber für die Expertise die dritte Phase der Lehrerbildung ausdrücklich ausklammert. Zu den Gründen dafür äußert sich Terhart nicht. Sein Verweis auf die „auftragsgemäße“ Auslassung der dritten Phase kann aber als Beleg dahingehend interpretiert werden, dass er mit der Position seiner Auftraggeber nicht unbedingt konform geht.

Lernforschung und daraus abgeleiteter Modelle gemacht. In Deutschland wurden die didaktischen Theorien von Aebli, im Gegensatz zu den entwicklungspsychologischen Theorien Piagets, dessen Schüler Aebli war, relativ wenig rezipiert. Es mag ein Zufall sein, jedoch ist es auffällig, dass die Standardgruppen von Oser in ihrer Anzahl den zwölf Grundformen des Lehrens von Aebli entsprechen. Die Fülle und der Detaillierungsgrad der von Oser vorgelegten handlungsorientierten Kompetenzprofile fallen speziell im Vergleich zu den Personenstandards, die Terhart für die erste und zweite Phase vorschlägt, ins Auge. Beide Modelle können als Ausgangspunkt zur Entwicklung von Forschungsdesigns dienen. Konsequent zu Ende gedacht bedingt aber das Stufenmodell von Terhart eine längsschnittliche Herangehensweise, während das Modell von Oser sowohl die Grundlage für einen Querschnitt als auch einen Längsschnitt bietet.

5.5 Fazit: Kritik an den Standardmodellen

Oelkers (2002) weist kritisch darauf hin, dass sich Standards zu einer Art „Lebensretter“ für das Bildungssystem entwickelt haben und als eine „Zauberformel“ Verwendung finden, der man die Lösung aller Probleme zutraut. Den expansiven Rückgriff auf den Standardbegriff setzt er mit einer neuen und durchsetzungsstarken Semantik gleich, die lediglich Forderungen aufstellt, ohne dass klar wird, wie diese genau überprüft werden können. An den Umgang mit dem Standardbegriff knüpft er damit drei Bedingungen.

- Standards müssen erfüllbar sein,
- beschränkt werden können und
- überprüfbar sein.

Bezogen auf das Konzept von Oser stellt Oelkers fest, dass die 88 vorgeschlagenen Standards die Ausbildungskapazität der Hochschulen bei weitem übersteigen und zudem eine wesentliche Größe nicht abdecken, nämlich die der Inhalte. Die fehlende Berücksichtigung von Inhaltsbereichen wurde auch von Terhart (2002, S. 24) kritisiert.

„Es ist auffällig, dass in diesen Standards von Wissen und Kompetenz im Fach bzw. in den Fächern nicht die Rede ist – ein Sachverhalt, der den Wert der Studie als Vorlage für eine die gesamte Lehrerbildung bzw. die gesamte Lehrerkompetenz umfassende Definition und Überprüfung von Standards stark einschränkt.“

An Terharts Gutachten übt Oelkers (2002) wiederum die Kritik, dass zwar zwischen Absolventenstandards und Standards für Institutionen und Steuerungssystemen unterschieden wird, externe Evaluationen aber zu wenig Berücksichtigung finden. Die Einführung von Standards würde so zu einem Teil der Organisationsentwicklung, der aber das Controlling fehlt, womit eine tatsächliche Weiterentwicklung unwahrscheinlich wäre. In einer historischen Analyse (Oelkers 2004a) zeigt er in diesem Zusammenhang auf, dass die Diskussion über die Einführung von Standardkonzepten auf ein System trifft, „dem Kontrollen fremd (sind, O.K.), das über

starke Bündnisse mit der Verwaltung verfügt und echte Leistungsbilanzen nicht kennt."

Die wohl massivste Kritik an der Einführung von Standardkonzepten in die Lehrerbildung, mit besonderem Bezug auf das Modell von Oser, wurde von Herzog (2005) vorgebracht. Er macht seine Kritik an fünf Punkten fest (Herzog 2005, S. 252):

1) Diffuse Begrifflichkeit
2) Theoretische Schwäche des Ansatzes
3) Latenter Ideologiegehalt
4) Fragwürdiges Verständnis von Lehrerprofessionalität
5) Gefährlicher Hang zu Normierung

Herzog[37] wirft Oser vor, dass er Standards „nicht nur als Maßstäbe für die Beurteilung und Evaluation von Kompetenzen" heranzieht, sondern Standards und Kompetenzen gleichsetzt. Zudem moniert Herzog eine zu enge Verwobenheit der von Oser formulierten vier Gütekriterien, die z.T. Ähnliches fordern, insofern diffus bleiben und als theoretische Absicherung für den Standardbegriff zu wenig leisten. Weiterhin erscheint dem Kritiker Osers Ansatz als theoretisch zu schwach, weil keine Gesamtsystematik darüber zu erkennen sei, wie es zu den Festlegungen der jeweiligen Standards gekommen ist. Der Verweis auf Expertengespräche und Metaanalysen erscheint Herzog als wenig aussagekräftig. Die Formulierung von Oser (2001, S. 217), ein Standard habe die „Bestätigung absoluter Notwendigkeit" bzw. die von Oser erteilte Zuschreibung zum Expertenstatus aus der rückwärts gehenden Logik heraus, in Befragungen formulierte Gütekriterien im Standardmodell zu berücksichtigen, erwecken bei Herzog den Anschein, dass hier der „ideale Lehrer" wiederauferstanden ist, wobei aus seiner Sicht trotzdem unklar bleibt, wie ein diesbezüglicher Geneseprozess auszusehen hat. Herzog wirft Oser ein fragwürdiges Professionsverständnis vor, weil letzter die Unterscheidung des Experten vom Laien

[37] Die sehr grundsätzliche Kritik von Herzog an dem Oserschen Standardkonzept wird hier nicht geteilt. Gleichwohl kann man einräumen, dass in dem Konzept z.T. Unschärfen auftauchen und der Begriff des „Standards" ein Eigenleben im Sinne eines ubiquitären Einsatzes anzunehmen scheint. Die theoretische und konzeptionelle Aufgeladenheit des Begriffes kann einerseits für den – zumindest schweizerischen Erfolg – des Konzeptes verantwortlich gemacht werden, andererseits steht es damit zugleich in Verdacht, an analytischer Schärfe zu verlieren. Verwendete Formulierungen – wie: „Standards werden transzendiert" (Oser 1997b, S. 226; Oser 2001, S. 331) – tragen ebenfalls nicht zur begrifflichen Präzisierung bei. In der analytischen Betrachtung von Beiträgen zur Standarddiskussion von Oser in chronologischer Reihenfolge ihres Veröffentlichungszeitpunktes kann man allerdings eine zunehmende Präzisierung feststellen, wofür z.B. die aktuelle Aussage von Oser, Curcio & Düggeli (2007, S. 14) spricht: „Sobald das Postulat der Messbarkeit und damit der Qualität konkretisiert werden, sprechen wir von Standards. Es geht dann um den Anspruch der Messung einer bestimmten Qualitätsausprägung, also darum, wie gut oder eben weniger gut eine Lehrperson beispielsweise Gruppenunterricht planen, durchführen und die Resultate in den weiteren Unterricht integrieren kann". Herausgestellt wird die Hauptdimension von Standards, nämlich die durch sie ermöglichte Messung von Kompetenzprofilen, die wiederum Aussagen über die Standarderreichung im Sinne von Expertiseentwicklung erlaubt. Für die Fassung anzustrebender Ziele der Lehrerausbildung erscheint mir der Kompetenzbegriff in seiner Dimensionalität hinreichend zu sein. Trotzdem findet in der eigenen Studie auch für solche Zielbeschreibungen der Standardbegriff Verwendung, einfach deshalb, weil der Terminus im deutschen Sprachgebrauch mittlerweile eine verbreitete Durchsetzung erfahren hat und einen allgemeinen Verständnishorizont bedient.

an der Beherrschung von Standards definitorisch festmacht. Das Spezifikum der Lehrerprofession sieht Herzog hingegen vielmehr in der Vermittlung von privatem und öffentlichem Leben, von Familie und Gesellschaft. Demgemäß stellen Schule und Unterricht intermediäre Bereiche mit einem hermeneutischen Charakter dar. Schließlich sieht Herzog eine Gefahr in der Normierung, die grundlegend mit der Einführung von Standardkonzepten verbunden ist, speziell dann, wenn sie tatsächlich auf der Ebene des Unterrichts in Erscheinung treten. Unterricht als komplexes System, so die kritische Position, ist von situativ unbestimmbaren Bedingungen geprägt, die auch durch definierte Standards keinem normierten Schema unterliegen.
Die Replik von Oser (2005, 2006) soll hier nicht detailliert dargelegt werden. Er befindet die Ausführungen von Herzog zwar grundsätzlich als interessant, sieht in ihnen aber keine grundlegende Kritik in Richtung der Infragestellung seines Konzepts und versucht alle Kritikpunkte zu widerlegen.
Grundlegend kann man sagen, dass hier zwei Deutungshorizonte aufeinanderprallen, die in jeweils unterschiedlichen Referenzsystemen und Argumentationslinien verankert sind, weshalb der Schlagabtausch insgesamt nicht als sehr konstruktiv bewertet werden kann.

Über einzelne Autorenpositionen hinaus können Kritikpunkte zu Standardkonzepten aus den Erfahrungen gezogen werden, die vorwiegend in den USA mit der Einführung von Standards in das Bildungssystem auf unterschiedlichen Ebenen gesammelt worden sind (Keller 2002). Sie unterscheiden sich bezüglich ihrer eher pragmatischen oder eher grundsätzlichen Ausrichtung, wenn diese Abgrenzung in der Literaturanalyse dabei teilweise auch nur unscharf auszumachen ist. Die Kritikpunkte beanspruchen zunächst einen länderübergreifenden Status, wobei länderspezifische Einschränkungen bzw. Besonderheiten bestehen, die etwa kulturellen Einfluss auf die Implementierung von Standardkonzepten ausüben und sich in jeweiligen Ausprägungen der Kritikpunkte abbilden (Tulodziecki & Grafe 2006).
Standards bergen aus dieser erfahrungsbezogenen Debatte heraus die Gefahr in sich einen bestimmten Status Quo der Gegenwart auch für die Zukunft festzuschreiben. Diese Gefahr ist zwar potentiell existent, lässt sich aber auf jede Theoriebildung übertragen. Begegnet werden kann ihr mit einer Antizipation notwendiger zukunftsbezogener Zielbereiche, so weit diese bestimmbar sind, und der prinzipiellen Entwicklung von Standardkonzepten als zeithistorische und jederzeit reformulierbare Fragmente.
Bei der Definition von Standards stellt sich zudem die Frage, welcher Personen- oder Institutionenkreis geeignet ist bzw. die notwendige Expertise mitbringt, um auch in legitimatorischer Perspektive Standardkonzepte abzusichern. In demokratischen Gesellschaften sollte es tendenziell selbstverständlich sein, alle betroffenen Akteure zu beteiligen. Tulodziecki (2004, S. 53) erkennt allerdings diesbezügliche Diskrepanzen zwischen einer wünschenswerten Beteiligung und der notwendig vorauszusetzenden und für die Standardentwicklung unverzichtbaren Expertise. Dies gilt umso mehr, wenn die Standards auf „psychologisch fundierten Kompetenzmodellen, auf geeigneten Theorien des Unterrichts, der Bildung und Erziehung sowie auf empirisch bewährten Annahmen beruhen sollen“ (Tulodziecki &

Grafe 2006, S. 42). Genauer bezeichnen Tulodziecki & Grafe das notwendige Profil zur Standardentwicklung, das aber einer demokratischen Einbeziehung aller Akteure entgegensteht, als das auch im demokratietheoretischen Diskurs diskutierte und Diskrepanzen implizierende Prinzip der Expertokratie. Demokratie und Expertokratie stehen insofern in einem Spannungsverhältnis zueinander, als dass Expertise mit der Definitionsmacht über Wissenslücken einhergeht, die dem Ideal eines prinzipiell herrschaftsfreien demokratischen Kommunikationsforums widerspricht. In eher pragmatischer Sichtweise führt Oelkers (2002, S. 4) dazu aus:

„Besonders die Lehrerbildung scheut Standards, einerseits weil sie wie eine ungewünschte Beschneidung der Freiheit wahrgenommen werden, andererseits weil sie als nicht ‚objektiv' erscheinen. Aber Standards sind einfach Festlegungen, die sachgemäß und plausibel sein müssen, ohne allen Beteiligten zu gefallen. Sie können entwickelt und revidiert werden, aber dürfen nicht warten, bis auch der Letzte zustimmt und dann über ihre ‚Objektivität' entschieden wäre."

Weitere grundlegende Kritikpunkte beziehen sich auf die mit einer Standardimplementation verbundene vermeintliche Technokratisierung und Ökonomisierung der Lehrerbildung (Gruschka et al. 2005). Die intendierte Umsteuerung der Bildungssysteme hin zu einer Outputorientierung wird hier grundsätzlich mit der Begründung in Frage gestellt, die einseitige Fixierung auf die Erreichung von Standards würde den Blick dafür verstellen, dass Bildungsziele oftmals nur über selbstinitiierte Umwege erreicht werden können. Die Outputorientierung erfährt eine Gleichsetzung mit der ökonomischen Kosten-Nutzen-Maximierung, worin die Kritiker kein adäquates Modell für die Lehrerbildung entdecken können, werde damit doch der Versuch unternommen, Lernprozesse zu standardisieren, die aber gerade nicht standardisierbar seien. Zudem findet sich der kritische Hinweis, dass die durch den Konstruktivismus indizierte Vorstellung einer Individualisierung von schülerbezogenen Lernprozessen ebenso für angehende Lehrer gelten müsse; ein an sich richtiger Sachverhalt, der jedoch hier eine instrumentelle Verwendung findet, indem das Gegensatzpaar „Individualisierung" vs. „Standarderreichung" konstruiert wird.
Sprechen die Kritiker der Standardisierung ein Gefährdungspotential zu, können sie aber nicht schlüssig aufzeigen, inwiefern die „traditionelle" Lehrerbildung der Individualisierung von Lernprozessen gerecht geworden ist. So ist der Zusammenhang zwischen einer Veränderung von Systemmerkmalen und ihren Auswirkungen auf der Individualseite für die Lehrerbildung bislang nur in Ansätzen geklärt (Blömeke 2004). Zurzeit führt die IEA (International Association for the Evaluation of Educational Achievement) eine internationale Vergleichsstudie mit dem Titel „Teacher Education and Development Study. Learning to Teach Mathematics (TEDS-M)" durch, an der auch Deutschland teilnimmt. Auf die Ergebnisse darf man gespannt sein (vgl. Blömeke 2006a). Ähnliches gilt für die deutsch-schweizerische Studie „Standarderreichung in der Lehrerinnen- und Lehrerbildung: Analyse der berufsfeldorientierten Ausbildung", an der ich mitarbeite. Konsistente Ergebnisse, die auf qualitative Unterschiede zwischen einer standardbasierten Ausbildung im Vergleich zu einer traditionellen Ausbildung hindeuten, liegen erst in Ansätzen vor (vgl. Baer et al. 2007).

Tulodziecki & Grafe (2006) führen noch weitere Kritikpunkte gegen die Implementation von Standardkonzepten an. Resümierend kann man aber sagen, dass die internationale Ausrichtung der Lehrerbildung hin zu einer Orientierung an Standards einen mittlerweile lang anhaltenden Trend darstellt, dem sich auch die deutsche Lehrerbildung zumindest nicht komplett verschließen kann. Damit sind nicht zu unterschätzende Problemlagen verbunden, die aber keine prinzipielle Gegenposition fundieren. Von der KMK (2004) wurden bereits Standards formuliert, die aber bislang nur wenig und in vereinzelten Initiativen Einfluss auf die deutsche Lehrerbildung ausüben. Angesicht gewählter vager Formulierung ist auch nicht zu erwarten, dass hier in Zukunft ein grundlegender Richtungswandel zu verzeichnen sein wird (Blömeke 2006b).
Die Entwicklung von Standards für die Lehrerbildung beinhaltet das Potential für die deutsche Erziehungswissenschaft, ihre theoretisch und empirisch gewonnenen Erkenntnisse in eine Form zu gießen, die einem bundesweit einheitlichen Ausbildungskonzept zugrunde liegt.[38] Geäußerte Kritikpunkte sind insofern ernst zu nehmen, als dass die Entwicklung von Standards für die Lehrerbildung nicht per se zu einer besseren Ausbildung und damit einer höheren Unterrichtsqualität führt, von der Schüler profitieren. Die appellative Setzung eines Outputs in Form von Standards gibt keinen Garant für seine Erzielung ab. Hingegen sind Studierende des Lehramtes in komplexen Lernprozessen dazu herausgefordert, eine Standarderreichung anzustreben.
Gleichzeitig kann man schon jetzt festhalten, dass sowohl die Hochschulbildung für die angehenden Lehrer als auch das soziale Feld Schule, das Schüler jeden Tag neu vorfinden, bereits Formen von Standards aufweisen, die zum Teil lediglich durch andere Bezeichnungen etikettiert sind. Oelkers (2004a, S. 16) konstatiert in einem historischen Abriss zur Entwicklung von Standards:

„Der historische Befund lässt sich in dieser Hinsicht so zusammenfassen. Wenn mit dem Ausdruck ‚Standards' niveaudifferenzierte Vorgaben gemeint sind, die überall angestrebt werden sollen und nirgendwo fehlen dürfen, dann ist die Schule voll von Standards. Das Problem ist nicht, dass es keine Standards gibt, sondern dass sie unverbindlich sind, unpräzise formuliert werden und keine wirklichen Kontrollen verlangen. Standards beziehen sich auf die Qualität der Lehrkräfte, also die Lehrerbildung, die Inhalte des Unterrichts, also die Lehrmittel, sowie die Klassen des Systems, also die Schulorganisation. Standards sind daher nicht einfach präzisere Fassungen der Lehrpläne, sondern Instrumente für das Gesamtsystem, die nicht neu erfunden, sondern für ein System entwickelt werden müssen, das darauf nicht wartet."

[38] Die damit verbundenen Reformbemühungen wurden schon 2002 kritisch von Luhmann gewürdigt, indem er konstatiert: „Beobachtet man das jeweils reformierte System, hat man den Eindruck, dass das Hauptresultat von Reformen die Erzeugung des Bedarfs für weitere Reformen ist. Reformen wären demnach sich selbst generierende Programme für die Veränderung der Strukturen des Systems" (Luhmann 2002, S. 18). Weiterhin weist er darauf hin, dass Zielsetzungen von Reformen häufig in der Vergangenheit schon einmal angestrebt wurden und gescheitert sind, woraus er schließt: „Die wichtigste Ressource der Reformer scheint daher eine Leistung des Systemgedächtnis zu sein, nämlich das Vergessen" (ebd.). Bezüglich der Einführung einer standardbasierten Umsteuerung der Lehrerbildung im Sinne einer kontrollierbaren Outputorientierung wird hier die These vertreten, dass es sich tatsächlich um ein Novum für Deutschland handelt, das nicht mit vergangenheitlichen Anstrengungen verglichen werden kann.

Folgt man der Lesart von Oelkers, so können durch die Formulierung von Standards Transparenzpotentiale für die Lehrerbildung erschlossen werden, die für ihre Absolventen eine wichtige Orientierungsfunktion haben und die Ausbildung darüber hinaus in einer Art hinterfragbar machen, wie das zum jetzigen Zeitpunkt nicht möglich ist. Larcher & Oelkers (2004, S. 135f.) betonen in diesem Zusammenhang die Funktion von Standards als Steuerungselemente der Qualitätssicherung für die Lehrerbildung und konstatieren, dass nur in Deutschland[39] eine „generelle Skepsis gegenüber neuen Instrumenten der Qualitätssicherung" festzustellen ist, während umgekehrt „niemand den gravierenden Qualitätsschwund der deutschen Lehrerbildung bestreitet". Damit ist das Spannungsverhältnis der Standarddiskussion in Deutschland mit Bezug auf nationale und internationale Erkenntnisbefunde kritisch beschrieben. Während viele Erziehungswissenschaftler einer Einführung von Standards skeptisch gegenüberstehen, können sie aber nur wenig plausibel die Vorteile der bestehenden Lehrerausbildung herausstellen.

[39] Diese pauschalisierende Feststellung kann wenig Anspruch auf Angemessenheit beanspruchen. Skepsis gegenüber den Effekten einer Standardisierung auf allen Ebenen des Schulsystems gibt es etwa auch in den USA, wie die Darstellung von Mintrop (2006) zeigt. Seine Ausführungen dokumentieren zudem, dass eine Vorsicht gegenüber einer übertriebenen Standardisierung durchaus angebracht ist, besonders wenn diese funktional an spezifische Interessen der Politik und Wirtschaft gekoppelt ist. Larcher und Oelkers begründen eine grundlegende Skepsis in Deutschland mit der „frustrierenden Struktur der (hiesigen) Lehrerbildung".

6 Spezifika des Lernens in Praxislernsituationen des Lehrerberufes

Das Theorie-Praxis-Verhältnis betrifft einen zentralen Diskussionspunkt innerhalb der Lehrerbildung und stellt sich häufig als Problemverhältnis dar. Wenn auch der Praxisbezug als Norm kommuniziert wird und dabei „auf merkwürdige Weise unstrittig“ (Oelkers 1999, S. 2) erscheint, so gibt es doch vergleichsweise wenig gesichertes Wissen darüber, was er im Lehramtsstudium tatsächlich leistet bzw. es bestehen verschiedene Konzeptionen davon, wie die Sequenzierung von Theorie- und Praxisphasen angelegt werden sollte. Praxisbezug ist dabei ein unscharfer Begriff und kann unterschiedlichste Dimensionen beinhalten, wie Bekanntmachung mit dem schulischen Feld, Berufsvorbereitung oder Überprüfung der Berufseignung. In jedem Fall soll der Praxisbezug das Hochschulstudium begleiten, die Inhalte bündeln und überprüfbar machen. Dabei können zwei Lesarten expliziert werden. Zum einen steht der Praxisbezug für eine Aufwertung des Studiums, wobei der Praxis mit der Begründung eine vergleichsweise höhere Dignität eingeräumt wird, dass ansonsten das Studium nur „Elfenbeinturmwissen“ generieren würde. Die zweite und häufigere Lesart lautet, wissenschaftliches Theoriewissen solle die reflektierte Handlungsfähigkeit ihrer Absolventen im späteren Berufsfeld gewährleisten, womit das Studium eine Ausrichtung auf die spätere Praxis hin erfährt. Die Praxis wird zum Bewährungsfeld der Theorie.
Grundsätzlich stellt sich aber die Frage, inwieweit zwischen Theorie und Praxis unterschieden werden kann, denn jede Theorie hat Praxis und jede Praxis Theorie. Die von amtierenden Lehrern vorgefundene und konstruierte Praxis ist von Theorien durchdrungen, die nicht unbedingt den Status von wissenschaftlichen Theorien beanspruchen können, aber funktional in Form von Alltagstheorien konvergieren. Setzt man Deskription, Präskription, kausale Attribuierung und Komplexitätsreduktion als die wichtigsten funktionalen Kennzeichen von wissenschaftlichen und impliziten Theorien voraus, so verdeutlicht sich der ähnliche epistemische Status. Umgekehrt ist die sogenannte Theorie der Hochschule, abhängig von der Perspektiveneinnahme, ebenso in einem Praxisfeld verwurzelt, das sich durch die Bereiche Forschung und Lehre konstituiert. Zudem ist es ein Spezifikum der pädagogischen Hochschulausbildung, dass ihre „theoretischen“ Veranstaltungsformen Vorlesung und Seminar strukturell der unterrichtlichen Praxis an Schulen entsprechen, für die sie theoretisches Wissen bereitstellen wollen.
Damit ist offensichtlich, warum die grundsätzliche Unterscheidung zwischen Theorie und Praxis auch als ein „Missverständnis“ gedeutet werden kann (Nakamura, Böckelmann & Tröhler 2006). Als symptomatisch erweist sich die beobachterabhängige Beschreibung der Theorie aus einer Praxisperspektive bzw. der Praxis aus einer Theorieperspektive heraus. Diesbezüglich lässt sich feststellen, dass alltagstheoretische Praxispositionen wissenschaftlichen Theorien skeptisch gegenüber stehen, während die Wissenschaft, so sie sich denn überhaupt empirisch mit der Praxis auseinandersetzt, häufig gleichfalls zu eher skeptischen Positionierungen mit Bezug auf Alltagstheorien gelangt. Diese wechselseitig vorbehaltlichen Kenntnisnahmen führen dazu, dass auf Seiten der Praxis ein nur geringes Interesse an wissenschaftlicher Theoriebildung besteht und die Sozialwissenschaften ihr Desinteresse an der Praxis wiederum mit der angeblich

unumkehrbaren Trennung zwischen Theorie und Praxis festschreiben. Dabei bleibt oftmals unberücksichtigt, dass auch die Gewinnung wissenschaftlichen Wissens in besagten Praxiskontexten verortet ist. So ist es zwar theoretisch und empirisch evident, dass ein Unterschied zwischen wissenschaftlichem Wissen und der empirischen Anschlussfähigkeit dieses Wissens an Anforderungen der Praxis besteht. Gleichwohl bringen aber einzelne Personen wissenschaftliches Wissen hervor, welches wiederum in diskursiven Prozessen innerhalb der sog. „science community" eine Absicherung und Festigung erfährt. Auch diese Prozesse konstituieren sich wesentlich in Praxisvollzügen, mit den von den Sozialwissenschaften analysierten Eigendynamiken als selbstreferentielle Automatismen.
Jenseits dieser hochkomplexen Zusammenhänge von Theorie und Praxis muss eine Studie, die sich mit der Kompetenzentwicklung von Studierenden des Lehramtes in verlängerten Praxislernsituationen beschäftigt, die Frage beantworten, welche Potentiale und Hemmnisse in entsprechenden hochschulorganisatorischen Konzeptionen auszumachen sind. Im Weiteren soll es deshalb um die Bedeutung des Erfahrungswissens für die Herausbildung professioneller unterrichtlicher Kompetenzen sowie um die Frage gehen, inwieweit Expertise durch das unterrichtliche Handeln in Praxislernsituationen generiert werden kann. Zu diskutieren ist schließlich die Rolle des Kontingenztheorems als Voraussetzung und Problembeschreibung für die Bewältigung von Praxislernsituationen.

6.1 Erfahrungswissen

Als Folge der durch Brezinka (1971) angeregten Diskussion darüber, inwieweit sich die Pädagogik zur Erziehungswissenschaft entwickeln kann,[40] womit wesentlich die Frage berührt wird, „welche Voraussetzungen erfüllt sein sollen, damit den Verfahren und Ergebnissen erziehungswissenschaftlicher Aktivitäten das Prädikat ‚Wissenschaft' zuerkannt werden darf" (Pollak & Heid 1994, S. 1), wurde das Erfahrungswissen zunehmend als unzulänglich und praxeologisch gebrandmarkt und sollte durch ein erziehungswissenschaftlich begründbares Fachwissen ersetzt werden. Umgekehrt lassen sich seit Mitte der 1980er Jahre Diskurse ausmachen, die Prozesse der Verwissenschaftlichung kritisch nachzeichnen, dabei aber nicht eine Rückwendung in vorwissenschaftliche Zeiten anstreben, sondern Korrekturen und Ergänzungen hinsichtlich der Anwendungstiefe wissenschaftlichen Wissens für spezielle Anforderungs- und Erfahrungsbereiche der Praxis aufzeigen (Böhle 2005). Weniger geht es um eine Rangfolge von wissenschaftlichem und Erfahrungswissen, sondern vielmehr um das Kenntlichmachen von jeweiligen Grenzen und damit auch die Wiederannäherung der Wissenstypen sowie eine diesbezügliche Ergänzung bis dahin vorherrschender Perspektiven.
Erfahrungswissen kann nur im praktischen Handeln erworben werden und ist somit individuell und situativ verankert. Böhle (ebd.) diskutiert vier mögliche Ursachen, die

[40] Pollak & Heid (1994, S. 1) machen darauf aufmerksam, dass diese Entwicklungslinie nicht als linear betrachtet werden darf, sondern prinzipiell – ohne normative Wertung – reversibel bleibt, was in ihrem Buchtitel „Von der Erziehungswissenschaft zur Pädagogik" zum Ausdruck kommt.

für die Wiederentdeckung des Erfahrungswissens im wissenschaftlichen Diskurs verantwortlich sind:

- Die Konfrontation mit unerwarteten Grenzen der Technisierung im Arbeitsbereich
- Die Beschleunigung des Wandels von Arbeitsanforderungen
- Die Erkenntnis, dass sich wissenschaftlich begründetes Fachwissen nicht umstandslos in praktisches Arbeitshandeln umsetzen lässt
- Die Auseinandersetzung mit nicht institutionalisierten informellen Lernprozessen, speziell in der Thematisierung des lebenslangen Lernens

Der erste Punkt verweist auf die Komplexität technischer Systeme, deren Funktionsfähigkeit z.T. von so vielen Parametern bestimmt wird, dass sie nicht mehr in technischen und empirischen Modellen abgebildet oder simuliert werden können. Zudem funktionieren hochkomplexe Systeme evtl. zwar unter Laborbedingungen in gewünschter Weise. Wenn sich bei der Implementation externe Störgrößen auf Systemmerkmale auswirken, kommt es aber zu Veränderungen des Gesamtsystems bis hin zum systemischen Zusammenbruch. Im günstigen Fall kann Erfahrungswissen das Aufkeimen solcher Störfälle verhindern, indem Unregelmäßigkeiten frühzeitig erkannt und behoben werden. Nicht zuletzt zeugen verschiedenste weltgeschichtliche Katastrophen aber von der Begrenztheit menschlicher Interaktionspotentiale innerhalb komplexer Situationsanforderungen. Schulze (2001) hebt die Bedeutung des Erfahrungswissens besonders für die Bewältigung solcher kritischen Situationen hervor, in denen komplexe und z.T. nicht vorhersehbare Einzelmerkmale unter Zeitdruck miteinander in Verbindung gebracht werden müssen.
Weiterhin führt ein beschleunigter Wandel von Arbeitsanforderungen (Punkt 2) dazu, dass in schulischen und hochschulischen Bildungsgängen nicht immer in ausreichendem Maße die Anforderungsqualitäten zukünftiger beruflicher Arbeitserfordernisse zuverlässig prognostiziert werden können. Insofern wird das Erfahrungswissen zur wichtigen Humanressource, weil ggf. alternatives Wissen, als Innovierungswissen, nicht zur Verfügung steht.
Die empirische Anschlussfähigkeit wissenschaftlichen Wissens für die Bearbeitung komplexer praktischer Herausforderungen wurde in dieser Studie bereits breit diskutiert (Punkt 3). Die Debatte um das Erfahrungswissen zeigt nun übergreifend, dass wissenschaftliches Wissen vor dem Hintergrund von erfahrungsbezogenem Wissen rekontextualisiert werden muss, weil Systematik und Typologie wissenschaftlichen Wissens nicht den Dimensionen von konkreten Anwendungserfordernissen entsprechen. Trotz einer gewissen Evidenz sei damit aber nicht auf die selbstreferentielle Unterscheidung zwischen wissenschaftlichem und praktischem Wissen verwiesen, wie sie in der professionssoziologischen Wissensverwendungsforschung zementiert wurde. „Erfahrung und Erfahrungswissen werden hier … als Grundlagen eines ‚Arbeitsprozesswissens' verstanden und sowohl als Voraussetzung wie auch als Ergebnis der Verbindung von wissenschaftlich begründetem Wissen mit praktischem Handeln gesehen" (Böhle 2005).

Der vierte Punkt hebt schließlich die Notwendigkeit einer lebenslangen Lernbereitschaft hervor,[41] um den in Veränderung begriffenen Arbeitsplatzbedingungen und -anforderungen gerecht werden zu können. Entsprechende Lernprozesse können nur in Anteilen institutionalisiert werden, womit der Blick auf informelle Lernprozesse fällt, die innerhalb und außerhalb beruflicher Tätigkeit ablaufen. Im Zusammenhang mit informellen Lernprozessen wird häufig der Begriff „implizites Wissen" gebraucht, der wiederum teilweise synonym zum Begriff Erfahrungswissen Verwendung findet (Neuweg 2005, 2006). Während allerdings das implizite Wissen nicht zwingend als Ergebnis von Praxisbezügen generiert wird (Neuweg spricht in Anlehnung an Polanyis (1985) Konzept des Tacit Knowledge „Vom Schweigen der Könner"), ist der Praxisbezug für das Erfahrungswissen konstitutiv. Im Alltagsverständnis sowie in wissenschaftlichen Diskursen markiert dieser Differenzierungsansatz den Unterschied zwischen Erfahrungswissen und wissenschaftlich begründetem Wissen.

Gruber & Rehrl (2005, S. 9) halten es für einen Fehlschluss, wissenschaftliches Wissen so didaktisch aufbereiten zu können, dass ein direkter Transfer wissenschaftlicher Theorien auf Anforderungen des Praxisfeldes umstandslos möglich ist. Hier soll aber die These vertreten werden, dass diese Unterscheidung keine ausreichende Begründung darstellt. Speziell im Zuschnitt der modernen Wissenschaften, die sich im Falle der Naturwissenschaften bereits traditionell empirisch begründen, kann auch die Entwicklung der Geisteswissenschaften als Bewegung hin zu einer stärkeren Einbeziehung von empirischen Bezügen beschrieben werden. Wenn sich, möglicherweise mit Ausnahme der Theologie und Philosophie, auch die Geisteswissenschaften einer empirischen Orientierung nicht entziehen können, sind damit Fragen der kriteriellen Nützlichkeitserwägung verbunden.

In Anlehnung an Böhle (2005, 2006, 2007) kristallisiert sich als Spezifikum des Erfahrungswissens heraus, dass es zwar individuell und situationsbezogen, nicht aber im Rahmen methodisch geleiteter Überprüfungen und also zweckgebunden erworben wird. Grundlegende Voraussetzung für die Generierung von Erfahrungswissen wird in dieser Konzeption damit die Ermöglichung von hohen Freiheitsgraden in der Wahrnehmung und Handlung, wodurch Erfahrungen erst als Basis für die Entwicklung von Erfahrungswissen individuell genutzt werden können.[42]

Im Kontext der Diskussion um die Reichweite des Erfahrungswissens wird auch deutlich, dass dessen Erwerb und Anwendung in berufspraktischen Situationen nicht allein den Zielsetzungen zweckrationalen Handelns folgt, sondern wesentlich über subjektive Wahrnehmungen und damit verbundene Emotionen gesteuert wird. Hierin

[41] Für die Erwachsenenbildung ist der Begriff des lebenslangen Lernens konstitutiv geworden (Brödel 1998). Kritisch kann man einwenden, dass mit der Ausweitung von notwendigen Lernprozessen auf die gesamte Lebensspanne Bildung eine „lebenslängliche" Aufgabe bedeutet. In diesem Zusammenhang sprechen Erpenbeck & Sauer (2001) von der „Entgrenzung des Lernens".

[42] Vor dem Hintergrund dieses Sachverhalts wird deutlich, dass die in den Pädagogischen Hochschulen Baden-Württembergs installierten Tagespraktika wenig Potential zur Entwicklung von Erfahrungswissen bieten, weil sie, als institutionalisierte Beobachtungssituationen, nur in Anteilen eine zweckfreie Erprobung erlauben. Hierin kann ein Vorteil von verlängerten Praxisphasen gesehen werden, wenn auch noch unklar bleibt, über welche Zeitspannen solche Praxisphasen angelegt werden müssen, um nachhaltiges Erfahrungswissen generieren zu können. Diese Frage wird auch in der vorliegenden Studie nicht endgültig beantwortet und kann angesichts der abhängigen Variablen individueller Voraussetzungen von Studierenden vermutlich nicht allgemeinverbindlich gelöst werden.

liegen zum einen Chancen, im erfahrungsbasierten Modus über mentale und emotionale Prozesse einer Anforderungssituation gerecht zu werden. Experten entwickeln ein „Gespür“ für die Mikrobereiche einer Situation und werden ihrer Gesamtdynamik potentiell besser als in einer wissenschaftlichen Annährung gerecht. Gleichzeitig liegen hier die Gefahren einer Routinisierung von Wissen und Handeln, die sich weder von speziellen Situationsmerkmalen und -anforderungen, noch von wissenschaftlich orientierter Reflexion leiten lässt. Wenn ein intuitiver Zugang situativ Vorteile hinsichtlich der Problembewältigung verspricht, führt er doch insgesamt ein hohes Fehlerpotential mit sich und ist auf Korrektive angewiesen, will er nicht Gefahr laufen, zu einem selbstreferentiellen Mechanismus zu werden.
Ein weiteres Spezifikum des Erfahrungswissens betrifft die Abhängigkeit von sozialen Akteuren, deren Handlungen oder Reaktionen auf das Handeln anderer nur bruchstückhaft antizipierbar sind und auch in wissenschaftlichen Modellen nur ansatzweise beschrieben werden können, ohne gleichwohl einen Geltungsanspruch für die jeweilige Praxissituation beanspruchen zu können.
Erfahrungswissen speist sich zu großen Teilen aus gemachten Beobachtungen (Rose 1992), die sich auf das eigene Handeln oder auf das Handeln anderer Akteure (Hospitationen) beziehen. Insbesondere beim Aufbau von Erfahrungswissen über Beobachtungen des eigenen Handelns ist mit Verzerrungen durch Eigenbetroffenheit zu rechnen (Hascher 2005, S. 44). In ähnlicher wenn auch abgeschwächter Weise besteht dieses Problem auch für Hospitationen. Als Ausweg aus „Erfahrungsfallen“ können Kooperationen und Erfahrungsaustausch angeführt werden, die interaktiv regulierte Reflexionsprozesse mit Bezug auf situative Beobachtungskontexte ermöglichen. Reflexivität wird damit zur Schlüsselkompetenz im professionellen Lehrerhandeln (Combe & Kolbe 2004, S. 835). Sie beinhaltet die Fähigkeit zur aktiven Distanzierung und erlaubt so ein bewusstes Heraustreten aus der unmittelbaren Situation von Lebensvollzügen (Häcker & Rihm 2005, S. 362).
Es besteht kein allgemeingültiges Reflexionsverständnis und man kann von einem unscharfen Begriff sprechen. Hier soll er aber weit gefasst als eine denkende Betrachtung von Merkmalen, Interaktionen und Beziehungen verstanden werden. Im Sinne der lateinischen Wurzel[43] blickt das Individuum in der Reflexion auf das Vergangene zurück und rekonstruiert es im Denken (Wildt 2003, S. 73). Avanciert zum pädagogischen Topos werden die Möglichkeiten reflexiver Prozesse aber nicht selten überschätzt. Ciompi (1988, 1997) macht auf die Eingewobenheit von Reflexionsprozessen in biographisch erworbene affektiv-kognitive Bezugssysteme aufmerksam. Individuelle Reflexionsprozesse können immer nur aus der Perspektive des betroffenen Subjekts, nicht von einem objektiven Standpunkt aus erfolgen. Deshalb sind Austauschprozesse als soziales Korrektiv umso wichtiger. Für das Berufsfeld Schule kann man in diesem Zusammenhang die plausible These aufstellen, dass genau diese Austauschprozesse selten vorzufinden sind, während ein gefördertes Einzelgängertum überwiegt.
Reflexionen sollen die individuellen Handlungsmöglichkeiten erweitern oder verbessern. Damit wird die Herausbildung reflexiver Potentiale zu einer Aufgabe der Lehrerbildung, wobei sich eine zunehmende Professionalität als Steigerung von Reflexivität ausdrückt (Reh 2004, S. 369). Vor dem Hintergrund der Komplexität

[43] Der Begriff Reflexion leitet sich aus dem lateinischen Verb „reflectere“ ab und bedeutet übersetzt „zurückbeugen“, „wenden“, „umkehren“ (Kluge 2002).

unterrichtlicher Prozesse stellt sich nun die Frage, welche Merkmale vorrangig kritisch reflektiert werden sollen bzw. inwieweit etwaige Ergebnisse für das didaktische Handeln fruchtbar gemacht werden können. Ziele von Reflexionsprozessen betreffen bspw. Einsichten darüber, wie sich im Stundenverlauf auftuende Diskrepanzen zwischen intendierten Handlungszielen und tatsächlichen Handlungsvollzügen beschreiben und erklären lassen. Dafür ist jedoch vorauszusetzen, dass Betroffene derartige Diskrepanzerfahrungen überhaupt wahrnehmen und thematisieren wollen/können. Da mit der Offenlegung notwendigerweise eine Infragestellung der eigenen Handlungsentwürfe verbunden ist, kann nicht zwingend davon ausgegangen werden, dass Praxiserfahrungen per se Reflexionspotentialen zugute kommen.
Die Diskussion um die prinzipielle Verschiedenheit von wissenschaftlich abgesichertem Wissen und Erfahrungswissen schließt an die Grundsatzdiskussion zum vermeintlichen Gegensatzpaar Theorie vs. Praxis an. Nakamura et al. (2006) sehen ein „Missverständnis“ darin, Theorie und Praxis als zwei verschiedene Wissensformen zu fassen, weil diese sich im Kern nicht aufspalten lassen und vielmehr zwei Dimensionen menschlichen Handelns beschreiben würden. Die Diskussion verkompliziert sich durch oftmals unklar zugrunde gelegte Theorie- und Praxisverständnisse. Die prototypische Modellierung der Kategorien „Denken“ bzw. Wissen“ und „Handeln“ wirkt häufig schablonenhaft. Das von Watzlawick et al. (1990) herausgestellte kommunikationstheoretische Axiom, man kann nicht „nicht“ kommunizieren, lässt sich ebenso an ein Praxisverständnis anlegen. Der Mensch kann nicht „nicht“ in der Praxis sein. In einer Annäherung an das zu bestimmende Verhältnis von Theorie und Praxis stellt sich grundlegend die Frage, was denn die Theorie der Praxis bzw. die Praxis der Theorie sein könnte. Da die eigene Studie aber keinem erkenntnisphilosophischen Interesse folgt, soll die Frage in Richtung des Aufspürens funktionaler Äquivalenzen angegangen werden.
Was also kann Praxis, hier die Funktion verlängerter Praktikumsphasen, für die Theorie leisten und umgekehrt, was kann Theorie, hier die wissenschaftliche Lehrerausbildung an einer Pädagogischen Hochschule in Baden-Württemberg, für die Praxis leisten? Die Funktionen von Theorien lassen sich abstrahierend für drei Dimensionen beschreiben:

- Theorien erlauben es Phänomene zu beschreiben (deskriptive Funktion)
- Theorien erlauben z.T. Vorhersagen (präskriptive Funktion)
- Theorien können im besten Falle Wenn-dann-Zusammenhänge erklären (kausal-attribuierende Funktion)

Für das komplexe Handlungsfeld „Unterricht“ beschreiben wissenschaftliche Theorien in den meisten Fällen allerdings nur Wahrscheinlichkeiten für Handlungszusammenhänge in der Lehrer-Schüler-Interaktion,[44] die unterschiedliche Relevanzen beanspruchen. Damit wird ein Grundproblem im Umgang mit pädagogischen Theorien deutlich. Auf der einen Seite wollen sie der Komplexität unterrichtlicher Prozesse abstrahierend gerecht werden. Sie müssen dabei viele

[44] Klare Zusammenhänge können nur für Extremwerte ausgewiesen werden, wobei es zu banalen Ergebnissen kommt. „Wenn Du sehr unverständlich sprichst, ist die Wahrscheinlichkeit groß, dass dich keiner versteht“, ist eine Information von geringem Erkenntniswert.

Einzelinformationen berücksichtigen und miteinander verschränken, was veränderbare Abstraktionsniveaus voraussetzt. Es ist ein traditionsreiches Merkmal der Geisteswissenschaften, die Güte ihrer Theorien an der Höhe des Abstraktionsniveaus festzumachen. Auf der anderen Seite ist es daher problematisch, dass Theorien ihre empirische Anschlussfähigkeit weitgehend verlieren, wenn das Komplexitäts- und Abstraktionsniveau zu hoch ansetzt. Transfer- und handlungsrelevante Hinweise können dann nicht mehr gegeben werden.
Ob ein hohes Abstraktionsniveau für Reflexionsprozesse von Praxiserfahrungen konstruktiv genutzt werden kann, darf angesichts der geringen Verschränkung wissenschaftlicher Theoriebildung mit praktischer Anforderungsbewältigung teilweise bezweifelt werden. Gruber und Rehrl (2005, S. 9f.) führen in Bezugnahme auf Patry (2005) thesenartig zwei Gründe an, warum die Nützlichkeit von Theorien für die Praxis häufig allgemein gering eingeschätzt wird. Zum einen nennen sie das „Unangemessenheitsargument" im Sinne der o.g. geringen Erklärungsdichte, die Theorien für das Handeln in der Praxis angeblich haben. Gruber und Rehrl (ebd.) weisen aber in meiner Sicht zu Recht darauf hin, dass Theorien notwendigerweise allgemein formuliert sind und damit nicht einzelne Handlungssituationen, sondern Handlungsklassen beschreiben, deren situative Übertragung immer erst individuell geleistet werden muss, allerdings prinzipiell geleistet werden kann, wenn Wissen als Erfahrungswissen konzeptualisiert wird. Damit sind hohe Anforderungen an die Lehrerausbildung umschrieben und die Integration von Praxisphasen erlangt den Status eines konstitutiven Elements.
Zum anderen verweisen die Autoren auf das sog. „Trivialitätsargument", nach dem erziehungswissenschaftliche Theorien immer nur im Nachhinein das beschreiben, was den Praktikern des Berufsfeldes ohnehin schon hinlänglich bekannt ist. Dieses Argument trifft teilweise zu, gilt allerdings m.E. vor allem für den deskriptiven Funktionsaspekt von Theorien.[45]
Zurückkommend auf die Bedeutung des Erfahrungswissens in pädagogischen Kontexten sei auf den Stellenwert des „pädagogischen Wissens" nach Nakamura et al. (2006, S. 13) verwiesen, die damit das Wissen bezeichnen, welches Schülern vermittelt werden soll. Pädagogisches Wissen und Erfahrungswissen beeinflussen sich wechselseitig. Im Gegensatz zu dem an Hochschulen vermittelten „wissenschaftlichen Wissen" nimmt das pädagogische Wissen einen konstitutiven Stellenwert für das Lehrer-Schüler-Interaktionsverhältnis ein, indem es auf Aufgabenbereiche von Lehrpersonen zielt und insofern als ein Konglomerat von wissenschaftlichem Wissen und Erfahrungswissen bezeichnet werden kann. Formen der didaktischen Reduktion transformieren dabei wissenschaftliches Wissen in pädagogisches Wissen (Jank & Meyer 2005, S. 81; Köck 2005, S. 146). Letztes zeichnet sich durch folgende Charakteristika aus:

- Es bezieht sich nur auf Ausschnitte (Exemplarisches Wissen);
- vereinfacht das zu vermittelnde Wissen;
- setzt es in einen möglichst logisch erscheinenden Aufbau;

[45] Es zeichnet Wissenschaft aus, Praxisfelder vor dem Hintergrund theoretisch fundierter Sprache, Methodik und Modelle hermeneutisch zu analysieren. Die diesbezüglichen deskriptiven Ergebnisse beschreiben selbstverständlich Merkmale und Dimensionen der Praxis, die auch den Praktikern bekannt sind oder zumindest bekannt sein können.

- vermeidet Widersprüche und Ausnahmen;
- macht sich mit Hilfsmitteln attraktiv (Dias, Experimente, Filme, Wandbilder);
- wird möglichst effizient vermittelt.

Merkmale pädagogischen Wissens liegen damit in der Selektion, Komprimierung, Komposition, im didaktischen Arrangement und Engagement sowie in der Effizienz. Die Prinzipien transformieren wissenschaftliches Wissen in pädagogisches Wissen. Da dem Wissen Bedeutung als Lehrgegenstand zukommt und hier Widersprüche sowie Ausnahmen möglichst vermieden werden sollen, weist es tendenziell einen dogmatischen Charakter auf. Wissenschaftliches Wissen wird dagegen per se als methodisch gewonnenes und damit transparentes Wissen betrachtet, welches als Untersuchungsgegenstand jederzeit einem reversiblen Reflektionsprozess unterliegt bzw. dessen Funktion sich genau durch diesen Prozess legitimiert (Drerup 1987; Heid & Harteis 2005).
Anknüpfend an die Ausführungen zum pädagogischen bzw. wissenschaftlichen Wissen stellt das Verhältnis von Theorie und Praxis kein konträres dar, hingegen wurzelt jede Theorie in der Praxis und bezieht sich gleichzeitig auf sie. Nach Perrig (2006, S. 65) stehen wissenschaftliche Theorien für Propositionsgefüge, die erst dann integrative und fruchtbare Wirkung entfalten können, wenn sie Anschluss an den Erfahrungs- und Erlebnisbereich des Lerners finden. Dagegen speisen sich naive Erkenntnisse oder naive Theoriebildungen ausschließlich über unmittelbare Erfahrungen oder Erinnerungen.
In der Diskussion um die Bedeutung des Erfahrungswissens für eine korrekte und sachadäquate mentale Repräsentation der physikalischen Umwelt (auf die der handelnde Akteur Einfluss nimmt bzw. von der er beeinflusst wird) als Grundlage des Handelns ist die Vorstellung eines bestehenden Kontinuums von reflexiv automatisierten, eher vorbewussten Repräsentationen bis hin zu bewussten Erlebnis- und Erkenntnisstrukturen verbreitet. Perrig (2006, S. 68) unterscheidet für das menschliche Handeln folgende Lernerfahrungen, die in der Chronologie ihrer Reihenfolge ein entsprechendes Kontinuum abbilden:

- automatische Reaktionen
- automatisierte Lernerfahrungen
- unbewusste (implizite) Lernerfahrungen
- intuitive Lernerfahrungen
- bewusste (explizite) Lernerfahrungen
- sprachlich vermitteltes deklaratives Wissen (Propositionen/ mentale Modelle)

In der Reihung drückt sich einerseits eine Hierarchie bezüglich der Elaboriertheit kognitiver Lernprozesse aus, andererseits bleibt unklar, inwieweit solche Lernerfahrungen als Grundlage für das Handeln ebenfalls eine qualitative Stufung für sich beanspruchen können. Für die Lehrerbildung muss es trotz dieser Unklarheiten in erster Linie um die Ermöglichung von Lernerfahrungen gehen, die über den sozialen Austausch und Reflexionsprozesse explizit gemacht werden bzw. um sprachlich vermitteltes deklaratives Wissen. Dabei sind die eher vorbewussten Lernerfahrungen mitzudenken.

6.2 Didaktisches Handeln des Experten als Ergebnis von Praxislernsituationen?

Wurde im Vorangehenden das Erfahrungswissen als Komponente des Lehrerhandelns in Praxislernsituationen herausgearbeitet, soll es im Folgenden um die Lernpotentiale gehen, die Praxisphasen innerhalb der Lehrerbildung einnehmen bzw. einnehmen können. Wenn Praktika auch häufig als das Herzstück der Lehrerbildung bezeichnet werden (Hascher 2006, S. 130), bleibt gleichzeitig verschwommen, was sie im Rahmen der Lehrerbildung tatsächlich leisten oder auch leisten können. Diese Unschärfe resultiert nach Oelkers (1999) daraus, dass Lehrerbildung sich traditionell auf folgende Effekterwartungen bezieht:

- auf Personen, nicht auf Standards;
- auf Gesinnung, nicht auf Wissen;
- auf Dogmen, nicht auf revisionsfähige Theorien;
- auf engagierte Sprache, nicht auf distanzierte Analyse;
- auf Methodenideale, nicht auf fallible Entwicklungen.

Folgt man dieser grundsätzlichen Kritik an der Lehrerbildung, erklärt sich auch, warum sie relativ wenig kritische Distanz zu ihrer eigenen Praxis einnimmt und sogenannte schulpraktische Studien als Ausbildungselement gegenüber der Theorie favorisiert. Mit den von Oelkers (polemisch) thematisierten Effekterwartungen der Lehrerbildung werden Ansprüche geweckt, die sich am Bild des „pädagogisch Guten“ orientieren, womit das Berufsfeld Schule als Ort der Verwirklichung dieser Ideale angesehen wird. Insofern darf es nicht überraschen, dass auch Studierende, deren Biographie zudem in den meisten Fällen noch stark von den eigenen Erfahrungen als Schüler geprägt ist, oftmals berufspraktische Erfahrungen dem hochschulischen Lernfeld vorziehen. Oben genannte idealistische Erwartungen sind jedoch individuell kaum einlösbar und begründen damit das „unstillbare Verlangen nach Praxisbezug“ (Hedtke 2000).
Auf der anderen Seite erscheint mir Oelkers Gegenüberstellung der Effekterwartungen kritisch überpointiert zu sein, da es schon seit längerer Zeit verschiedene Versuche gibt, die angesprochenen Kritikpunkte aufzulösen und den Praxisbezug innerhalb der Lehrerbildung konstruktiv zu nutzen (Obolenski & Meyer 2003; Lemmermöhle, Schellack & Brümmer 2006; Bayer, Carle & Wildt 2001; Merkens 2003). Besonders mit der Implementation von Zentren für die Lehrerbildung in den 1990er Jahren als Innovierungsfaktor für die Lehrerausbildung an bundesdeutschen Universitäten hat man versucht sich der angesprochenen Problematik anzunehmen (Rinkens 1999; Hilligus & Rinkens (2005); Merkens 2005) . Aufgrund der im föderalistischen System unterschiedlichen Zielsetzungen und Ausstattungen, die mit der Implementation in den einzelnen Bundesländern verbunden sind, bleiben die empirischen Auswirkungen bislang allerdings diffus (Merkens 2005, S. 103).
Die von Oelkers angesprochene Orientierung am „pädagogisch Guten“ mag nicht selten zu einer übertriebenen Idealisierung der Zielsetzungen pädagogischen Handelns führen, jedoch wird damit sichergestellt, dass es einen Kanon verbindlicher

Werte in der Interaktion mit Kindern und Jugendlichen gibt (v. Hentig 2006; Ebke & Kliemann 2006; Honnefelder 2002; Oser 1998). Zusammenfassend kann man diese Wertorientierung mit den Begriffen „Anerkennung und Schutz der Würde des Kindes“ und „bestmögliche Förderung der Entwicklungschancen“ bezeichnen. Die Grundlage eines wertorientierten Lehrerhandelns liegt dabei auch wesentlich in der Reflexion der eigenen Vorbildfunktion. Insgesamt sind damit die Säulen eines pädagogischen Berufsethos als Bestandteil einer Profession (vgl. Kap. 4.3.5) kenntlich gemacht.
Innerhalb der Lehrerbildung wird die Frage der Integration von Praktika in die akademische Ausbildung kontrovers diskutiert. Hascher (2006, S. 131) erkennt folgende Kriterien einer qualitätsvollen Integrationspraxis, die aus ihrer Sicht in bestehenden Umsetzungen weitgehend vermisst bleiben:

- die Orientierung an verbindlichen Lernzielen,
- die regelmäßige Kooperation zwischen Ausbildungsstätten und Schulen,
- die kontinuierliche und kompetente Beratung und Betreuung der Studierenden,
- die Integration der Studierenden in den Schulalltag und in das Kollegium,
- der fachliche Diskurs zwischen Experten und Novizen,
- die Verknüpfung der verschiedenen Praktika im Laufe der Ausbildung.

Die als Defizite herausgestellten Punkte sind maßgeblich dafür verantwortlich, dass der in dieser Studie untersuchte Modellversuch ins Leben gerufen worden ist. Ob damit allen Kritikpunkten begegnet werden kann, soll hier nicht entschieden werden.
Die Unklarheit darüber, was Praktika leisten können oder sollen, fügt sich in ein übergreifendes Bild, nach dem die empirische Bildungsforschung wenig dazu sagen kann, was Lehramtsstudierende am Ende ihrer Ausbildung wissen oder können sollen (Terhart 2004, S. 87). In Konsequenz dieser schmalen empirischen Datenlage kann auch nur unzureichend formuliert werden, wie denn die „optimale Lehrerausbildung“ auszusehen hätte (Blömeke 2004, S. 75). Oelkers (1999, S. 1) etwa definiert pragmatisch „Berufsfähigkeit“ als das Ziel der Lehrerbildung. Dieses Ziel steht im Spannungsfeld der individuellen Verknüpfung von angemessener Wahrnehmung und adäquatem Handeln (auf der Seite der Studierenden) und der objektiv angebotenen Vorbereitung (durch die Ausbildung). Die von Oser & Oelkers (2001) vorgelegten Befunde legen es nahe, darin eine Bestätigung der von Lorties (1975) vorgestellten These zu finden, dass es sich bei der Lehrerbildung um ein „low impact enterprise“ handelt. Im Rahmen einer solchen Positionierung sind Mechanismen zu beobachten, grundsätzliche Fragen und Probleme der Lehrerbildung auszublenden und für eine Vergrößerung des Praxisbezuges zu votieren.
Die Erwartungen von Studierenden an Praktika sind einerseits hoch, andererseits leicht zu befriedigen, weil sie tendenziell vielseitige Erfahrungen im berufspraktischen Feld sammeln wollen und dabei nicht selten Lernen mit Erfahrungen gleichsetzen (Hascher 2006, S. 133). In dieser Perspektive gibt die Berufspraxis die Ziele der Lehrerbildung vor. Hier ist auf die Gefahr der frühzeitigen Einsozialisierung in Handlungsformen hinzuweisen, die sich einseitig an den Kriterien von Funktionalität orientieren. Das Berufsfeld Schule unterliegt dem Vorwurf Funktionalitätskriterien an die erste Stelle zu setzen und deshalb Innovierungspotentiale zu hemmen. Gründe

dafür können nicht zuletzt in der schwierig zu bewerkstelligenden wachsenden Heterogenität der Schülerschaft als Abbild der sich weiterhin pluralisierenden Gesellschaft gesehen werden. Hinzu kommt die mangelnde Abstimmung zwischen den drei Ausbildungsphasen der Lehrerbildung in Deutschland, die zu der fast schon sprichwörtlichen Aussage von Vertretern jeweils nachfolgender Phasen geführt hat, die in vorangegangenen Phase gelernten Inhalte doch getrost vergessen zu können.
Die Einstellungspraxis im Schulbereich, die sich trotz der prinzipiell möglichen langfristigen Vorausplanung zumeist an kurzfristigen Bedarfslagen orientiert und deshalb innerhalb kompakter Zeiträume vergleichsweise viele Personen in den Lehrerberuf befördert, wogegen in übrigen Zeiträumen nur wenige Personen eine Einstellungschance erhalten, vulgo „Schweinezyklus“ genannt, führt zu einer relativ homogenen Altersstruktur der Lehrerschaft. Die altersbezogene Homogenität geht einher mit einer zeitkulturell bzw. biographisch begründbaren Ähnlichkeit des jeweiligen Theoriehintergrundes im besagten Personenkreis, der sich in diskursiven Prozessen des berufspraktischen Handelns wiederum stabilisiert. Dass das Lehrerhandeln wesentlich einem Handeln „unter Druck“ (Wahl, 1991) bzw. einem Handeln im Emergency Room (Oser 2001) entspricht, trägt ebenfalls nicht konstruktiv dazu bei, die eigene Person in Frage zu stellen. Erst dann kann aber eine Distanz zum eigenen Handeln gewonnen werden, womit Innovierungsprozesse freigelegt werden können.

In der Zusammenführung der diskutierten Stränge über erfahrungsbezogene (vgl. Kap. 6.1) bzw. handlungsbezogene (vgl. 6.2) Lernpotentiale kann festgehalten werden, dass Praxisphasen im Studium vorurteilsbelastete Einstellungen mit Bezug auf pädagogische Handlungsfelder verstärken können, besonders dann, wenn Studierende wenig Theoriehintergrund und berufsbiographische Erfahrungen mitbringen. Die Studie setzt jedoch Praxisphasen innerhalb des Studiums oder auch nach der ersten Phase der Lehrerbildung weder per se als bildend noch als „verbildend“ voraus (Altrichter & Hascher 2005, S. 5). Problematisch erscheint gleichwohl, dass die Handlungsunsicherheit von Novizen durch vielfältige Formen der pädagogischen Idealisierung noch gesteigert wird. Hinsichtlich des Verhältnisses zwischen angehenden Lehrern und erfahrenen Praktikern besteht die Gefahr, dass der Erfahrungsvorsprung von Praktikern gleichgesetzt wird mit der Fähigkeit, ein „guter Lehrer“ zu sein. Eine Erklärung liefert der generell positiv konnotierte Erfahrungsbegriff, der häufig auch in der Alltagssprache eine Gleichsetzung mit dem Expertisebegriffe erfährt.
Vor diesem Hintergrund können Praxisphasen ihr für die Entwicklung unterrichtlicher Kompetenzen bestehendes Potential erst entfalten – und hierbei handelt es sich neben dem Freiheitskriterium um eine ergänzende Zuschreibung zu konstruktiven Erfahrungsräumen (vgl. Kap. 6.1) – , wenn sie verantwortungsvoll vorbereitet und nachbereitet werden und damit die Gelegenheit bieten, individuelle und kritisch reflektierte Erfahrungen aufzuschließen. Eine unreflektierte Übernahme von Erfahrungen birgt nach Hascher (2005, S. 41) andernfalls die Gefahr der „Erfahrungsfalle“. Hier zuzuordnende Problembereiche sieht sie dann, wenn man dem Irrglauben verfällt, Erfahrungen würden stets zu Expertise führen bzw. Lernen garantieren, womit gleichzeitig die Neigung zu unkritischen und unreflektierten

Positionen gegenüber dem Erfahrungswissen oder entsprechenden Praxisfeldern verbunden sein kann.

6.3 Kontingenz als strukturelles Merkmal von Praxislernsituationen

Es gehört zu den zentralen Ergebnissen sowohl der empirischen Unterrichtsforschung als auch der Systemtheorie, so sie denn ihre Analysefolien auf unterrichtliche Prozesse richtet, dass Unterricht von hoher Ungewissheit und Komplexität geprägt ist. Ein Erklärungsmuster bietet die Instabilität sozialer Interaktion, die sich lediglich in Form struktureller Kopplungen beschreiben und analysieren lässt (Luhmann & Schorr 1988).
Zur Darstellung entsprechender Prozessgefüge hat sich das Kontingenztheorem[46] durchgesetzt und wurde zu einem systemtheoretischen Grundbegriff. Für die pädagogische Interaktion von Lehrern und Schülern kommt Vanderstraeten (2004, S. 63) zu dem Analyseergebnis, dass durch die Vielzahl der am Unterricht beteiligten Akteure Kontingenz derart zunimmt, dass sie „von ihrer eigenen, selbst generierten Komplexität überfordert wird". Das Kontingenztheorem bezieht sich dabei sowohl auf die durch den Konstruktivismus beförderte Erkenntnis der idiosynkratischen Welt- und Sinnkonstruktion als auch auf den sozialen Aushandlungsprozess in zwischenmenschlichen Kommunikationen und Interaktionen.[47] Bezogen auf diese beiden Prozessdimensionen des Unterrichts spricht Vanderstraeten (ebd.) von der „doppelten Kontingenz" (vgl. auch Baumert & Kunter 2006).
Proske (2006, S. 139) stellt die Kontingenz erzeugende Differenz von Unterrichten bzw. Erziehen auf der einen und Lernen bzw. Aneignen auf der anderen Seite heraus. Lernprozesse kann der einfallsreichste Lehrer dem Schüler nicht abnehmen. Unterrichten wird damit nach Luhmann zu einem „opportunistischen Prozess" (Luhmann 2002, S. 104). Im Gegensatz zu einem alltagssprachlichen Opportunitätsverständnis über Anpassungsbemühungen zwecks Wahrnehmung möglichst vieler Gelegenheiten verwendet Luhmann hier den Begriff um zu veranschaulichen, dass Lehrerhandeln lediglich einen Möglichkeitsraum schaffen kann, in dem Schüler Lerngelegenheiten nutzen oder auch nicht nutzen.
Die systemtheoretische Perspektive lenkt den Blick auf die Analyse kommunikativer Prozesse im Unterricht. Er erscheint als ein „Medium, in dem unter kommunikativen Aushandlungsbedingungen die jeweilige Bedeutung von Wissen erst situationell hervorgebracht wird bzw. emergiert" (Krummheuer 2002, S. 42f.; zit. nach Proske 2006, S. 144). Luhmann (2002, S. 102ff.) fasst Kommunikationsprozesse im Unterricht als Prozesskategorie mit den drei operativen Elementen:

[46] Der Begriff Kontingenz leitet sich aus dem lateinischen Infinitiv „contingere" ab und kann mit „sich ereignen" übersetzt werden. Im Zuge einer spätlateinischen Entwicklung bekommt das Subjektiv die Bedeutung „Möglichkeit". Als philosophischer Begriff wird damit allgemein das Notwendige und nicht Unmögliche bezeichnet, das auch als das Zufällige (das eintreten kann, aber nicht muss) gilt (Meyer 2007).
[47] Aus Perspektive der empirischen Forschung bleibt es strittig, inwieweit sich „Kontingenz" beobachten oder methodisch rekonstruieren lässt. Die alleinige Tatsache, dass kommunikative Botschaften im Unterricht (und auch in anderen Zusammenhängen) über verschiedene Bedeutungshorizonte vom Adressaten entschlüsselt, verstanden und an den Sender rückgekoppelt werden müssen, verweist auf forschungsmethodische Schwierigkeiten bei der Analyse solcher Phänomene.

a) Information
b) Mitteilung
c) Verstehen

Die systemtheoretische Analyse der Unterrichtskommunikation hebt die Bedeutung von Verstehensprozessen hervor. Erst durch Verstehen konstituiert sich die Bedeutung von unterrichtlichem Wissen. Proske (2006, S. 145) beschreibt Verstehensleistungen als ein „im Unterricht stattfindendes Ereignis, das weder mit den Intentionen des Lehrers übereinstimmen muss noch in der Interaktion dauerhaft fixiert“ werden kann. Verstehen schließt in diesem Sinne das Missverstehen mit Bezug auf die Ursprungsintention des mitteilenden Lehrers ein. Es ist prinzipiell produktiv und eigensinnig, wenngleich strukturierende Erwartungen z.B. der Organisation oder des sozialen Gedächtnisses der Schulklasse das Eintreten „richtig-verstehender“ Anschlusskommunikation wahrscheinlicher machen.
Das Kontingenztheorem als Analysefolie für den Interaktionsraum Unterricht verweist auf grundlegende darin auszumachende Technologiedefizite. Scheunpflug (2004, S. 84) definiert in diesem Zusammenhang Unterricht als „nicht beliebige Kommunikationsofferte, deren Resonanz nicht determinierbar ist“. Als Zielperspektive für die Lehrerbildung lässt sich ableiten, dass sie ihre Absolventen in die Lage versetzen muss, statt einer rigiden Unterrichtsverlaufsplanung ein variables Schema an den Unterricht anzulegen und das didaktische Handeln adaptiv an die Unvorsehbarkeiten anzupassen. Eine solche systemtheoretisch orientierte Didaktik schätzt Scheunpflug (2004, S. 86) wie folgt ein:

„Eine teleonome Theoriebildung verzichtet darauf, das Lernen auf zielgerichtetes Lehren zurückzuführen. Sie kultiviert die Erwartung von Vielfalt und macht misstrauisch gegenüber engen normativen Erwartungen. Eine teleonome Theoriebildung entlastet, da Lehrende sich damit nicht mehr für das Gelingen des Lehr-Lernprozesses allein verantwortlich fühlen müssen: Die Selektion von Bildungsofferten liegt nicht in ihrer Hand. Gleichzeitig werden aber auch klare Verantwortlichkeiten aufgezeigt; denn das Variationsangebot, die kommunikative Offerte liegt in ihrer Hand“.

Damit sind die zentralen Anforderungsfelder für die Lehrer-Schüler-Interaktion beschrieben, gleichzeitig scheinen aber mögliche Überforderungstendenzen durch. Die systemtheoretische komplex-abstrakte Analyse unterrichtlicher Prozesse vernachlässigt zuweilen, dass die Berücksichtigung der von ihr generierten, theoretisch überzeugenden Ergebnisse im realen unterrichtlichen Handeln durch Lehrpersonen kaum einzuholen sind. Aus diesem Blickwinkel muss der Unterricht in der Praxis zwangsläufig als defizitär eingeschätzt werden. Im Gegensatz dazu deklariert Proske (2006, S. 148f.) die z.T. als problematisch bewertete pragmatische Engführung des Unterrichts (Seidel et al. 2003b, S. 161) zur notwendigen Strategie der Kontingenzbewältigung durch Kontingenzreduktion. Er weist darauf hin, dass „Unterricht grundsätzlich nicht umhin kann, bestimmte Lerngelegenheiten zu übergehen und andere zu nutzen“ (Proske 2006, S. 148f.). So kann das Einräumen von schülerorientierten Partizipationsmöglichkeiten mit unterrichtlichen Zielorientierungen kollidieren und ein Konfliktpotential im anzustrebenden

kooperativen Verhältnis von Lehrern und Schülern eröffnen, dass sich mit daraus ergebenden divergierenden Interessenlagen erklären lässt.
Proske bewerkstelligt den Spagat zwischen einer systemtheoretischen und empirischen Analyse von unterrichtlichen Prozessen. Unterricht steht damit im Spannungsfeld von Kontingenzerzeugung und Kontingenzreduktion, wobei die Notwendigkeit von Reduktionsstrategien nicht als defizitär, sondern als unausweichliche, strukturelle Gegebenheit didaktischen Handelns von Lehrpersonen beschrieben wird, die sich gleichwohl aus einer normativen Position heraus kritisieren lässt.

Die bisherigen Ausführungen beziehen sich wesentlich auf die didaktische Frage, wie es Lehrpersonen gelingen kann, sowohl Kontingenz, als wesentliche Bedingung von Unterricht als sinnbasiertes System (Proske 2006, S. 146), zu erzeugen sowie gleichzeitig über Strategien der Kontingenzreduktion handlungsfähig im Unterricht zu bleiben und nicht von prinzipiell geltender Uneindeutigkeit und Komplexität überrollt zu werden. Die Zielperspektive liegt jeweils in der schülerbezogenen Ermöglichung von Lerngelegenheiten.
In der eigenen Studie geht es vor allem um die Frage, welche Lernpotentiale verlängerte Praxisphasen innerhalb des Studiums für angehende Lehrer enthalten. Praxissituationen bieten Lernräume, in denen Studierende sich erproben, üben und Erfahrungen sammeln können (Reinhoffer, Barthold & Küster 2007, S. 89). Die Kontingenz unterrichtlicher Verlaufsdynamiken, als nur z.T. antizipierbare Reaktionen auf das eigene Handeln, benennt dabei sowohl Herausforderungen als auch daraus erwachsene Lernchancen für Studierende in Praxissituationen. Vor dem Hintergrund des Kontingenztheorems werden Praxissituationen zu Praxislernsituationen. Sie schaffen Lerngelegenheiten für selbst gesteuerte Aktivitäten der Studierenden, die unmittelbar an die eigene Wahrnehmung rückgekoppelt werden und weitgehend kontingenten Verlaufsmustern unterliegen.[48]
Die dargestellten Wechselspiele zwischen Kontingenzerzeugung und Kontingenzreduktion im unterrichtlichen Handeln stellen einen wesentlichen Theorierahmen für die Entwicklung der eigenen methodischen Zugänge sowie die Ergebnisinterpretation dar.

6.4 Fazit

Wenn Theorie und Praxis auch nur modellhaft unterschieden werden können und man insofern von einem Missverständnis bezüglich dieser Trennung sprechen kann, so hat sich ein entsprechender Sprachgebrauch trotzdem dauerhaft durchgesetzt. Für die Lehrerbildung, so kann man kritisch einwenden, ist die hochschulische Tendenz auszumachen, sich den Erfordernissen der späteren beruflichen

[48] Auch im Praxisjahr unterliegen die Studierenden zwar gewissen Kontrollmechanismen z.B. durch Mentoren oder institutionelle Gepflogenheiten und Routinen, die sich in jeweiligen Schulformen verfestigt haben. Der von Oelkers (1999) vorgetragene Vorwurf, dass der Praxisbezug in einer hochschulischen Ausbildung lediglich als Idealisierung und Modell in Form schulpraktischer Studien simuliert werde, wobei Praxis im Regelfall mit „Unterricht“ gleichgesetzt würde, ohne die den Unterricht beeinflussenden Kontextfaktoren der Schule mit zu berücksichtigen, trifft aber für das Praxisjahr nur in deutlich abgeschwächter Form zu.

Einmündung ihrer Absolventen weitgehend zu entziehen und diese Verweigerungshaltung mit Verweis auf akademische Grundsatzpositionen zu legitimieren. Generell stellt sich die Frage, welche Wissenstypen vorwiegend aus Praxisbegegnungen als Lerngelegenheiten generiert werden können, die Gütekriterien der Theoriebildung standhalten oder zumindest Anschlussoptionen eröffnen. Umgekehrt ist zu fragen, welche Funktionen Theorien für die Bewältigung von Praxisanforderungen anbieten können. Entsprechende Möglichkeiten sind oben dargestellt worden (vgl. Kap. 6.1). Dabei darf es nicht überraschen, wenn Studierende des Lehramtes Probleme der Akzeptanz von erziehungswissenschaftlicher Theoriebildung bekunden, wenn der Zweck dieser Theoriebegegnung dahingehend begründet wird, für spätere Handlungssituationen ein Reflexionswissen zur Verfügung zu stellen, welches innerhalb des Studiums von den Studierenden aber wegen ihrer fehlenden berufspraktischen Erfahrungen nicht nach Nützlichkeitskriterien eingeschätzt werden kann. Diese Argumentation liefert plausible Argumente für die Integration verlängerter Praxisphasen in die Lehrerbildung.

Im Theorie-Praxis-Verhältnis liefert die Theoriebildung aus Sicht der Praxis funktionale Leistungen, womit Handlungsmöglichkeiten beschrieben, erklärt oder sogar im Hinblick auf ein Ursache-Wirkungs-Verhältnis analysiert werden können. Umgekehrt bieten Praxisbegegnungen durch Kontingenzerfahrungen die Möglichkeit einer Überprüfung dieser Theorieoptionen in Situationen, die nur teilweise vorweggenommen werden können. Praxiserfahrungen als potentielle Lerngelegenheiten bieten zudem die Herausforderung sich mit der ganzen Person in berufliche Situationen einbringen zu können und auch zu müssen. Sie schaffen Lerngelegenheiten, die zutreffend mit der Erlangung von Erfahrungswissen beschrieben werden können und insgesamt den Wert der schulpraktischen Studien innerhalb der Lehrerbildung legitimieren.

Man darf diesbezüglich fragen, zu welchem Zeitpunkt und mit welcher zeitlichen Intensität Studierende innerhalb des Studiums mit Praxisbegegnungen konfrontiert werden sollen. Eine zu frühe Begegnung birgt die Gefahr von Überforderungen bzw. wirft die Frage auf, wie Überforderungen durch Beratungsleistungen verhindert werden können. Eine zu späte Begegnung bewirkt möglicherweise, dass die Potentiale des Erfahrungswissens nicht mehr mit Fragen der wissenschaftlichen Theoriebildung individuell verknüpft werden. Dagegen kann eine frühe Begegnung mit den Anforderungen der beruflichen Praxis wiederum dazu führen, dass die Bedeutung wissenschaftlicher Theoriebildung gering geschätzt wird. Dies wird vor allem dann der Fall sein, wenn keine funktionalen Äquivalenzen der Theorien für das Handeln gefunden werden.

Theoretisches Wissen dient dann nicht der Handlungsbefähigung in einem funktionalen Sinne der Komplexitätsreduktion, sondern vergrößert eher noch die Kontingenzerfahrung. Aus einer wissenschaftlichen Perspektive ist dies wünschenswert, weil damit die These verbunden wird, der Komplexität der Praxis in einem höheren Maße gerecht werden zu können. Eine empirische Untersuchung der Auswirkungen eines in das Studium integrierten Halbjahrespraktikums an der Universität Bremen (Hoeltje et al. 2002, 2004) kam zu dem Ergebnis, dass die zugeschriebene Bedeutung erziehungswissenschaftlicher Theoriebildung für das eigene pädagogische Handeln im Anschluss an das Praktikum zugenommen hat,

während die Bedeutung fachwissenschaftlicher Inhalte in der Relation eher abgenommen hat (gemessen an dem Wunsch diesbezügliche Veranstaltungen in der Fortführung des akademischen Studiums besuchen zu wollen). Die Gefahr einer „praxeologischen“ Verkürzung der Handlungsoptionen durch eine verlängerte Praxisphase innerhalb des Studiums konnte in der genannten Studie somit nicht bestätigt werden. Als problematisch wird eher der hohe Sozialisationseffekt herausgestellt, der darin besteht den Unterrichtsstil der betreuenden Mentoren zu kopieren, mit dem tendenziellen Ergebnis einer Favorisierung des Frontalunterrichts. Daraus kann man folgern, dass Praxisbegegnungen nicht per se positive oder negative Konsequenzen zeitigen, sondern wesentlich die jeweilige Gestaltung eine Rolle spielt. Betrachtet man den Sozialisationseffekt (den man auch für die zweite Phase der Lehrerbildung in Deutschland annehmen kann), so spricht einiges für die Integration verlängerter Praxisphasen in das akademische Studium, weil man damit zumindest die Hoffnung verbinden darf, dass es im weiteren Verlauf des Studiums zu einer Reflexion der gemachten Erfahrungen kommt, worin eine wichtige Korrektivfunktion für das praktische Handeln liegt.

7 Ergebnisse der empirischen Bildungsforschung

Bislang liegen erst wenige Studien darüber vor, welche Kompetenzen bei angehenden Lehrern durch eine institutionalisierte Lehrerbildung ausgebildet werden (vgl. Kap. 7.2). Mehr Befunde liefern dagegen die nationalen und internationalen Schulleistungsstudien (vgl. Kap. 7.1). Für die eigene Fragestellung liefern diese Studien einen empirischen und theoretischen Referenzrahmen, weil sie weit reichende Aussagen darüber generiert haben, welche Lernpotentiale das Feld Schule über die Unterrichtsgestaltung für Lernende bietet. Mittelbar kann man den Versuch unternehmen die Daten auf das unterrichtliche Handeln von Lehrpersonen rückzubeziehen und daraus Folgen für die unterrichtliche Praxis und – in einem weiter reichenden Schlussverfahren – für die Lehrerbildung an sich bzw. die damit intendierten Ziele abzuleiten. Im Rahmen dieser Studien wurde auch in Form von schriftlichen Befragungen der Schüler und Lehrpersonen Aufschluss über die konkrete Unterrichtspraxis in verschiedenen Ländern gewonnen, womit die aggregierte Beschreibung verschiedener Unterrichtskulturen und deren Differenzierung nach Ländern möglich wurden. Über derart transparente Erklärungsmuster konnten bessere oder schlechtere Schülerleistungen im internationalen Vergleich plausibel nachvollzogen werden. In diesem Zusammenhang wurden zudem Qualitätsdimensionen des Unterrichts untersucht, die sich förderlich oder hemmend auf komplexe Lernprozesse von Schülern auswirken und sich fachspezifisch unterscheiden lassen. Dabei kam es auch zum Einsatz von Videostudien, die den unmittelbaren Unterrichtsprozess in den Blick nehmen. Die in diesen Studien gewonnenen Aussagen über Erfolgskriterien von unterrichtlichen Lehr-Lernprozessen werden im Folgenden voraussetzend für die eigene Untersuchung nachvollzogen. Die Darstellung der Fragestellungen, methodischen Vorgehensweisen und übergreifenden Ergebnisse dient der Kontextuierung der Befunde zu den Erfolgskriterien unterrichtlicher Lehr-Lernprozesse.

Für die deutsche Diskussion ist zu beobachten, dass die Ergebnisse der empirischen Bildungsforschung erst durch die PISA-Studie (Baumert et al. 2001) eine zumindest öffentlich erhöhte Aufmerksamkeit erfahren haben. Die TIMMS-Studie (Baumert et al. 1997; Baumert, Bos & Lehrmann 2000a, b), die bereits früher tendenziell ähnliche Ergebnisse für das Fach Mathematik und die Naturwissenschaften aufwies, blieb in der Rezeption einem Expertenkreis vorbehalten und konnte sich nur wenig öffentliche Resonanz verschaffen. Das im internationalen Vergleich relativ schlechte Abschneiden der deutschen Schülerschaft hat auch die Frage aufgeworfen, inwieweit die deutsche Lehrerbildung ihre Absolventen zu wenig auf schulische Anforderungen vorbereitet.[49] Bislang fehlt eine repräsentative Studie, die Auswirkungen der deutschen Lehrerbildung darlegt. Die bereits genannte Studie TEDS-M intendiert das Feld aufzuhellen und befindet sich gegenwärtig noch im Planungsstadium (Blömeke 2006a).

[49] Es bleibt strittig, ob eine so weitgehende Schlussfolgerung aus den Daten abgeleitet werden kann. Generell ist festzustellen, dass die Formulierungen „nach PISA“ oder „PISA sagt“ schon den Charakter geflügelter Worte annehmen, wobei die aus der Studie gezogenen Ableitungen häufig nicht haltbar sind. Der darin sich andeutende, auch politisch getragene Überschwang kann auf eine vorangegangene Leerstelle in der Rezeption empirischer Studien zum Bildungssystem zurückgeführt werden.

Im weiteren Vorgehen werden die Ergebnisse von Studien vorgestellt, die sich unmittelbar mit den Ergebnissen und also der Wirksamkeit der Lehrerbildung beschäftigen (vgl. Kap. 7.2). In Form größerer quantitativer Studien liegen sie überwiegend aus der Schweiz und den Vereinigten Staaten von Amerika vor. In den Blick rückt zunächst die Untersuchung von Oser & Oelkers (2001) zur Wirksamkeit der Lehrerbildungssysteme in der Schweiz. Es handelt sich um die für den deutschsprachigen Raum umfangsreichste Untersuchung, die auch bei konzeptionellen Unterschieden der Lehrerausbildung in der Schweiz und Deutschland tendenziell Übertragungen auf die deutsche Situation zulässt. Im Anschluss an Wirksamkeitsuntersuchungen zum schweizerischen Lehrerbildungssystem interessieren Evaluationsstudien zur Untersuchung der Lehrerbildung aus den USA, die weitgehend eine outputorientierte Richtung verfolgen. Die Ergebnisse dieser empirischen Untersuchungen werden allerdings unterschiedlich gedeutet und werfen die Frage auf, wie Inputs der Lehrerbildung gestaltet werden müssen, um Ergebnisse zu verbessern. Da die Diskussion um eine standard- und outputorientierte Lehrerbildung in Deutschland zunehmend virulent ist, werden die Diskussionslinien in den USA kritisch nachgezeichnet.
Mit der Arbeit von Frey (2004) liegt eine europäische Vergleichsstudie der erworbenen Kompetenzen von Studierenden des Lehrerberufes vor, in der auch deutsche Studierende untersucht wurden. Mit der expliziten Thematisierung von Kompetenzdimensionen ist diese Studie auch für die eigene Herangehensweise interessant. Für Deutschland selbst liegen nur wenige Untersuchungen vor, die sich mit den konkreten Wirkungen der Lehrerbildung befassen. Die größte Aufmerksamkeit gilt dabei den Auswirkungen von Praxisphasen. Während sich Bodensohn, Frey & Balzer (2004) und Bodensohn & Schneider (2006) mit den Kompetenzentwicklungen von Studierenden in Blockpraktika befassen, untersucht Abs (2006) die Kompetenzentwicklung von Studierenden in der zweiten Phase der deutschen Lehrerbildung, die konzeptionell als Praxisphase angelegt ist. Nölle (2002) schließlich analysiert Unterschiede, die sich aus einer veränderten Lehrerbildung im Vergleich zu „traditionellen“ Ausbildungsformen ergeben. In den neuen Studienkonzepten erhalten Studierende die Möglichkeit, ihre Praxiserfahrungen unmittelbar in die Gestaltung der Lehrerbildung einzubringen, wodurch Reflexionsprozesse angeregt werden sollen.
Insgesamt liefern die deutschen Studien Aufschluss darüber, inwieweit Praxisphasen oder veränderte Konzeptionen, die Praxisphasen als Reflexionsfolie nutzen wollen, Auswirkungen auf Kompetenzentwicklungsprozesse bei Studierenden ausüben können.
Die sowohl mittelbaren als auch unmittelbaren Befunde zu Wirksamkeitsdimensionen der Lehrerbildung geben den theoretischen Hintergrund für die eigene Studie ab, die in Anschluss an bestehende Ergebnisse die Wirksamkeit von Praxislernphasen videobasiert untersucht.

7.1 Schülerleistungen

Hier zu nennende nationale (BIJU / MARKUS) bzw. internationale (DESI, IGLU, PISA, TIMMS) Schulleistungsstudien fokussieren die Leistungsentwicklung von Schülern und versuchen Zusammenhänge zwischen dieser Entwicklung und ausgewählten didaktischen Handlungsformen von Lehrpersonen herzustellen. Darüber hinaus wird der Versuch unternommen, Kontextvariablen (z.B. Elternhaus, Vorwissen, Intelligenz etc.) in ihrem möglicherweise moderierenden Einfluss auf Leistungsverläufe bei Schülern zu kontrollieren. Diese Anlage ist der Erkenntnis über multipel determinierte Schulleistungen geschuldet.

„Implizit gehen die meisten multipel angelegten Untersuchungen und Modelle von einer unbegrenzten wechselseitigen Kompensierbarkeit verschiedener Determinanten und von einem additiven Zusammenwirken aus“ (Helmke & Weinert 1997, S. 139).

Die wichtigsten Studien werden nachfolgend in ihrer Relevanz für die eigene Fragestellung dargestellt. Zentral sind Ergebnisse, die als Indikatoren bei der Entwicklung von methodischen Beobachtungsverfahren angelegt werden können, weil sie Aufschluss über Erfolgskriterien unterrichtlichen Handelns gewähren und damit die eigene Untersuchung wesentlich tangieren.

7.1.1 BIJU (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter)

Die 1991 gestartete Längsschnittstudie wurde vom Max-Planck-Institut für Bildungsforschung in den drei Bundesländern Mecklenburg-Vorpommern, Nordrhein-Westfalen und Sachsen-Anhalt (1995 mit einer Stichprobe auch in Berlin) durchgeführt (Baumert & Köller 1998; Gruehn 2000). Sie war fachspezifisch angelegt und berücksichtigte die Fächer Mathematik, Physik und Biologie in allen Schulformen. Der Längsschnitt begann bei siebenten Klassenstufen und verfolgte die Schülerentwicklungen bis zum Abitur bzw. zum Übergang in das Studium oder den Beruf. Für die Unterrichtsgestaltung konnten lernförderliche Effekte von Formen direkter Instruktion herausgestellt werden. Genauer wurden Faktoren der effizienten Klassenführung, der Regelklarheit und ein der Klasse angemessenes Unterrichtstempo sowie die Klarheit und Sequenzierung des Unterrichtsablaufs positiv expliziert. Unterrichtsansätze, die einer konstruktivistischen Lernauffassung entsprechen und deshalb z.B. auf die lernende Rekonstruktion übergreifender Konzepte abzielen (vgl. Reinmann-Rothmeier & Mandl 1997, 2001), erwiesen sich ebenfalls als leistungsförderlich.
Gleichzeitig zeigen sich jedoch negative Auswirkungen auf die Leistungsentwicklung der Schüler durch offene Unterrichtsformen sowie Versuche der Binnendifferenzierung und Individualisierung, wobei als Haupterklärungsmoment von den Autoren eine damit verbundene Zeitabsorption angeführt wird. Dieses Ergebnis lässt sich jedoch nicht generell auf alle Schulformen und Altersstufen übertragen. Allerdings macht es darauf aufmerksam, dass offene Unterrichtsformen und Versuche einer Binnendifferenzierung und Individualisierung ein hohes und im Vergleich zu Formen der direkten Instruktion möglicherweise höheres Fehlerpotential

aufweisen. Nach wie vor stellt ein lehrerinstruierender Unterricht die häufigste Unterrichtsform dar, womit Schwierigkeiten mit offenen Formen auch auf fehlende Gewohnheiten und Routinen zurückgeführt werden könnten. Aus der Lehrerperspektive wird deutlich, dass offene Unterrichtsformen mit einem hohen organisatorischen und diskursiven Aufwand verbunden sind, womit unter Umständen weniger Zeit für das eigentliche, auch eigenverantwortliche, Lernen verbleibt. Überraschenderweise konnten in der Studie auch keine Korrelationen zwischen der emotional-affektiven Lehrer-Schüler-Beziehung und dem fachlichen Leistungszuwachs gefunden werden.[50]

7.1.2 DESI (Deutsch-Englisch Schülerleistungen International)

Die DESI-Studie wurde unter der Leitung des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) als Längsschnittstudie mit zwei Messpunkten durchgeführt (Beck & Klieme 2007). Gegenstand der Untersuchung sind sowohl die reproduktiven als auch produktiven schriftlichen und mündlichen Leistungen von Neuntklässlern in den Fächern Deutsch und Englisch als erster Fremdsprache, wobei Messungen über alle Schulformen vorgenommen wurden. An der Studie haben alle 16 Bundesländer teilgenommen. Ziel war die Erfassung von Kompetenzen im aktiven und passiven Gebrauch des Deutschen und Englischen als nationale Ergänzungsstudie zu PISA 2000, die im sprachlichen Bereich nur die Untersuchung von Lesekompetenz vorsah.
Für die Kompetenzentwicklung im Englischen wurden detaillierte Ergebnisse ausgewiesen. Danach erreichen zwei Drittel aller Schüler ein Kompetenzniveau, welches sie befähigt, Einzelinformationen aus Zusammenhängen alltäglicher Kommunikation zu verstehen und daran anzuknüpfen. Zudem konnte eine starke Leistungsspitze von etwa 10 bis 15 Prozent der Schüler (vor allem an den Gymnasien) ausgemacht werden, deren Kompetenzniveau weit über das Anforderungsniveau der Lehrpläne herausragt (Klieme 2006a). Auf der anderen Seite des Leistungsspektrums offenbarten sich Defizite von Hauptschülern, wonach nur ein Drittel der Schüler das in den für die Studie definierten Bildungsstandards formulierte Kompetenzniveau erreicht. Insgesamt fällt auf, dass die Kompetenzunterschiede zwischen den Bildungsgängen hoch ausfallen.
Hinsichtlich der Kompetenzentwicklung im Deutschen sind geringere Unterschiede zwischen Schülergruppen zu verzeichnen. Zwei Drittel der Schüler sind am Ende der neunten Jahrgangsstufe in der Lage einen Brief so zu verfassen, dass sie problemlos mit ihrem Anliegen verstanden werden. Auch hier gibt es eine verstärkt an den Gymnasien zu findende Leistungsspitze von ca. 13 Prozent. Wiederum werden Probleme der Hauptschulen offensichtlich, indem etwa die Hälfte ihrer Schüler am Ende der Jahrgangsstufe neun nicht in der Lage ist, einen Brief adressatengerecht zu formulieren.

[50] Auch dieses Ergebnis, so meine These, ist nicht auf alle Schulformen und Altersstufen übertragbar. Im Sinne der konstruktivistischen Erkenntnisperspektive lässt sich daraus aber ableiten, dass Schüler ab einem gewissen Alter (gemäß der Studie ab der siebenten Klassenstufe) nicht vorrangig lernen, um dem Lehrer gegenüber gefällig zu sein oder von ihm wertgeschätzt zu werden.

Als Ergebnis einer in Rahmen von DESI durchgeführten Videostudie konnten im Vergleich zu den TIMMS-Daten andere Unterrichtsmerkmale in ihrer Bedeutung für den Kompetenzzuwachs herausgestellt werden. Die im mathematischen und naturwissenschaftlichen Unterricht herausragende Bedeutung von Strukturierungen für den Lernerfolg spiegelte sich in der DESI-Studie so nicht wider.[51] Am Beispiel der Förderung des Hörverstehens lassen sich folgende fachübergreifende – d.h. auf die Fächer Deutsch und Englisch bezogene – Merkmale erfolgreichen Unterrichts empirisch herausstellen (Klieme 2006b):

- Effiziente Klassenführung und Störungsfreiheit
- Anspruchsniveau: Wichtigkeit der Lernziele, „Korrektheit“ und Flüssigkeit im Lehrverhalten
- Integration der Hausaufgaben in den Unterricht
- Konstruktiver Umgang mit Fehlern
- Thematische und instrumentelle Motivierung
- Ein positives Lehrer-Schüler-Verhältnis
- Verwendung der englischen Sprache in Phasen des Unterrichtsgespräches
- Hohe Sprechanteile für Schüler/innen
- Viele Gelegenheiten zur Fehler-Selbstkorrektur
- Ausreichende Wartezeit nach Lehrerfragen
- Wenige „Ein-Wort-Sätze“
- Lehrer-Schüler-Dialoge über mehr als drei Stationen

Neben expliziten Strukturierungselementen stehen atmosphärische und beziehungsrelevante Faktoren guten Unterrichts. In der Auflistung finden sich einige Merkmale, die wenig fachspezifisch erscheinen, wie z.B. Klassenführung und Störungsfreiheit, thematische und instrumentelle Motivierung oder eine ausreichende Wartezeit nach Lehrerfragen. Darüber hinaus scheinen für die sprachlichen Fächer aber Spezifika auf. Der Spracherwerb ist, unabhängig davon, ob es sich um die Muttersprache oder eine Fremdsprache handelt, von Lerngelegenheiten durch Sprachanwendung bedingt, wenn diese zugleich mit einem Anspruch durch die Lehrperson versehen werden. Weiterhin bildet der eingeforderte hohe Sprechanteil von Schüler/innen oftmals nicht die realen Bedingungen ab. So konnte in einer Videostudie, die Sprechanteile im Fach Englisch untersuchte, aufgezeigt werden, dass der Redeanteil von Lehrpersonen bei etwa 50% liegt, wobei in einer anschließenden Befragung der jeweils eigene Redeanteil z.T. deutlich geringer

[51] Die unterschiedliche Bedeutung von Strukturierungselementen lässt sich mit unterschiedlichen Fachkulturen erklären. Der mathematische und naturwissenschaftliche Unterricht orientiert sich deduktiv an der Bearbeitung spezifischer Aufgabentypen oder der Rekonstruktion naturwissenschaftlicher Phänomene. Nicht selten geht es um Aufgabenstellungen, die bei falscher Bearbeitung einen Lernabbruch zur Folge haben. Strukturierungselemente von Lehrpersonen können so betrachtet als Versuch gewertet werden, Lernabbrüche zu verhindern, womit sie tendenziell unersetzlich für den Lernprozess werden. In sprachlichen Fächern erwachsen Kompetenzen durch die Verwendung von Sprache und den Umgang mit ihr an sich bzw. werden Kompetenzen mit Bezug auf die Sprachverwendung definiert. Insofern erscheint es tautologisch, wenn Klieme (2006b) „Sprechen“ als wichtige Ressource des Unterrichts für den Fremdspracherwerb darstellt. Für diese Erkenntnis hätte es keiner aufwendigen empirischen Studie gebraucht.

eingeschätzt wurde (Klieme 2006b).[52] Zudem ließ sich beobachten, dass Lehrpersonen kaum eigene Fehler noch die von Schülern korrigieren. Schließlich konnte die DESI-Studie einen nicht unbedeutenden Zusammenhang zwischen der Klassengröße und der Kompetenzentwicklung im Englischen nachweisen. In der Vergangenheit zeugten empirische Studien von einem nur geringen Zusammenhang zwischen der Klassengröße und dem instruktionalen Einfluss von Lehrpersonen bzw. dem Lerngewinn von Schülern (Helmke & Weinert 1997, S. 95). Nachweise gelangen nur für Extremwerte (unter 16 Schüler pro Klasse/ mehr als 37 Schüler pro Klasse). Man folgerte aus diesen Ergebnissen, dass Lehrpersonen sich zwar nach eigener Aussage in kleinen Klassen wohler fühlen, die damit verbundenen Potentiale aber nur wenig nutzen. Der in DESI nun explizit nachgewiesene Zusammenhang von Klassengröße und Kompetenzentwicklung spiegelt den Sachverhalt wider, dass in größeren Klassen nur wenige Schüler die Möglichkeit erhalten, durch aktives Sprechverhalten am Unterricht zu partizipieren. Zusammenfassend können somit offensichtliche Unterschiede zwischen erfolgreichen Unterrichtsmerkmalen im Sprachunterricht sowie im mathematisch-naturwissenschaftlichen Unterricht festgestellt werden.

7.1.3 IGLU (Internationale Grundschul-Leseuntersuchung)

Bei der IGLU-Studie[53] handelt es sich um die deutsche Teilstudie der internationalen Lesestudie der IEA PIRLS (Progress in International Reading Literacy Study). Die Hauptuntersuchung fand 2001 unter Beteiligung aller 16 Bundesländer in vierten Klassen der Grundschule statt (Bos et al. 2003). Im Mittelpunkt stand die Lesekompetenz als weit gefasstes Konstrukt, indem auch Einstellungen zum Lesen und zu der Nutzung des Gelesenen erhoben wurden (Spinner 2003).
Die IGLU-Konzeption für Grundschulen rekurrierte auf die international anschlussfähige Vorstellung einer Grundbildung („literacy“), die insbesondere auf die kulturelle Bedeutung von Bildungsinhalten verweist. Grundbildung stellt zudem eine wichtige Ressource für anschließende Lernprozesse in weiterführenden Schulen und Bildungseinrichtungen dar.
Wenn der Schwerpunkt von IGLU auf die Untersuchung des Leseverständnisses gelegt wurde, erhob man doch in Teilstichproben zusätzlich Grundkompetenzen im Fach Mathematik und den Naturwissenschaften. Lesekompetenz wurde insgesamt in vier Kompetenzstufen modelliert, die wiederum eine Zuordnung zu verschiedenen Punktwerten erfuhren. Im Ergebnis erscheinen Testleistungen am Ende der vierten Klasse, deren Punktwerte lediglich die Zuordnung zu den ersten beiden Kompetenzstufen erlauben, problematisch und verweisen auf einen erhöhten

[52] Generell ist es nicht unproblematisch von einem vergleichsweise hohen oder niedrigen Sprechanteil von Lehrpersonen auf Faktoren von Unterrichtsqualität zu schließen. Stellt die DESI-Studie jedoch die Bedeutung einer aktiven Verwendung der Unterrichtssprache durch die Schüler heraus, kann man das extensive Sprechverhalten der Lehrpersonen zumindest als wenig förderlich für die Erreichung zentraler Bildungsstandards bewerten. Das gilt besonders dann, wenn Lehrpersonen ihre eigenen Sprechanteile falsch einschätzen und diesbezüglich einen geringen Bewusstheitsgrad dokumentieren, was einer Veränderung des Kommunikationsverhaltens Grenzen setzt.
[53] Aktuell liegen die Ergebnisse der IGLU-Studie 2006 in Teilen vor, die aufgrund ihrer noch unvollständigen Gesamtdarstellung aber keine Berücksichtigung in dieser Arbeit finden.

Förderbedarf. Er betrifft in Deutschland ca. ein Drittel der Schüler. Als Risikogruppe gelten Schüler, die Kompetenzstufe eins oder darunter erreichen. In Deutschland zählen etwa 10% zu dieser Gruppe (OECD-Mittel: 14,1%, EU: 11,1%). Indizien sprechen dafür, dass die Risikogruppe in weiten Teilen mit der Gruppe von Schülern identisch ist, die ihre Schullaufbahn ohne Abschluss beendet.
International schneidet Deutschland damit vergleichsweise gut ab. Deutsche Schüler erreichen im Mittel einen Punktwert von 539, der einer Zuordnung zu dem unteren Bereich der Kompetenzstufe drei entspricht. Etwa zwei Drittel der Schüler können mindestens der Kompetenzstufe drei zugeordnet werden. Im Vergleich zu anderen Ländern fallen die Streuungsmaße gering aus, woraus man schlussfolgern kann, dass die deutschen Grundschulen vergleichsweise hohe Kompetenzprofile im Leseverständnis generieren und dies gleichzeitig in homogener Breite erreichen. Eine Erklärung für das gute Abschneiden bietet ein Unterrichtsverständnis von Lehrpersonen, das auf Prozessen der Differenzierung und individuellen Förderung basiert. Diese an sich positiven Differenzierungs- und Individualisierungstendenzen beinhalten gleichzeitig problematische Gewichtungen im Grundschulunterricht, denn sie zielen hauptsächlich auf die Förderung schwächerer Schüler ab. Die Spitzengruppe der Schüler, die Kompetenzstufe vier oder darüber erreicht (in Deutschland 18,1 %; im OECD-Durchschnitt: 18,0%; in der EU: 20,6%), erfährt nach Aussagen befragter Lehrpersonen wenig Förderung. Hinsichtlich des Geschlechterverhältnisses kann man feststellen, dass sich Jungen und Mädchen im Leseverständnis nur wenig unterscheiden.
Neben dem Leseverständnis wurden auch mathematische und naturwissenschaftliche Kompetenzen der Grundschüler in den Blick genommen. Für deren Untersuchung wurden Testverfahren entwickelt, die fünf Kompetenzstufen differenziert abbilden. Deutsche Viertklässler erreichten in beiden Fächern die Kompetenzstufe IV. Betrachtet man die Punktwerte genauer, so stellt man fest, dass sie im mathematischen Bereich mit 545 deutlich (internationaler Durchschnitt 529) und im naturwissenschaftlichen Bereich mit 560 (internationaler Durchschnitt 524) sehr deutlich über den internationalen Vergleichswerten liegen. In den Naturwissenschaften sind die deutschen Grundschüler damit im oberen Leistungsdrittel platziert.
Die für Deutschland insgesamt recht positiven Ergebnisse können jedoch ein Hauptmanko des bundesdeutschen Schulsystems nicht verdecken. Vergleicht man die erreichten Punktwerte der getesteten Schüler mit ihren Noten und den großenteils darauf basierenden Empfehlungen für die weiterführenden Schulen, so streuen alle drei Kompetenzstufen über die drei Schularten. Bewertet man die für die Studie international entwickelten Verfahren zur Kompetenzdiagnostik vor dem Hintergrund der schon vor 30 Jahren benannten „Fragwürdigkeit der Zensurengebung“ (Ingenkamp 1977, 1988) an Schulen als vergleichsweise objektiv, valide und vor allem als reliabel, so kommt man zu dem Schluss, dass es den am Ende der Grundschule praktizierten Ausleseverfahren zu wenig gelingt, eine valide Zuordnung zum dreigliedrigen Schulsystem zu erreichen.[54] Die entwickelten

[54] Das aus Teilen der Politik und dem Praxisfeld Schule zu vernehmende Argument, solcherlei Testverfahren berücksichtigen zu wenig die ökologische Validität der Schulen, überzeugt vor dem Hintergrund der IGLU-Ergebnisse nur wenig. Umgekehrt ist der Fingerzeig auf Schulen und Lehrer

Kompetenzstufen und das schulische Notenraster verlaufen zwar nicht völlig parallel zueinander, bemerkenswert ist allerdings, dass 8% der Schüler, die in der IGLU-Studie nur die Kompetenzstufe II erreichten, die Deutschnote 1 erhielten, während 11% der Schüler mit der Deutschnote 5 die Kompetenzstufe III erreichten. Bedenkt man, welch weit reichende Folgen die Schulempfehlungen für individuelle Bildungs- und Berufskarrieren haben können, so ergibt sich einige Evidenz dafür, Schüler erst zu einem späteren Zeitpunkt in unterschiedliche Bildungsgänge zu separieren. Gerade in der Grundschule sind bisher übliche Leistungsbewertungsverfahren stärker noch als in der Sekundarstufe von der sozialen Herkunft abhängig. Das Problem der Diagnostik in schulischen Handlungsfeldern würde mit einer späteren Separierung allerdings nicht automatisch gelöst. Helmke (2003, S. 88) wirft die Frage auf, inwieweit die Orientierung an psychometrischen Gütekriterien überhaupt für Lehrkräfte gelten solle (könne). Mit Weinert & Schrader (1986, S. 18f.) kommt er zu einer skeptischen Einschätzung. Setzt man voraus, dass der in unterrichtlichen Prozessen Urteilende sich der Ungenauigkeit, Vorläufigkeit und Revisionsbedürftigkeit seiner Urteile bewusst ist und sie deshalb permanent überprüft, wäre nicht eine Zustands-, sondern eine Verlaufsdiagnostik anzustreben.

Die insgesamt erfreulichen Ergebnisse der IGLU-Studie führen zu der Frage, weshalb die in der Grundschule angelegten Potentiale von den weiterführenden Schulen, auch wenn sich hier Konfrontationen mit neuen Problemstellungen etwa disziplinärer Art ergeben, anscheinend wenig genutzt werden.

7.1.4 MARKUS (Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext)

Im Rahmen der Studie wurde eine Vollerhebung der achten Jahrgangsstufe im Bundesland Rheinland-Pfalz mit Blick auf die Mathematikleistungen durchgeführt. In den Fokus gerieten Unterrichtsmerkmale des Mathematikunterrichts sowie Lernvoraussetzungen und persönliche Hintergründe der Schüler (Helmke & Jäger 2002). Dabei legen die Autoren Wert darauf, dass es sich um kein Forschungsprojekt handelt, sondern um ein wissenschaftlich fundiertes Evaluationsprojekt zur Standortbestimmung von Schulen (ebd., S. 5f.). Entsprechend stehen erhobene Daten dem Qualitätsmanagement der Schulen vor Ort zur Verfügung.
Die Studie hat umfassende Ergebnisse geliefert. Erwartungsgemäß stellen sich Leistungen zwischen den Bildungsgängen Hauptschule, Realschule und Gymnasium signifikant unterschiedlich dar. Schularten mit mehreren Bildungsgängen (Gesamtschulen) zeugen im separaten Vergleich der Bildungsgänge von ähnlichen Leistungen wie jeweils entsprechende Schulen mit nur einem Bildungsgang.
Dabei wirken sich Bildungsgänge der weiterführenden Schulen im Vergleich zur Grundschule leistungshomogenisierend aus, so dass sozioökonomische oder soziokulturelle familiäre Hintergründe einen geringeren Einfluss auf Leistungsvariablen ausüben. Dieses Ergebnis kann in zwei Richtungen gedeutet werden. Positiv fällt auf, dass Bildungsgänge Schüler mit und ohne

kaum weiterführend. Ziel müsste es sein, Verfahren der pädagogischen Diagnostik weiterzuentwickeln und sie den Schulpraktikern an die Hand zu geben sowie sie in die Lehrerbildung zu implementieren.

Migrationshintergrund offenbar gleich fördern. Kritisch kann damit angemerkt werden, dass eigentliche bildungsgangsrelevante Entscheidungen am Ende der Grundschule fallen. Schüler nicht-deutscher Sprachherkunft sind in den anspruchsvolleren Bildungsgängen deutlich unterrepräsentiert. Kompensatorische Maßnahmen zum Ausgleich solcher Disparitäten wären an den Grundschulen zu verorten. An den Hauptschulen gleichen sie bisweilen einer Symptombehandlung.
Auch in der Markus-Studie wurden wie für ältere Studien nur minimale Auswirkungen der Klassengröße auf Leistungen sowie den Unterricht festgestellt.[55] Ein weiterer Befund wirft ein insgesamt kritisches Licht auf den Mathematikunterricht speziell an Realschulen und Gymnasien. Während 40% der Hauptschüler (absolut betrachtet kein hoher Wert) der Aussage zustimmen, der Mathematikunterricht wäre nützlich für den Alltag, liegt der Anteil von Positivantworten bei den Realschülern bzw. Gymnasiasten bei nur 20% bzw. 16%. Einhergehend mit solchen Einschätzungen kann eine niedrige fachbezogene Lernmotivation beobachtet werden. Gerade den Gymnasien gelingt es also offensichtlich zu wenig einen Alltags- oder Lebensweltbezug herzustellen.
Schließlich legt die Studie bestimmte Merkmale offen, die zu Störungen des Unterrichts führen. So stellen Lehrpersonen einen Zusammenhang von übermäßigem TV- und Videokonsum und einer hohen Unterrichtsbeeinträchtigung heraus. Auf Leistungsbeeinträchtigungen schließen sie jedoch nur bei extremen Mediennutzungsgewohnheiten. Als weitere Gründe für Unterrichtsstörungen nennen befragte Lehrer Desinteresse, Begabungsunterschiede innerhalb der Klasse und störende Schüler. Insofern überrascht es nicht, dass die Studie die große Bedeutung einer effizienten Klassenführung über alle Bildungsgänge hinweg betont. Helmke (2003, S. 138) stellt aber strukturelle Unterschiede der Bildungsgänge hinsichtlich der Umsetzung von Klassenführungsansprüchen heraus:

„Die Klassenführung ist in den Spitzengruppen aller Bildungsgänge überdurchschnittlich effizient. Darüber hinaus lässt sich die Spitzengruppe in der Hauptschule wie folgt charakterisieren: entwickelte Aufgabenkultur, große Methodenvielfalt. Bei der Spitzengruppe in der Realschule kommen hinzu: anspruchsvolles Üben, ein hohes Anspruchsniveau, verbunden mit mehr Kleingruppenarbeit. Im Gymnasium sind die Spitzenreiter charakterisierbar durch entwickelte Aufgabenkultur und ein hohes Anspruchsniveau."

Erkennbar sind notwendige Unterscheidungen von Effizienz fördernden Klassenführungselementen je nach Profil der einzelnen Bildungsgänge. Gleichzeitig zeugt die vergleichende Betrachtung von weit reichenden Überschneidungen.

[55] In der Studie wird die interessante These vertreten, dass sich die nur geringe Auswirkung (nach Plausibilitätsargumenten wäre eine größere Bedeutung vorstellbar) unterschiedlicher Klassengrößen auf das Leistungsverhalten durch eine Nivellierung bzw. Kompensierung von größenbezogenen Klassenunterschieden auf der Schulebene erklärt. Fähigeren Lehrern würden demnach größere und von der Schülerzusammensetzung schwierigere Klassen zugewiesen. Die höheren Anforderungen durch größere Klassen würden durch vermehrte Anstrengungen in der Vorbereitung kompensiert.

7.1.5 PISA (Programme for International Student Assessment)

Die PISA-Studie wurde von der OECD als Längsschnitt mit einer Laufzeit von 1998 bis 2007 initiiert, wobei eine Fortführung der Studie über den Zeitraum hinaus als ausgemacht gilt. Ziel war die detaillierte und ländervergleichende Untersuchung von Lernprozessen bei 15jährigen Schülern. Die erste Untersuchungsreihe (PISA 2000) legte ihren Schwerpunkt auf das Leseverständnis (es wurden aber auch Kompetenzen im Fach Mathematik und den Naturwissenschaften untersucht) und wurde federführend vom Max-Planck-Institut für Bildungsforschung (Berlin) durchgeführt (Baumert et al. 2001). Die zweite Untersuchungsreihe (PISA 2003) stellte das Fach Mathematik in den Mittelpunkt und stand unter der federführenden Verantwortlichkeit des Instituts für die Pädagogik der Naturwissenschaften (Kiel). Schwerpunkt in der wiederum unter der Federführung des IPN gestarteten dritten Untersuchungsreihe (PISA 2006) liegt schließlich auf den Naturwissenschaften.[56]
Die PISA-Studie wurde an vielen Stellen ausführlich dargestellt und soll hier nur mit Bezug auf das Design sowie zentrale Ergebnisse diskutiert werden, die aufschlussreich für die eigene Untersuchungsrichtung sind. Die Grundfragestellung lautete, inwieweit die nachwachsende Generation über kulturelle Basiskompetenzen, im Sinne des anglo-amerikanischen „literacy"- Konzepts, verfügt. Baumert & Prenzel (2002, S. 98) spezifizieren, wie die Kompetenzen in der PISA-Studie 2000 modelliert worden sind:

- Lesekompetenz, d.h. die Fähigkeit, mit authentischen Texten unterschiedlicher Art verständig umgehen zu können,
- mathematische Modellierungsfähigkeit und
- die Verfügbarkeit basaler naturwissenschaftlicher Konzepte sowie ein Verständnis über Prinzipien naturwissenschaftlichen Arbeitens und Argumentierens.

PISA-2000 umfasste eine repräsentative Stichprobe der 15jährigen Schüler als Individualstichproben. Damit folgte zugleich ein Verzicht auf Klassenstichproben.[57]
Die Schüler wurden zu folgenden Unterrichtsdimensionen befragt:

- Leistungserwartungen
- Lehrerunterstützung
- Disziplinprobleme
- Klarheit und Regeltreue
- Überforderung
- Individuelle Bezugsnormorientierung
- Anspruchsvolles Üben

[56] Die Ergebnisse von PISA 2006 werden aktuell vorgestellt, konnten aber in der Arbeit nicht mehr berücksichtigt werden.
[57] Da keine Klassen, sondern Individuen untersucht wurden, eignet sich PISA 2000 nur bedingt dafür, Aussagen über Unterrichtsqualität zu leisten (Helmke 2003, S. 140f.). Zudem wurden zwar Schulleitungen, aber keine betroffenen Lehrkräfte befragt.

Die Untersuchung der Unterrichtsdimensionen ergab hinsichtlich der Auswirkungen auf Schülerleistungen nur signifikante Werte der Überforderungsdimension. Je stärker sich Schüler überfordert fühlen, desto geringere Leistungen erbringen sie. In den anderen Dimensionen wurden keine übergreifenden signifikanten Zusammenhänge gefunden. Für die Dimension „Lehrerunterstützung“ fällt allerdings auf, dass Schüler an Gymnasien – im Vergleich zu Probanden anderer Schularten – eine signifikant geringere Lehrerunterstützung wahrnahmen.
Die PISA-Studie hat eine verbreitete Popularität erlangt, weil die Ergebnisse für die Öffentlichkeit, Politik und die Schulen selbst von überraschenden Fakten zeugten. Mit Blick auf die bereits früher durchgeführte erste TIMMS-Studie, die jedoch in geringerem Ausmaß zur Kenntnis genommen wurde, überraschen die Ergebnisse weniger. Deutschland nimmt in den Bereichen Lesekompetenz und Mathematik bzw. Naturwissenschaften nur einen Mittelplatz ein. Dieser nun weit geteilte und empirisch belegte mittlere Rangplatz erklärt die außerordentlich hohe und übergreifende Resonanz auf die PISA-Ergebnisse. Als allgemein besorgniserregend wurde insbesondere der hohe Zusammenhang von Lesekompetenz und der Zugehörigkeit zur Sozialschicht kommuniziert. Im Kontext von Mediatisierungsprozessen gibt es in Deutschland relativ breite soziale Schichten, die sich durch Bildungsferne auszeichnen. Kinder aus diesen Schichten werden unzureichend an Literatur herangeführt, was in mangelnder Lesekompetenz Ausdruck findet. Der besagte Zusammenhang ist in keinem untersuchten Land so groß wie in Deutschland. Dementsprechend stellen Kinder aus unteren Sozialschichten eine Risikogruppe dar. Ähnliches lässt sich für Kinder mit Migrationshintergrund sagen. Hieran knüpft die Kritik der PISA-Studie mit der Benennung von diagnostischen Defiziten bei Lehrpersonen zur Identifizierung solcher Risikogruppen an.

7.1.6 TIMMS (Third International Mathematics and Science Study)

TIMMS ist die dritte Fortführung einer von der IEA initiierten Untersuchungsreihe für die Bereiche Mathematik und Naturwissenschaften.[58] Um Entwicklungstendenzen sichtbar zu machen, wurde die Studie auf einen vierjährigen Turnus abgestellt; Hauptuntersuchungen fanden 1995, 1999 und 2003 statt, wobei Deutschland an der 1999 durchgeführten TIMMS-R („Repeat“) Studie und auch 2003 nicht teilgenommen hat. Übergreifend umfasst TIMMS drei Populationen, die bezüglich ihrer Alters- bzw. Schulstufe unterschieden werden:

- TIMMS/I bezieht sich auf die vierte Klassenstufe, in Deutschland also das Ende der Grundschule
- TIMMS/II bezieht sich auf die achte Klassenstufe, in Deutschland also die Sekundarstufe I

[58] Die in Deutschland weniger bekannten Vorgängerstudien behandelten die Bereiche Mathematik und Naturwissenschaften zunächst getrennt voneinander: FIMS (First International Mathematics Study), FISS (First International Science Study), SIMS (Second International Mathematics Study) und SISS (Second International Science Study). An diesen Studien hat sich Deutschland allerdings nicht beteiligt.

- TIMMS/III bezieht sich auf die Leistungen in der Sekundarstufe II, in Deutschland also auf das letzte Jahr des allgemeinbildenden Schulwesens

Methodisch wurden verschiedene Instrumente in mehreren Teilstudien eingesetzt, deren Ergebnisse quantitativ kaum noch zu überblicken sind. Neben Leistungsstudien zur Untersuchung von Dimensionen unterrichtlicher Lernprozesse wurden Befragungen von Schulleitern und Fachlehrern, Curriculums- sowie Videostudien durchgeführt. Die unterrichtsrelevanten Ergebnisse werden hier nur in ihren zentralen Tendenzen und mit Bezug auf das eigene Frageinteresse wiedergegeben.
Erstmals 1997 veröffentlichte Befunde von TIMMS (Baumert et al. 1997) lösten in der Fachwelt und z.T. auch in der allgemeinen Öffentlichkeit starke Reaktionen aus, so dass, in Anlehnung an den „Sputnik-Schock“ der 1960er Jahre, vom „TIMMS-Schock“ gesprochen wurde (Kohler 2002, S. 158). Eine Feststellung lautete, dass deutsche Schüler in den Bereichen Mathematik und Naturwissenschaften durchaus Stärken aufweisen, wenn es um Faktenwissen und Routineaufgaben geht. Sind zur Lösungsfindung allerdings tiefer gehende Reflexions- und Denkprozesse erforderlich, welche die Anwendung erworbenen Wissens verlangen oder Aufgabenstellungen, die insgesamt auf Problemlösekompetenzen abzielen, demonstrieren die Schüler im Bewältigungsverhalten Schwächen (Baumert et al. 1997; Baumert, Bos & Lehmann 2000a, b; Seidel 2003a).
Eine genauere Deutung dieser Ergebnislage setzt an der Analyse unterrichtlicher Lehr-Lernprozesse an. Die TIMMS-Videostudie (Stigler, Gallimore & Hiebert 2000; Stigler et al. 1999; Stiegler & Hiebert 1997), die einen transkulturellen Vergleich zwischen den USA, Japan und Deutschland durchführte, belegt jeweils typische kulturelle Skripts, die als Determinanten des didaktischen Ablaufs von Unterrichtsstunden fungieren. Für Deutschland konnte als charakteristisch vorfindbares Verhalten von Lehrpersonen das „fragend entwickelnde Unterrichtsgespräch“ identifiziert werden. Ein solcher Unterrichtsstil, einhergehend mit einer schematischen Engführung auf einzig richtige Lösungen, so kann angenommen werden, führt wiederum zu Problemen deutscher Schüler bezüglich eigenständiger Lösungsfindung im Rahmen anspruchsvoller Aufgabendefinitionen. Insgesamt ist als Ergebnis der TIMMS-Videostudie festzuhalten, dass der untersuchte Unterricht erstaunlich gleichförmig und aus Schülersicht wenig variantenreich erscheint. Konkret ging der transkulturelle Vergleich folgenden drei Fragestellungen nach (Stigler et al. 1999):

- Wie werden Mathematikstunden strukturiert und organisiert?
- Wie hochwertig ist der mathematische Inhalte?
- Welche Denkprozesse werden bei den Schülern angeregt?

Als Ergebnis der Analysen komprimierten Baumert et al. (1997, S. 225f.) typische Unterrichtsskripts für Japan und Deutschland.

Mathematikunterrichtsskript in Japan:

- Der Lehrer stellt der Klasse ein komplexes und kognitiv anspruchsvolles mathematisches Problem vor, dessen Lösung nicht evident ist.
- Die Schüler arbeiten einzeln oder in Gruppen- oder Partnerarbeit an der Problemlösung.
- Verschiedene Schüler, die vom Lehrer aufgrund der eingeschlagenen unterschiedlichen Lösungswege ausgewählt werden, präsentieren der Klasse alternative Aufgabenlösungen und präsentieren sie an der Tafel.
- Im Unterrichtsgespräch werden die verschiedenen Lösungswege diskutiert.
- Der Lehrer fasst die Ergebnisse des Unterrichtsgesprächs in einem kurzen Lehrervortrag zusammen und notiert sie an der Tafel.
- Die Schüler bearbeiten in Einzel- oder Gruppenarbeit ähnliche, aber auf neue Anwendungskontexte übertragene Aufgaben.

Mathematikunterrichtsskript in Deutschland:

- Die Stunde beginnt mit der Durchsicht und Besprechung der Hausarbeiten.
- Es folgt eine kurze Wiederholungsphase.
- Variante 1: Der neue mathematische Stoff wird im fragend entwickelnden Unterrichtsgespräch, das auf eine einzige Lösung hinführt, relativ kurzschrittig erarbeitet und vom Lehrer an der Tafel dokumentiert.
- Variante 2: Wenn das Thema schon in der vorangegangenen Stunde vorbereitet wurde, entwickelt ein Schüler – unterstützt von der Klasse und dem Lehrer – einen aufgabenbezogenen Lösungsweg an der Tafel.
- Zur Einübung des Verfahrens lösen Schüler in Stillarbeit ähnliche Aufgaben.

Vor diesem Hintergrund erklären sich auch Befunde zum Physikunterricht durch die TIMSS-Studie. So geben 56% der befragten deutschen Schüler im ersten Turnus an, sie hätten „in den meisten Stunden oder allen Stunden der Lehrkraft zuzusehen oder zuzuhören, wenn sie physikalische Konzepte entwickelt“ (Baumert & Köller 2000, S. 284). Der fragend entwickelnde Unterricht, so das Befragungsergebnis, findet in den meisten Stunden durch Demonstrationsexperimente eine Unterstützung (89% der Befragten), während von den Schülern durchgeführte Experimente fast nie vorkommen (51% der Befragten).
Baumert und Köller (2000) identifizieren in globaler Perspektive drei mögliche Großformen der Unterrichtsführung im Physikunterricht:

- Verständnisorientierter Unterricht: Modellgeleitete anspruchsvolle Aufgabenstellungen; ein vom Phänomen ausgehendes, induktives, aber auf Systematisierung zielendes Vorgehen sowie die Anwendung naturwissenschaftlicher Konzepte auf Alltagsprobleme.
- Unterricht auf der Basis von Schülerexperimenten: Der Schwerpunkt liegt auf der Durchführung von gelenkten Schülerexperimenten.
- Demonstrationsunterricht: Anhand von Vorführexperimenten wird ein physikalisches Konzept entwickelt; die Schüler übernehmen von der Tafel Notizen; Kombination von Lehrerexperimenten und fragend entwickelndem Unterrichtsgespräch mit starker Lehrerlenkung (Seidel 2003a, S. 41f.).

Im deutschen Physikunterricht überwiegt der Demonstrationsunterricht. Auch damit wird deutlich, weshalb deutsche Schüler anspruchsvolle Aufgaben nur teilweise eigenständig lösen können. Die Reduktion komplexer mathematisch-physikalischer Modelle auf kleine Wissenseinheiten, die von Schülern in kleinschrittigen Verfahren des fragend entwickelnden Unterrichts nachvollzogen werden sollen, vermindert nach Baumert (1998) die Chance, dass Schüler übergreifende Konzepte aufbauen und vernetzte Denkstrukturen entwickeln. Zudem erhalten sie nur wenig Raum, um eigenen Fragestellungen nachzugehen.
Was Ergebnisse zu dem Fähigkeitsniveau von Schülern im Mathematik- und Physikunterricht betrifft, erfolgt an dieser Stelle eine Konzentration auf Daten zur gymnasialen Oberstufe. Erwartungsgemäß können deutliche Unterschiede des Fähigkeitsniveaus zwischen Grund- und Leistungskursen der Fächer Mathematik und Physik ausgemacht werden. In den mathematischen Grundkursen bleiben 60% der untersuchten Schüler auf einem Niveau der Anwendung elementarer Konzepte und Regeln verhaftet. Bei Änderungen des Kontextes von Aufgabenstellungen zeigen Schüler erhebliche Probleme in der Aufgabenlösung. Zwar fällt das Leistungsniveau in Leistungskursen deutlich höher aus, auch hier können aber nur 10% der Schüler Aufgaben erfolgreich bewältigen, deren Lösungen nicht unmittelbar evident sind. Auch in den Physikgrundkursen kann beobachtet werden, dass ein großer Anteil der Schüler erhebliche Schwierigkeiten damit hat, bekannte Aufgabentypen in veränderten Beschreibungskontexten zu bewältigen. Dagegen erreichen in den Physikleistungskursen 25% der Schüler ein Fähigkeitsniveau, das es ihnen erlaubt, fachlich selbständig zu argumentieren.[59] Im internationalen Vergleich steigt der Abstand zu den leistungsfähigeren Nationen mit zunehmender Aufgabenschwierigkeit. Im Verhältnis der Geschlechter können deutliche Leistungsunterschiede zu ungunsten der Mädchen ausgemacht werden, die sich mit zunehmendem Alter (s. TIMMS/II, TIMMS/III im Vergleich zu TIMSS/I) noch vergrößern.

Die in der TIMMS-Studie herausgestellten Ergebnisse, so das bildundungspolitische Fazit, können einem Land nicht genügen, dessen Wirtschaftskraft auf wissensbasierten Innovations- und Industrialisierungsprozessen basiert. Für die gymnasiale Oberstufe fällt besonders das schlechte Abschneiden im Fach Mathematik in den Blick. Der Mathematik- und Physikunterricht der gymnasialen Oberstufe weist in zweifacher Richtung Defizite auf. Zum einen misslingt eine anspruchsvolle fachspezifische Ausbildung für eine breite Gruppe der Schülerschaft. Zum anderen findet nur eine kleine Gruppe der Schüler den Anschluss an international vergleichbare Spitzenleistungen. Ausprägungen in die ein oder andere Richtung könnten erklären, warum entweder eine breite Schülerschaft ein nur unteres bzw. mittleres Fachniveau erreicht bzw. warum fachspezifische Spitzenleistungen nicht homogen bei den Schülern auffindbar sind. Beide Erklärungsmuster für die Zielsetzungen von Unterricht – entweder die Orientierung auf breite Förderung oder die Konzentration auf Eliten – bildeten die TIMMS-Ergebnisse jedoch nicht ab.

[59] Im Vergleich der Mathematik- und Physikleistungskurse kann herausgestellt werden, dass die Physikkurse stärker positiv vorselektiert werden und im Regelfall eine kleinere Teilnehmerzahl aufweisen. Beides wirkt sich förderlich auf Leistungsverhalten aus.

Nachdem im Vorangehenden empirische Studien zu Schülerleistungen in ihren mittelbaren Aussagen für die Wirksamkeit von unterrichtlichem Lehrerhandeln reflektiert wurden, behandelt der nächste Abschnitt Untersuchungen, die explizit und damit unmittelbaren Aufschluss über Wirksamkeitsdimensionen der Lehrerbildung – in internationaler und auf Deutschland bezogener nationaler Perspektive – erlauben.

7.2 Wirksamkeit der Lehrerbildungssysteme

Blömeke (2004) liefert einen umfassenden Überblick über die empirischen Befunde zur Wirksamkeit der Lehrerbildung. Dabei hebt sie hervor, dass die durchaus nicht unumstrittene Variablenauswahl zur Wirksamkeitsuntersuchung, die in Abhängigkeit von einem zugrunde gelegten Modell des Zusammenhangs zwischen Lehrerbildung, Lehrerhandlung und Schülerleistungen erfolgt, eine zentrale Herausforderung darstellt. Innerhalb der Lehrerbildung lassen sich verschiedene Prozesselemente, z.B. die Inhalte der Lehrangebote, ihre didaktisch-methodische Präsentation, der Grad der Strukturierung der Ausbildung, die Abfolge ihrer Phasen, Ort oder Länge der Ausbildung etc. als mögliche Wirkfaktoren identifizieren. Welche dieser Faktoren sich allerdings auf das spätere Lehrerhandeln positiv auswirken, ist bisher nicht eindeutig beantwortet worden. Letztendlich erfolgt sogar die Auswahl der positiven Wirkkriterien nicht konsensual. Blömeke (ebd., S. 61) stellt drei mögliche Herangehensweisen für die Untersuchung von Wirksamkeitskriterien vor:

- eine Erfassung der fach- und berufsbezogenen Kognitionen von Lehrenden
- die Beobachtung des Lehrerhandelns im Unterricht
- die Erfassung der Leistung von Schülern

Unter Berücksichtigung des professionstheoretischen Diskurses schlägt sie zudem eine Einbeziehung des Ethos von Lehrenden als zusätzlich zu untersuchendes Kriterium vor. Die von Blömeke herausgearbeiteten Wirksamkeitskriterien der Lehrerbildung setzen aber die These voraus, dass die Lehrerbildung überhaupt untersuchbare Wirkungen generiert und die Kriterien von der Lehrerbildung auch intendiert sind. Beide Erwartungen können nicht voraussetzungslos gesetzt werden. Gleichwohl liefern Blömekes Wirksamkeitsattributionen wichtige Impulse für die Zielrichtung empirischer Forschung. In der hier vorliegenden Studie erfolgt eine Beschränkung auf die Untersuchung des unterrichtlichen Handelns von angehenden Lehrpersonen.
Für die Bundesrepublik Deutschland lässt sich generell ein Defizit an empirischer Forschung zur Wirksamkeit der Lehrerbildung feststellen. Im Folgenden werden die Ergebnisse der Untersuchung der Wirksamkeit des schweizerischen Lehrerbildungssystems vorgestellt (Oser & Oelkers 2001), um anschließend die Befunde der diesbezüglich profilierteren Diskussion in den Vereinigten Staaten von Amerika aufzuzeigen und am Ende Beispiele der empirischen Forschung in Deutschland zu diskutieren.

7.2.1 Untersuchungen zur Wirksamkeit des Lehrerbildungssystems in der Schweiz

Oser (2001, S. 67) stellt fest, dass die in seiner Studie angestrebte Untersuchung der Wirksamkeit des Lehrerbildungssystems vor dem grundlegenden Problem steht, die Qualität der Ausbildung als eine noch unbekannte Größe fassen zu müssen.[60]

„Die Überprüfung der Qualität der Ausbildung steht aber vor dem Problem, dass genau diese Qualität eine noch unbekannte Größe darstellt. Eigentlich müssten wir im Sinne von Antezedentien der Wirkungsattribution im voraus entscheiden, was wir unter Qualität verstehen, um sie überprüfen zu können. Aber dies ist deshalb ein schwieriges Unterfangen, weil es für die Profession der Lehrperson noch kein festes, international anerkanntes und als absolut notwendig betrachtetes Set von Kompetenzen gibt. Desweiteren wandeln sich die gesellschaftlichen Erfordernisse an der Schule andauernd, was zu neuen Forderungen hinsichtlich der professionellen Ausbildung führt."

Die Überprüfung der Wirksamkeit der Lehrerbildung ist nach Oser von vier zu klärenden Merkmalsbereichen abhängig (ebd., S. 68-70):

- Offen zu legen sind Qualitätskriterien, die eine Bestimmung dahingehend zulassen, ob eine Lehrerbildung (auch aus Sicht der Absolventen) qualitativ „gut" ist.
- Das in der Lehrbildung formulierte notwendige Sach- und Handlungswissen muss einer empirischen Überprüfung standhalten.
- Normative Aussagen der Lehrerbildung sind darauf zu überprüfen, inwieweit sie rein rhetorischen Charakter haben.
- Erforderlich ist die Bestimmung eines übergreifenden Wirkungsmodells für die Lehrerbildung.

In Anlehnung an Scheerens & Bosker (1997) identifiziert und diskutiert Oser (2001, S. 71) zunächst noch sieben Modelle, die allerdings nicht orthogonal zueinander stehen, sondern z.T. Überschneidungen aufweisen.

1. Das „Unser-Ziel-Modell" der Wirksamkeit
2. Das Effektivitätsmodell der Wirksamkeit (what works)
3. Das Andachtsmodell der Wirksamkeit (Schulkultur und Sozialklima)
4. Das Organisationsmodell der Wirksamkeit

[60] Vor dem Hintergrund dieser Feststellung wird das bereits in Kap. 5 kritisierte Zusammenfallen der formulierten Ausbildungsziele mit den Instrumenten zur Überprüfung derselben in Form von Standarddefinitionen deutlich. Oser (2001, S. 67) hebt hervor, dass „im Sinne von Antezedentien der Wirkungsattribution" im vorhinein entschieden werden muss, was man unter Qualität versteht, um sie überprüfen zu können. Genau diesen Weg ist er, wenn auch über den Umweg einer Quasi-Delphi-Studie, gegangen. Seine Untersuchung hätte aber vermutlich nicht so viel Furore gemacht, wenn nicht gleichzeitig viele an der Lehrerbildung beteiligte Akteure der Meinung gewesen wären, dass eine Reformierung ansteht. Die tumultartige Resonanz der PISA-Ergebnisse steht für ein ähnliches Phänomen. Auf der Systemebene lassen sich bestimmte Zeitpunkte beobachten, an denen notwendige Korrekturen plötzlich alternativlos erscheinen. Bei der Rekonstruktion solcher Prozesse stellt sich die Frage, warum zum einen notwendige Umsteuerungsaktionen nicht schon früher initiiert worden sind und warum zum anderen plötzlich bestimmte, immer auch anzweifelbare Ergebnisse eine derartige Dynamik entfalten.

5. Das Ökonomie-Modell der Wirksamkeit
6. Das Defizit- bzw. Kompensationsmodell der Wirksamkeit
7. Das Professionsgenerierungs-Modell der Wirksamkeit

Das erste steht dem sechsten Modell diametral gegenüber. Während erstes eine Überprüfung von Wirksamkeitskriterien lediglich an den selbst formulierten Zielen erlaubt, trägt letztes eine Defizitperspektive an zu evaluierende Systeme heran, indem die Forderung auf eine gesteigerte systemische Leistungsfähigkeit zielt. Damit können jedoch Überforderungstendenzen auf Adressatenseite verbunden sein.[61] Das sog. Andachtsmodell der Wirksamkeit führt nach Oser zu einer Therapeutisierung der Lehrerbildung und fokussiert zu stark die Frage des Wohlbehagens anstatt der Leistung. Damit können seiner Meinung nach Professionalisierungsprozesse nur schwerlich gelingen, wenn man die schulische Praxis einseitig idealistisch mit einem guten Klima und weniger mit vielfältigen Spannungsverhältnissen assoziiert. Oser stellt die These auf, dass dieses Modell unterschwellig in vielen Lehrerbildungsinstitutionen Verwendung findet, obwohl ein gutes kommunikatives Klima nicht per se zu einer Leistungsverbesserung führt. Mit dieser Polemik, die sich schon in der quasireligiösen Begrifflichkeit „Andacht" ausdrückt, nimmt Oser jedoch eine einseitige Verengung des Klimabegriffes vor. Nimmt man den von ihm geforderten Praxisbezug in der Lehrerbildung ernst, so gibt es gute Gründe dafür, die vielfältigen Ergebnisse der Schulklimaforschung auch in der Lehrerbildung zu berücksichtigen. Fend (1977, 1998, 2000) konnte aufzeigen, dass die von Schülern wahrgenommenen Lernumweltmerkmale einen wichtigen sozialisatorischen Einfluss nehmen können und eine positive Lehrer-Schüler-Beziehung eine relevante Dimension für Lernmotivation darstellen kann. Die Schulklimaforschung hat zudem darauf aufmerksam gemacht, Schule insgesamt als pädagogische Handlungseinheit und nicht nur den Unterricht als Ursprung gelingender Lernprozesse zu betrachten (Fend 1988). Die Bedeutsamkeit eines lernförderlichen Klimas in der Unterrichtsgestaltung wurde von Meyer (2004, S. 47) herausgestellt. Insofern verwundert es, wenn Oser Merkmale einer positiven Schulkultur und des Sozialklimas einseitig negativ und leistungsfeindlich attribuiert.
Einseitige Schwerpunktsetzungen auf Fragen der Effektivität (Modell 2) bzw. Ökonomie (Modell 5) werden einer Evaluation der Lehrerbildung in Osers Perspektive ebenfalls nicht gerecht. So kann Lehrerhandeln und -wissen in einem unverkürzten Modell nicht pauschal vermeintlichen Zielkriterien, die zudem teilweise unklare Zuschreibungen aufweisen und nicht-lineare Wirkungsketten beschreiben, funktional nachgeordnet sein. Gegen die ökonomische Vorrangstellung im Sinne der Einengung von Lehrerhandeln auf wirtschaftliche Rentabilitätsaspekte und die damit verbundene Konzentration auf Ressourcenuntersuchungen und Outputorientierung,

[61] Oser ordnet z.B. die TIMMS-Studie diesem Modell zu und postuliert, dass die internationalen Leistungsstudien der IEA nicht zuletzt intendieren, über internationale Vergleichsdokumentationen nationale Defizitgefühle aufkommen zu lassen. Aus bildungspolitischer Perspektive trifft Osers Kritik in Teilen zu, wenn damit der Erkenntniswert der Studie auch nicht geschmälert ist. Auch Osers (2001) Studie kann aus der Perspektive des schweizerischen Lehrerbildungssystems als Defizitmodell wahrgenommen werden, trotz dem der Autor selbst von einem „Professions-Generierungs-Modell" spricht. Implizit und explizit wird damit aber ausgesagt, dass die bisherige Lehrerbildung in der Schweiz zu wenig professionell entwickelt und damit defizitär war. In seiner Studie legt er zudem nicht nur Defizite offen, sondern beschreibt erreichbare Verbesserungen auf der Systemebene sowie damit verbundene bildungspolitische Implikationen.

der auch Oser zuneigt, können ähnliche Argumente vorgebracht werden. Legt man einen Schwerpunkt auf die Frage der Organisation von Ausbildungsstätten (Modell 4), rücken Aspekte der institutionellen Organisation reibungsloser Kommunikation und Interaktion, der gegenseitigen Kontrolle oder der Teamarbeit in den Mittelpunkt. Bätz und Oser (1997) konnten allerdings nachweisen, dass Lehrpersonen, die sich für die Profilbildung ihrer Schulen engagieren, damit Organisationsentwicklung betreiben und den Stellenwert von arbeitsbezogenen und kommunikativen Teamprozessen zu erhöhen intendieren, Gefahr laufen, der Kommunikation mit den Kindern fortan weniger gerecht zu werden. Dieser herausgestellte Gefahrenpunkt kann jedoch nicht als unabwendbarer Automatismus gewertet werden. Problematisierend ist in diesem Zusammenhang anzumahnen, dass in den schweizerischen und deutschen Schulen zu wenige Funktionsstellen vorgesehen sind und diese Stellen zudem nicht selten von unqualifiziertem Personal besetzt werden. Schulische Organisations- und Personalentwicklung avanciert damit zu einem semiprofessionellen Unterfangen oder schlimmstenfalls zu einem „Hobby“ von Funktionsträgern. Dass hier Überforderungstendenzen vorgezeichnet sind, kann man nur zum Teil einzelnen Personen anlasten. Oftmals sind Defizite zunächst in der Systemstruktur angelegt.

Die Lehrerausbildung ist z.T. (so sie an Universitäten und nicht an Pädagogischen Hochschulen angesiedelt ist) in die Gesamtorganisation der Hochschule eingebunden, was eine genuine Evaluation erschwert. In der doch recht kritischen Diskussion von Oser wird deutlich, dass jedes Evaluationsmodell spezifische Vorteile mit sich bringt. Dabei können Wirksamkeitsziele der Lehrerbildung nur dann sinnvoll untersucht werden, wenn sie von beteiligten Institutionen getragen werden. So besteht heute allgemeiner Konsens darüber, dass eine moderne Lehrerbildung sich Ansätzen pragmatischer Effektivitätsstrategien, ökonomischer Ressourcenverteilung und also einer Outputperspektive nicht entziehen kann. Angehende Lehrerinnen und Lehrer sollen ihre Ausbildung in einem angenehmen Sozialklima erleben, gleichzeitig aber vor den Auswüchsen einer „Kuschelpädagogik“ geschützt werden. Die Lehrerausbildung ist Teil einer kontinuierlich zu verbessernden und weiterzuentwickelnden Organisation. Jede Organisation ist zudem Vergleichen, Kritik und externen Anregungen ausgesetzt, begibt sie sich in offene Entwicklungsprozesse. Die von Oser diskutierten Modelle teilen sich in diesem Zusammenhang den Nachteil, dass sie jeweils bestimmten Momenten im Sinne von Alleinstellungsmerkmalen einen überhöhten Stellenwert beimessen. Übergreifende Besonderheiten der Lehrerbildung finden so nur ungenügende Berücksichtigung.[62]

In seiner Diskussion der Wirksamkeitsmodelle plädiert Oser schließlich für das Professionsgenerierungs-Modell der Wirksamkeit. Im Mittelpunkt steht hier die Berücksichtigung späterer Anforderungssituationen für die Entwicklung

[62] In Osers Argumentation fällt allerdings auf, dass er Kritikpunkte an den Modellen uneinheitlich z.T. von den Erfordernissen der Ausbildungsinstitutionen und z.T. von den praktischen Erfordernissen der Schulen her denkt. Es leuchtet ein, dass Ausbildungsinstitutionen Erfordernisse der Schulen vor Ort mit bedenken müssen und sich damit Überschneidungen institutioneller Positionierungen zu Wirksamkeitserwartungen ergeben. Was die Angemessenheit von Evaluationsmodellen der Lehrerbildung betrifft, muss die Diskussion aber jeweils auf die spezielle Ausbildungssituation einerseits bzw. auf die Berufssituation andererseits bezogen bleiben, gilt es also, institutionenspezifische Bedingungskontexte zu berücksichtigen.

professioneller Lehrerkompetenzen, weniger die Frage, wie Schule an sich funktioniert. Oser hält fest (2001, S. 79):

„Ein solches Ausbildungsmodell fragt nicht so sehr nach dem Wohlergehen der jungen Professionskandidaten, sondern eher nach dem Wohlergehen der zukünftigen Kinder, die bei diesen Kandidaten zur Schule gehen werden. Nicht entspanntes Lernen und Spaß an einem in sich stimmigen Ausbildungsklima sind vorrangig. Die Zeit nach der Ausbildung ist im Visier".

Das professionsbezogene Konzept der Wirksamkeit orientiert auf die Ausbildung professioneller Kompetenzen durch das Lehrerbildungssystem. Nicht die Schule an sich, sondern das Wohlergehen zukünftiger Kinder ist zentral. Man kann angesichts der vagen Formulierungen nur vermuten, dass Oser die Lehrerbildung idealisierend so anzulegen intendiert, dass sie anspruchsvolle professionsbezogene Ziele anvisiert und sich nicht von den Gebrochenheiten der Praxis her definiert. Den Kern professionellen Arbeitens machen für den Autor Tätigkeiten des Unterrichtens, Beratens, Unterstützens, Gegenwirkens und Behütens aus. Sie sind im Rahmen der Lehrerbildung theoretisch zu reflektieren, in Übungsphasen zu verbessern und durch Praxiserfahrungen zu sichern. In der Ausführung des Professionsgenerierungs-Modells der Wirksamkeit der Lehrerbildung antizipiert Oser eine Vorstellung vom Lehrenden, die Parallelen zu klassischen Professionsverständnissen etwa über den Arzt oder den Psychotherapeuten aufweist. Für Evaluationen ergibt sich damit die Herausforderung zu überprüfen, ob die Ausbildung den Absolventen zu einem entsprechend anspruchsvollen Kompetenzprofil verhilft. Die in schulischen Praxissituationen vorzufindenden Probleme, Brüche und Konflikte stellen in dieser Perspektive keine vermeidenswerten Belastungen dar, im Gegenteil, sie geben nach Oser konstitutive Elemente des Lehrerberufes sowie Reflexionsanlässe in der Ausbildung ab. Das von ihm entwickelte Standardkonzept (vgl. Kap. 5.2) basiert als Evaluationsinstrument auf dem hier vorgestellten Professionsgenerierungs-Modell. Die damit hergestellte Koppelung der Kompetenzevaluation an die Definition von Kompetenzen bewertet Oser als methodologischen Vorteil. Problematisch schätzt er jedoch die Möglichkeit der Formulierung von übergreifenden Kompetenzen ein.

7.2.2 Ergebnisse zur Wirksamkeit des Lehrerbildungssystems in der Schweiz

In seiner Evaluation des Lehrerbildungssystems verwendet Oser drei auf die Standards bezogene Skalen:

- Verarbeitungstiefe
- Bedeutung
- Anwendungswahrscheinlichkeit

Die Verarbeitungstiefe wird als zunehmende und systematische Verknüpfung von Theorie, Übung und Praxis im Sinne eines Portfolioansatzes modelliert und in folgenden Items realisiert (Oser 2001, S. 251):

1 = habe nicht von diesem Standard gehört

2 = habe theoretisch davon gehört
3 = habe dazu Übungen gemacht oder mich in der Praxis damit auseinandergesetzt
4 = habe Theorie und Übung, oder Theorie und Praxis, oder Übung und Praxis miteinander verbunden
5 = habe Theorie, Übung und Praxis systematisch miteinander verbunden

Die Skala „Bedeutung“ enthält die Items:

1 = nicht bedeutungsvoll
2 = ziemlich bedeutungsvoll
3 = sehr bedeutungsvoll

Die Skala „Anwendungswahrscheinlichkeit“ (in meiner Lehrertätigkeit werde ich dies…) enthält die Items:

1 = nicht beachten
2 = ab und zu beachten
3 = häufig beachten

Oser entscheidet sich explizit gegen die Verwendung einer Lickert-Skala, mit Hilfe derer Studierende befragt werden könnten, in welchem Umfang sie einen Standard gelernt oder erreicht haben, weil er davon ausgeht, dass aus Gründen der Verunsicherung die Neigung bestehen würde, eher mittlere Werte anzugeben.[63] Die Bedeutung, die Befragte den Standards zumessen, dient als Indikator dafür, inwieweit ein Bewusstsein für den Wunsch einer Standarderreichung ausgeprägt ist. Besonders wenn eine Differenz zwischen mangelnder Verarbeitungstiefe und hoch eingeschätzter Bedeutung aufgezeigt werden kann, verspricht dies Aufschluss über die Qualität der Lehrerbildung. Schließlich wird noch danach gefragt, inwieweit Studierende die Standards in ihrer späteren Lehrertätigkeit anwenden würden, um die Praktikabilität der Standards zu untersuchen. Die Verarbeitungstiefe erscheint in der Befragung als zentrale und von den anderen Skalen „Bedeutung“ und „Anwendungswahrscheinlichkeit“ flankierte Skala.
In der Anlage ist die Untersuchung als Längsschnittsstudie konzipiert und berücksichtigt die verschiedenen Zielstufen der Lehrerbildung von der Vorschule bis zur Sekundarstufe II (Gymnasium). Lehrpersonen werden am Ende ihrer Ausbildung und nach ein oder zwei Jahren Berufserfahrung befragt. Im Ergebnisteil seiner Studie stellt Oser der Lehrerausbildung ein schlechtes Zeugnis aus. Die beurteilte Verarbeitungstiefe ergibt insgesamt niedrige Werte, wenn sich auch unterschiedliche Ausprägungen festmachen lassen. Ihre Bedeutung wird wiederum überwiegend hoch oder sehr hoch und auch die Anwendungswahrscheinlichkeit hoch bis sehr hoch eingeschätzt. Tendenziell (allerdings gilt dies nicht für alle Standards) liegt die Verarbeitungstiefe umso höher, je niedriger die Zielstufe festgelegt ist. Über alle Zielstufen hinweg zeigt sich ein positiver Zusammenhang zwischen der Verarbeitungstiefe und der Anwendungsprognostik.

[63] Dass dem so nicht sein muss, legen etwa Baer et al. (2007) dar.

Die Mittelwerte[64] für die Verarbeitungstiefe liegen zwischen 2 und 3. Lehrpersonen haben sich also vor allem nur theoretisch oder nur praktisch mit den Standards beschäftigt. Der höchste Wert (AM=3,01) wird beim Standard „Phasen des Unterrichts gestalten“ und der niedrigste Wert (AM=1,32) beim Standard „Persönliche Fortbildung“ erreicht. Generell weisen die befragten Personen die höchsten Mittelwerte im didaktischen und fachdidaktischen Bereich auf. Die Verarbeitungstiefe nimmt dann aber über die pädagogisch-psychologischen sowie auf die Schulorganisation bezogenen Bereiche kontinuierlich ab und liegt bei jenen Standards am niedrigsten, die eine Auseinandersetzung mit dem Selbstkonzept der Lehrpersonen enthalten. Diese Reihenfolge überrascht nicht und spiegelt in gewisser Weise die Ausbildungscurricula.

Oser schließt aus den Ergebnissen einer geringen Verarbeitungstiefe, jedoch ihrer gleichzeitig hohen Bedeutung und Anwendungsrelevanz auf eine dysfunktionale Ausbildungssituation, weil das Portfolio kaum für einen Standard erreicht wird. Die Ausbildung würde damit „nicht professionelles Können und Beherrschen, sondern bloß partikuläres, verinseltes Wissen“ vermitteln (Oser 2001, S. 310).
Die zweite Befragung des Längsschnitts ergibt ein etwas erfreulicheres Bild. Fast alle Standards weisen eine signifikante Verbesserung der Verarbeitungstiefe auf (ebd., S. 315), wobei diese Feststellung mit nur wenigen Beispielen belegt wird. Von keiner signifikanten Veränderung zeugt etwa der Standard „Vielfältiges Handeln ermöglichen“ und es kann hier kein nennenswerter Einfluss im Längsschnitt festgestellt werden. Erneut ist anzumerken, dass sich eine gewisse Selbstreferentialität in den Ergebnissen widerspiegelt. Das von Oser verwendete Konstrukt „Verarbeitungstiefe“ ist operational so definiert, dass sich verstärkte Praxiserfahrungen fast notwendigerweise in einem Anstieg der Skalenwerte niederschlagen. In diesem Zusammenhang findet auch Mayr (2006b) keinen Beleg dafür, dass die von Oser postulierten Verarbeitungsstufen in einem hierarchischen Verhältnis zueinander stehen; in seiner eigenen Untersuchung legt Mayr den Fokus auf Praxiserfahrungen als Determinanten für die individuelle Zuweisung von Kompetenz. Zwar stellt Oser schlüssig fest, dass sich an dem Anstieg eine Verstärkung der Ausbildungsintensität fest machen lässt. Unklar bleibt aber, ob damit zugleich Kompetenzentwicklungen verbunden sind, mögen auch Plausibilitätsargumente dafür sprechen.
Die von Oser durchgeführte Standarduntersuchung ist international breit rezipiert worden. Weniger Beachtung haben einige Teilstudien erfahren, die ebenfalls im Rahmen des Schweizerischen Nationalen Forschungsprogramms 33 „Wirksamkeit unserer Bildungssysteme“ durchgeführt wurden. So legt Wild-Näf eine Untersuchung vor, die u.a. die Berufsmotivation von Studierenden des Lehramtes darstellt. Er

[64] Streng methodisch betrachtet dürften bei Nominalskalen keine Mittelwerte gebildet werden bzw. sind diese nicht sinnvoll zu interpretieren. Es ist zudem strittig, inwieweit eine höhere Verarbeitungstiefe darin Ausdruck findet, dass Lehrpersonen sich übungs- und praxisbezogen mit Standards auseinander setzen im Unterschied zu einer rein theoretischen Beschäftigung (dieser Problematik widmet sich Oser nur äußerst knapp). Folgt man der Diskussionslinie, so lassen sich Unterschiede zwischen den Mittelwerten „2“ und „3“ überhaupt nicht im Sinne von „besser“ oder „schlechter“ interpretieren. Zudem werden mit einer Mittelwertorientierung Ausbildungsinstitutionen, die konzeptionell näher an der Praxis ausbilden, im selbstreferentiellen Sinne automatisch positiver eingeschätzt. Vor diesem Hintergrund erklärt sich die abnehmende Verarbeitungstiefe mit steigender Zielstufe bei gleichzeitig abnehmendem Erklärungswert.

kommt zu dem Ergebnis, dass 20% der Studierenden sich nicht sicher sind, ob sie den Beruf später ausüben wollen, während 78% der Befragten eine positive Berufsmotivation aufweisen, die über die Ausbildungsdauer auch stabil bleibt (Wild-Näf 2001, S. 163). Diese Stabilität über die Ausbildungsspanne kann damit erklärt werden, dass es sich anscheinend um eine intrinsisch gelagerte und damit um eine inhaltliche bzw. arbeitsbezogene Berufsmotivation handelt.

Die Ergebnisse der Studierendenbefragungen zu organisatorischen Ausbildungselementen verweisen auf interessante, aber konfliktträchtige Konstellationen. So kommen 70,9% der Studierenden zu der Einschätzung, dass die Ziele der Ausbildung nicht immer klar sind, 83,3% meinen, innerhalb der Ausbildung würde Zeit vertrödelt. Wiederum finden 88,2% der Studierenden, dass sie in der Ausbildung vielen interessanten Personen begegnen (Wild-Näf 2001, S. 168f.). Überpointiert kann man feststellen, dass ein nicht unerheblicher Teil der Studierenden sich im Studium unterfordert fühlt, diese leistungsfaktoriell defizitäre Situation aber zugunsten von Klimafaktoren nur gering gewichtet. Dabei schätzen die Studierenden das Engagement der Ausbilder höher als ihr eigenes Engagement ein (ebd., S. 181). Die in die Ausbildung integrierten Praktika werden als sehr relevant für die Verbesserung der eigenen Unterrichtskompetenzen erlebt, wobei dort anfallende Belastungen als erheblich gelten. Fast 90% der Studierenden unterliegen dem Eindruck, dass die berufspraktische Ausbildung sie mehr als die berufswissenschaftliche fordert. Zusammenfassend ist festzuhalten, dass Studierende des Lehrerberufes sich im berufswissenschaftlichen Anteil tendenziell unterfordert fühlen, diese Unterforderung allerdings als positives Klima verbuchen und Leistungsanforderungen diesbezüglich eher negativ attribuieren. Insofern wird nachvollziehbar, warum sie die Praktika zwar als wichtig, aber auch als belastend wahrnehmen.
Eine von Kersten (2001) im Rahmen der Längsschnittuntersuchung nochmals vorgenommene Befragung der Lehrpersonen, mit einer allerdings deutlich kleineren Personenanzahl (ca. 25% der zum ersten Messzeitpunkt im Studium befragten Lehrpersonen), bestätigt die Ergebnisse z.T., macht aber auch deutlich, dass die Lehrpersonen ihre Einschätzungen revidieren, nachdem sie einige Praxiserfahrungen gesammelt haben. So verliert die Bedeutung von Praktika, die an früher Stelle des Studiums als sehr wichtig für die Entwicklung von Unterrichtskompetenzen beurteilt werden, vor dem Hintergrund eigenverantwortlichen Unterrichtens an Stellenwert.[65] In einer weiteren Teilstudie hat Criblez (2001) die Ausbilder der Lehrerbildung u.a. dazu befragt, wie sie selbst die Qualität der Ausbildung wahrnehmen. 79,1% schätzen die Qualität als gut oder sogar sehr gut ein. 81% der Ausbilder befinden, dass der existierende Fächerkanon ihren Vorstellungen einer guten Lehrerausbildung entspricht (ebd., S. 465). Das fachliche Ausbildungsniveau schätzen 65% als eher hoch und 13% als hoch ein (ebd., S. 473). In einer Mittelwertsbetrachtung stimmen die Ausbilder eher dem zu, dass sie in ihren Lehrveranstaltungen die im Lehr-/Studienplan vorgegebenen Ziele erreichen und

[65] Dieses Ergebnis wurde von Hascher (2006) bestätigt. Die Begegnung mit der Berufspraxis nimmt offenbar im Rahmen des Studiums eine überbewertete Rolle ein. Mit Blick auf eigenverantwortliches Praxishandeln relativiert sich die Einschätzung dieses Ausbildungselementes.

stimmen der Aussage eher nicht zu, dass es in ihrem Fach an verbindlichen Standards mangelt (ebd., S. 472).
Zusammenfassend kann man sagen, dass die von Oser & Oelkers bzw. ihren Mitarbeitern vorgelegten Studien einen umfassenden empirischen Überblick über die Wirksamkeit des schweizerischen Lehrerbildungssystems bieten und in weiten Teilen eine Unwirksamkeit bzw. eine geringe Ausschöpfung möglicher Potentiale darlegen. Die von Oser herausgearbeitete selbst attribuierte mangelhafte Standarderreichung durch Lehrpersonen im Sinne der fehlenden Entwicklung professioneller Unterrichtskompetenzen hat in der fachlichen und bildungspolitischen Diskussion am meisten Furore gemacht und führte in der Schweiz bereits zu einer weitgehenden Unsteuerung in Richtung einer standardbasierten Ausbildung. Bislang fehlt aber noch der empirische Nachweis, dass die Umsteuerung eine bessere Ausbildungsqualität bewirkt. Es bleibt eine grundsätzliche Herausforderung der Lehrerbildung, Lernprozesse unter dem Verzicht auf rein technologische Konzeptvermittlung zu initiieren, die in professionelle Kompetenzen münden. Die Einführung von Standards kann diesbezüglich hilfreich sein, gleichzeitig darf eine Standardimplementierung auf institutioneller Ebene nicht mit dem Erwerb auf Studierendenseite appellativ verwechselt werden (Mayr 2002; Mayr & Paseka 2002). Nimmt man das von Oser verwendete Konstrukt der Verarbeitungstiefe – als Konglomerat von Theorie, Praxis und Übung – als prinzipiell unstrittige Referenz für die Entwicklung von professionellen Kompetenzen an, so kann damit keine automatische Qualitätssteigerung in der Lehrerbildung garantiert werden (Mayr 2006a, b). Gleichwohl ist die Ausrichtung an solchen Portfolioansätzen als Verbesserung von Rahmenbedingungen anzusehen, die sich evtl. positiv auf die letztlich unverfügbaren Selbstbildungsprozesse der Studierenden auswirken (Messner 2001). Nicht erst seit den systemtheoretischen Ansätzen von Luhmann gelten Aussagen über den Einfluss von Variablenveränderungen in dem übergreifenden System auf die Individualebene (beziehungsweise umgekehrt) jedoch als gewagt.
Methodisch weisen die von Oser durchgeführten Befragungen Schwächen auf, was die Datenqualität und die Reichweite der darauf basierenden Aussagen einschränkt. Das gilt sowohl für die Anlage der Untersuchung, die auf Selbsteinschätzungen basiert, als auch für die Konstruktvalidität der Skala „Verarbeitungstiefe“, die zudem keine weitergehende inferenzstatistische Analyse erlaubt. Mit diesen Einschränkungen erscheint auch das Ergebnis des in der Studie nur verkürzt dargestellten Längsschnitts nicht unproblematisch. So verwundert die Tatsache, dass die Studie für einschneidende Veränderung gesorgt hat und auch international auf großes Interesse stößt. Man kann vermuten, dass Osers Ergebnisse in der Schweiz auf einen Nerv getroffen sind und bereits bestehenden Stimmungen im Bildungssystem eine Verstärkung geliefert haben. Diese Deutung intendiert aber nicht die grundsätzliche Infragestellung der Ergebnisse.
Die von Brühwiler, Kersten und Criblez vorgestellten Teilstudien sind methodisch weniger angreifbar. Ergebnisse fassen das Verhältnis von Studierenden bzw. Absolventen zu ihrer Ausbildung bzw. ihren Ausbildern sowie das Bild, das Ausbilder, als Hauptakteure der Lehrerbildung (Criblez 2001), von der Ausbildung haben. Die dabei zu Tage tretenden unterschiedlichen Blickwinkel auf die Lehrerbildung dokumentieren nicht nur für sich genommen problematische Situationslagen, sondern zeugen auch davon, dass eine konstruktive Verständigung kaum gelingt.

Demnach kann die vergleichsweise positive Einschätzung der Lehrerbildner bezogen auf die Ausbildungsqualität eher als Indikator für Berufszufriedenheit denn als objektive Qualitätseinschätzung gewertet werden. Polemisch kann man vermuten, dass sich dahinter ein nur geringes Innovierungspotential verbirgt, was zusammen mit der hohen Bewertung sozialer Klimafaktoren bei gleichzeitiger Ablehnung von Leistungsanforderungen der Studierenden für eine Gemengelage sorgt, bei der man gespannt sein darf, inwieweit die Einführung von Standards diesbezüglich Veränderungen zeitigen kann.

7.3 Wirksamkeit der Lehrerbildung in den USA im Spannungsfeld von Prozessen der Deregulation und Professionalisierung

Für die vereinigten Staaten von Amerika lässt sich seit einigen Jahren eine Diskussionslinie verfolgen, die charakterisierend als Gegensatzpaar von Deregulierungsbemühungen und Professionalisierungsbestrebungen gekennzeichnet werden kann. Die den jeweiligen Richtungen zuzuordnenden Argumente sind bereits dargestellt worden (vgl. Kap. 5.1). Beide Lager nehmen für sich in Anspruch eine Verbesserung der Lehrerbildung zu intendieren. Die Implementation von Standardkonzepten steht nicht im Mittelpunkt der Auseinandersetzungen, wohl aber ihre Funktion. Während die sog. „Deregulierer“[66] Standards auf verschiedenen Ebenen als Ausgangspunkt dafür fassen, Ausbildungsanforderungen niedriger zu setzen (die Überprüfung der Standards fragt nicht nach dem Weg der Standarderreichung und entbindet damit Personen und Institutionen von ihrer Ausbildungsverantwortlichkeit), wollen die „Professionalisierer“ die Eingangsvoraussetzungen für den Lehrerberuf und damit die Ausbildung quantitativ und qualitativ verbessern, orientieren sich jedoch ebenfalls an Standards. Beide Gruppen liefern umfangreiche Metaanalysen empirischer Studien, nutzen zum Teil die gleiche Datenbasis, kommen jedoch zu unterschiedlichen Ergebnissen (Blömeke 2004; Abell Foundation 2001a, b; Darling-Hammond 2000). Darling-Hammond beispielsweise (2000), eine Hauptvertreterin des Professionalisierungsansatzes in den USA, stellt in ihrer Metastudie Merkmale von Lehrpersonen in ihrer Bedeutung für die Verbesserung von Lehr-Lernprozessen heraus und diskutiert die Bedeutung eines lehrerbildenden Studiengangs (in den USA gibt es eine Vielzahl von Studienangeboten, in der Regel handelt es sich aber um einen Bachelor mit einer Spezialisierung (Major) in einer Fachwissenschaft oder Erziehungswissenschaft): Lehrpersonen mit einem entsprechenden Abschluss erzielen im Vergleich zu Lehrpersonen mit einer Kurzausbildung oder keiner Ausbildung bessere Leistungen bei ihren Schülern und werden auch von Experten besser eingeschätzt. Zudem sind

[66] Grundsatzpositionen der Deregulierer folgen einer neoliberalen Logik, nach der der Markt mit seinen Prinzipien des Angebots und der Nachfrage am besten dafür sorgen kann, dass die befähigsten Lehrer Anstellungen finden und es dementsprechend zu einer Aussortierung der schlechteren Lehrer kommt. Zu wenig bedenkt diese Position allerdings mit, dass die Qualität der dem Markt zur Verfügung stehenden Bewerber wesentlich von der Ausbildungsqualität beeinflusst wird. Nimmt man in der Ausbildung ökonomisch begründete Einsparungen vor, die sich auf die Qualität der Ausbildung und damit auf das Absolventenprofil niederschlagen, kann der Markt bestenfalls noch die „Einäugigen unter den Blinden“ selektieren. Der Deregulierungsansatz wird von der einflussreichen und regierungsnahen Abell Foundation unterstützt.

sie zufriedener mit ihrer Berufsausübung, haben weniger Probleme zu Beginn ihres Berufslebens und verbleiben vor allem länger in ihrem Beruf (Mintrop 2006). Dabei macht es keinen Unterschied, ob der Major in einer Fachwissenschaft oder der Erziehungswissenschaft absolviert wurde (Murnane 1985). Bezüglich der Verbesserung von Schülerleistungen berichtet Darling-Hammond (2000, S. 8) auch von inkonsistenten Ergebnissen, die sie aber auf methodische Schwächen zurückführt. Höhere Studienabschlüsse (Vergleich: Master vs. Bachelor) bedingen wiederum erhöhte Schülerleistungen.

Betrachtet man nicht die Lehrerbildung im Ganzen als unabhängige Variable, sondern untersucht ihre Komponenten, z.B. die Qualität der Ausbildungsstätte oder Testergebnisse bei Berufseingangstests, so werden nicht durchgehend linear positive Korrelationen zwischen untersuchten Merkmalen und der Verbesserung von Schülerleistungen festgestellt (Wayne & Youngs 2006). Ähnlich sieht es aus, wenn man die Qualität der fachwissenschaftlichen Ausbildung als treatment für die Verbesserung von Lehrerhandeln oder Schülerleistungen untersucht. Auch hier gelangt man zu inkonsistenten Ergebnissen. Ein Erklärungsmuster liegt in der unterschiedlichen Auswirkung der Qualität der Fachausbildung auf verschiedene Schul- und Klassenstufen. Generell kann man die These aufstellen, dass Expertise in den Fachwissenschaften notwendig ist, um Lernprozesse bei Schülern zu unterstützen, dass sich aber ab einem gewissen Grad der Fachexpertise Deckeneffekte einstellen, die keine signifikante Verbesserung mehr mit sich bringen.[67] Ein zu hoher Grad an fachwissenschaftlicher Expertise kann zudem negativ mit einer für das Schülerlernen notwendigen didaktischen Reduktion in Form von Elementarisierung korrespondieren. Für das fachdidaktische Wissen („pedagogical content knowledge“) berichtet dagegen Darling-Hammond (2000) von konsistent positiven Zusammenhängen. Allerdings muss einschränkend gesagt werden, dass die Fachdidaktik in den USA eine andere Tradition als in Deutschland hat und eher als ein Methodentraining verstanden werden kann. Für die erziehungswissenschaftliche Ausbildung ergeben sich überwiegend (nicht ausschließlich) positive Zusammenhänge. Dabei hebt Darling-Hammond (2000, S. 5) hervor, dass die Befunde in Zusammenhang mit der fachwissenschaftlichen Ausbildung zu sehen sind. In der Konsequenz spricht sie sich sowohl für die Verstärkung der erziehungswissenschaftlichen als auch der fachwissenschaftlichen Ausbildung aus. Wiederum gilt, dass die erziehungswissenschaftlichen Ausbildungselemente – anders als in Deutschland – eher pragmatisch konzeptualisiert sind.

Insgesamt führt die Untersuchung der Hauptkomponenten der Lehrerbildung zu einer divergenten Befundlage. Eine Untersuchung berufsunspezifischer intellektueller Fähigkeiten von Lehrpersonen ergab ebenfalls widersprüchliche Zusammenhänge zu Schülerleistungen oder Lehrerhandeln, dagegen geben die verbalen Fähigkeiten

[67] Dieser Fragestellung geht zurzeit das Max-Planck-Institut für Bildungsforschung mit der Studie „COACTIV“ nach. Problematisch an der Anlage erscheint dabei, dass eine nähere Unterscheidung zwischen fachdidaktischer und fachwissenschaftlicher Kompetenz nicht vorgesehen ist (Brunner et al. 2006). Allerdings ist die Untersuchung von Kompetenzdimensionen immer von der zugrunde gelegten Konstruktmodellierung abhängig. In der Studie wird fachdidaktische Kompetenz wesentlich in Abhängigkeit von Fachwissen modelliert, weil erst auf dieser Grundlage die mehrperspektivische Darstellung und Erklärung von Phänomenen gelingen könne. Dass damit eine kompetenzbezogene Unterscheidung in der Ergebnisanalyse schwer fällt, überrascht nicht.

offensichtlich eine gute Prädiktorvariable ab (Blömeke 2004; Murnane 1985). Signifikante Auswirkungen verbaler Fähigkeiten liegen angesichts der überwiegenden Strukturierung von Unterricht über Kommunikationsprozesse nahe. Intellektuelle Fähigkeiten zeitigen dagegen nur unspezifische Effekte, weil für die unterrichtliche Gestaltung die Bereichsspezifität fehlt. Einflüsse des Intelligenzquotienten können vermutlich nur für Extremwerte sichtbar gemacht werden.

Im Gegensatz zu der bislang dargestellten Strömung, die empirisch eine Verbesserung der Lehrerausbildung in Richtung der Stärkung fachwissenschaftlicher und erziehungswissenschaftlicher Ausbildungselemente als Professionalisierungsbemühungen zu begründen sucht, besteht eine andere Strömung in den USA, deren Vertreter eine solche veränderte Ausrichtung der Lehrerbildung, die auch in Verbindung mit zeitlichen Ausdehnungen und dementsprechenden Ressourcenverwendungen zu sehen ist, im Kern ablehnen. Protagonistin dieser bereits angesprochenen Deregulierungsbestrebungen ist die Abell Foundation. In einer eigenen, groß angelegten Studie kommt sie zu dem Ergebnis, es könnten keine hinreichenden Belege dafür angeführt werden, dass auf der Grundlage bisheriger Konzeptionen eine Verbesserung der Lehrerbildung zu erzielen sei. Die Annahme lautet hingegen, es gäbe Personen, die Kraft ihrer Geburt oder auf Grund von Erfahrungen, die gerade jenseits der in der Lehrerbildung intendierten Zielsetzungen liegen, besonders geeignet zum Unterrichten in Schulen wären. Hier scheint also wieder die Lehrerpersönlichkeit durch, die unter Annahmen angeborener bzw. durch Erfahrungserwerb ausgebildeter Kompetenzen – jedoch nicht im genuinen Kontext pädagogischer professioneller Entwicklungsprozesse – modelliert wird. Diese Diskussionslinie ist auch in Deutschland – wenn auch weniger im wissenschaftlichem Diskurs – nach wie vor virulent. Die Abell Foundation zieht aus ihren Metaanalysen[68] den Schluss, dass ein Eingangstest, bestehend aus der Überprüfung verbaler Fähigkeiten, der Fähigkeit Leseunterricht zu erteilen sowie (bei Sekundarschullehrern) fachwissenschaftlicher Kenntnisse anhand eines Fachwissenstest die adäquate Grundlage dafür darstellt, lehrbezogene und also einstellungsrelevante Kompetenzen überprüfen zu können. Gleichzeitig wird dafür plädiert, weniger die grundständige Lehrerbildung zu unterstützen, sondern mehr Energie auf die Berufseingangsphase zu richten.
Die Diskussion darüber, inwieweit die Lehrerbildung durch Professionalisierungs- oder Deregulierungsprozesse verbessert werden kann, ist in den USA nicht abgeschlossen. Hoffnungen auf weiteren Aufschluss liegen in der zurzeit durchgeführten TEDS-M-Studie. Es bestehen Anzeichen dafür, dass sich in Deutschland im Zuge der Umstellung der Lehrerbildung auf Bachelor- und Masterstudiengänge ähnliche Positionskämpfe ergeben könnten. Die Studien von Mayr (2006a, b) weisen in eine Richtung, wonach man die Notwendigkeit einer institutionalisierten Lehrerbildung unter Berücksichtigung spezifischer Fachinhalte

[68] Von der Abell Foundation werden allerdings nur empirische Studien berücksichtigt, die als large scale-Studien konzipiert sind, die die Wirksamkeit der Lehrerbildung anhand von Schülerleistungen messen und in peer-reviewed-Veröffentlichungen publiziert wurden (Blömeke 2004, S. 62). Mit diesen Einschränkungen mögen bestimmte Verzerrungen verbunden sein. In jedem Fall werden damit aber qualitative Studien und Qualifikationsarbeiten ausgeschlossen.

ebenfalls in Frage stellen kann. Allerdings verwehrt Mayr sich gegen die Indienstnahme seiner Ergebnisse für die Deregulierungsargumentation. Folgerichtig wirft er die Frage auf (Mayr 2006b, S. 162):

„Was spricht gegen die Vorstellung, dass professionell handelnde Lehrer/innen in ihrer Berufausübung alle ihnen zur Verfügung stehenden kognitiven und emotionalen Ressourcen nutzen, gleichgültig wie diese entstanden sind? Und was spricht dagegen, dass diese Faktoren gleichwertig auch in die Diskussion und in die Erforschung von Professionalisierungsprozessen eingehen?“

7.4 Internationale Ergebnisse zur Kompetenzstruktur von Studierenden des Lehrerberufes

Frey (2004) legt eine internationale Studie zur Untersuchung der Kompetenzen von Lehramtsstudierenden vor. Einbezogen werden die Länder Deutschland, Österreich, Schweiz, Polen, und Italien. Die methodische Anlage sieht schriftliche Selbsteinschätzungen von Studierenden zu den Kompetenzklassen Fachkompetenz, Sozialkompetenz, Methodenkompetenz und Personalkompetenz vor, voraus sich länderspezifische Unterschiede ablesen lassen (vgl. Kap. 3.6). Als intervenierende Variablen werden die Außenbezüge Alter, Geschlecht, Semesteranzahl, Ausbildungszufriedenheit und Berufsaussicht kontrolliert, um mögliche Effekte auf die Kompetenzklassen erfassen zu können. In der methodologischen Diskussion seines Vorgehens wägt Frey Vor- und Nachteile des Instruments der Selbstbeurteilungen ab. Vorteile erscheinen ihm in forschungsökonomischer Hinsicht gegeben. Selbstbeurteilungen lassen sich in relativ kurzer Zeit und mit wenig Aufwand erheben. Zudem liegt die Auffassung nahe, dass betroffene Personen am besten selbst über sich Auskunft geben können. Als problematisch stellen sich jedoch vorausgesetzte Bedingungsannahmen für Selbstbeurteilungen dar, so die Erwartung, dass Probanden tatsächlich in der Lage sind, retrospektiv ihre Fähigkeiten realistisch einschätzen zu können. In diesem Zusammenhang geht es in der Studie zunächst um die Klärung der methodologischen Frage, ob über Selbsteinschätzungen reliable Ergebnisse zur Kompetenzdiagnostik erreicht werden können und ob die theoretisch plausible Einteilung von Kompetenzklassen dadurch Bestätigung findet. In einem zweiten Schritt interessiert, auf welchem Niveau Studierende der verschiedenen Länder ihre Kompetenzen einschätzen. Konkret wird nach der Anwendung von Kompetenzen gefragt und nach der darauf bezogenen Einschätzung von Studierenden auf einer sechsstufigen Skala mit den Antwortformaten „(1) = trifft völlig zu“ bis „(6) trifft gar nicht zu“.[69] Ein dritter Schritt untersucht abzuleitende Effekte aus den kontrollierten Außenbezügen. Die Stichprobe umfasst insgesamt 1841 Studierende, wobei es sich bei dem größten Anteil, nämlich bei 1313, um deutsche

[69] Strittig bleibt, ob über die Einschätzung der Anwendung nicht doch vorrangig eine Quantität und weniger eine Qualität von Handlungsoptionen der Studierenden ausgewiesen wird. Mit einiger Plausibilität kann man zwar einen Zusammenhang von Anwendungswahrscheinlichkeit und Kompetenzniveau konstatieren, der aber keinesfalls eindeutig besteht. Mit dieser Einschränkung relativieren sich auch die insgesamt recht guten Selbsteinschätzungswerte und es ist ein Erklärungsmuster dafür gegeben, warum die Studierenden höhere Anwendungshäufigkeiten für die Dimensionen angeben, die eine größere Nähe zu den didaktischen Arrangements der Hochschulen oder Akademien aufweisen.

Studierende handelt. Insgesamt, so zeigt der Ergebnisteil auf, sind die Studierenden mit der Ausbildung eher zufrieden und beurteilen ihre Berufsaussichten positiv.
Was die Reliabilität von Selbsteinschätzungen im Rahmen der Kompetenzdiagnostik betrifft, bestätigt eine Faktorenanalyse das theoretische Konzept der vier Kompetenzklassen mit z.T. sehr hohen Faktorenladungen. Die Selbsteinschätzungen in den einzelnen Fähigkeitsdimensionen zeugen übergreifend von recht guten Werten, wobei Mittelwertsanalysen Unterschiede zwischen einzelnen Fähigkeitsdimensionen sowie zwischen den Ländergruppen aufzeigen. Im Bereich der Fachkompetenz schätzen sich Studierende in den unterrichtsnahen Dimensionen (z.B. Lehrer-Schüler-Beziehungen) besser ein als in den auf die schulische Organisation bezogenen Dimensionen (z.B. Zusammenarbeit in der Schule). Spitzenreiter sind die österreichischen Studierenden, während die polnischen Studierenden sich selbst die schlechtesten Werte zuschreiben. Dazwischen liegen Werte der italienischen, schweizerischen und deutschen Studierenden. Vergleichbare Ergebnisse liefert Frey für die übrigen drei Kompetenzklassen; auch hier bilden sich analoge Länderabstufungen ab.
Eine Korrelationsanalyse von Selbsteinschätzungen und kontrollierten Außenbezügen (Alter, Geschlecht, Semesterzahl, Ausbildungszufriedenheit, Berufsaussichten) ergibt insgesamt kleine Werte, die aber länderspezifisch differieren. Für die schweizerischen und österreichischen Studierenden konnten keine Signifikanzen gefunden werden. Bei den deutschen, italienischen und polnischen Studierenden lassen sich nur in Einzelfällen signifikante Effekte aufzeigen, wobei Varianzaufklärungen gering ausfallen.
Obwohl die Selbsteinschätzungen der Studierenden über alle fünf Länder vergleichsweise positiv gewichtet sind, schließt Frey aus den z.T. signifikanten Länderdifferenzen einen nicht unerheblichen Lern- oder Kompetenzentwicklungsbedarf. Eine detaillierte Untersuchung der Mittelwerte zeigt auf, dass sich die Studierenden in den Fähigkeitsdimensionen Lehrer-Schüler-Beziehungen, Medien des Unterrichts, allgemeine didaktische Gesichtspunkte, soziale Verantwortung, situationsgerechtes Auftreten, Reflexivität, zielorientiertes Handeln, Arbeitstechniken, Hilfsbereitschaft und Einfühlsamkeit sowie Gelassenheit und Geduldigkeit besser einschätzen als in den Dimensionen Zusammenarbeit in der Schule, Schule und Öffentlichkeit, Selbstorganisation, Konfliktfähigkeit, Führungsfähigkeit und Analysefähigkeit. Frey folgert daraus einen nur geringen Stellenwert von Kompetenzdimensionen in den Ausbildungen an Hochschulen und Akademien, die zu einem anhaltend prozessualen und intensivierten Handeln in Praxissituationen mit Kindern, Jugendlichen und schulischen Kollegen befähigen. Welche Schlüsse aus diesen Ergebnissen für die Lehrerbildung gezogen werden können, führt Frey allerdings nicht aus.

7.5 Empirische Studien zur Lehrerbildung in Deutschland

Die empirische Untersuchung der Qualität der Lehrerbildung hat international erst seit den 1990er Jahren an Aufmerksamkeit und Intensität gewonnen. Larcher & Oelkers (2004, S. 129) stellen polemisch fest, „zuvor sind empirische Studien nur sporadisch unternommen worden, und die Theoriediskussion lebte von ideologischen

Gegensätzen, die gebetsmühlenartig in Stellung gebracht wurden, vor allem um Interessen zu schützen und Entwicklung zu blockieren“. Weiterhin stellen die Autoren fest: „Wenn es eine Krise in der Lehrerbildung gibt, dann ist es wesentlich eine Krise der fehlenden Daten“ (ebd.). Gründe für die prekäre Datenlage liegen nicht zuletzt an den diskutierten Schwierigkeiten, das Unternehmen Lehrerbildung hinsichtlich von Wirksamkeitsdimensionen empirisch zu fassen. Zumindest für Deutschland kann man unterstellen, dass viele Akteure der Lehrerbildung ohnehin der Meinung waren, dass der betriebene hohe (akademische) Aufwand und die Zweiphasigkeit der Lehrerbildung eine auch im internationalen Vergleich hohe Qualität garantieren würden. Dieser Glaube ist durch die internationalen Vergleichsstudien (TIMMS und vor allem PISA) ins Wanken geraten.
Mit der forcierten Etablierung des Kompetenzbegriffes als wissenschaftliche Kategorie[70] vermehren sich empirische Untersuchungen zum Lehrerhandeln. Es können beachtenswerte Ergebnisse aufgezeigt werden, auch wenn Konstruktunterschiede zur Modellierung des Kompetenzbegriffes bestehen. In Deutschland nehmen die Studien VERBAL[71] und das Nachfolgeprojekt REBHOLZ[72] als empirische Evaluationsprogramme wegweisenden Stellenwert ein. Zentrale Berücksichtigung erfahren hier die Schulpraktischen Studien sowie ein Rückmeldedesign zwecks Rückspiegelung der Ergebnisse an Studierende (Bodensohn, Frey & Balzer 2004; Bodensohn & Schneider 2006; Bodensohn, Schneider & Jäger 2007). In den Studien wurde eine von Frey & Balzer (2003) entwickelte Kompetenzklassifikation eingesetzt, mit dem Ziel Aufschluss über tatsächliche personale, fachliche, methodische und soziale Kompetenzen der Studierenden sowie deren möglichen Einfluss auf die Schulpraktischen Studien im Längsschnitt zu erlangen. Die Stichprobe umfasste 401 Studierende, die sich in 16 Kompetenzdimensionen vor und nach zwei Blockpraktika einschätzten, während die betreuenden Lehrer sie denselben Dimensionen zuordneten. Wiederum wurde ein Antwortformat gewählt (vgl. Kap. 7.4), das sich auf die Häufigkeit des jeweils zu beobachtenden Verhaltens bezieht. Als Fazit der Studie können folgende Ergebnispunkte herausgestellt werden (Bodensohn & Schneider 2006, S. 264):

- Die Selbstbeurteilung der Studierenden fällt über fast alle Dimensionen schlechter aus als die Fremdbewertung durch betreuende Lehrer, wobei beim ersten Blockpraktikum eine größere Diskrepanz als beim zweiten Blockpraktikum zu beobachten ist.
- Die Einschätzungen am Ende der Praktika liegen durchweg höher als zu Beginn der Praktika.
- Das Verhalten im zweiten Blockpraktikum wird sowohl in der Selbst- als auch in der Fremdbeurteilung konsistent besser als im ersten Blockpraktikum eingeschätzt.

[70] Ein aktueller Beleg dafür ist die Ausschreibung eines DFG-Schwerpunktprogramms zum Thema „Kompetenzentwicklung“.
[71] „Verbesserung der Beratungsqualität bei der Ausbildung beruflicher Handlungskompetenz von Lehramtsstudierenden“; Laufzeit 2002-2004.
[72] „Report Entwicklung Beruflicher Handlungskompetenzen im Organisationsbereich der Lehrerbildung“; Laufzeit 2005-2010.

- Die Diskrepanz zwischen Selbst- und Fremdwahrnehmung ist im ersten Blockpraktikum, mit der Tendenz höherer Werte in der Fremdbeurteilung, größer als im zweiten Blockpraktikum.
- Im Längsschnitt zeugen die Selbsteinschätzungen im Vergleich zu den Fremdeinschätzungen von einem stärkeren Kompetenzzuwachs über die Zeitspanne der beiden Praktika.

Diese interessanten, wenn auch methodisch bezüglich des Antwortformates und des Erhebungsdesigns nicht unproblematischen Ergebnisse können zumindest als vorsichtiger Beleg dafür gewertet werden, dass in den Blockpraktika eine Kompetenzentwicklung statt findet. Was übergreifend die höheren Einschätzungen zum Ende der Praktika im Vergleich zum jeweiligen Praktikumsbeginn betrifft, relativiert sich ein hier dokumentierter großer Effekt durch die Tatsache, dass Beurteilungen zu Beginn und am Ende der Praktika auf demselben Beurteilungsbogen festgehalten wurden und keine Kontrollen darüber stattfanden, zu welchen Zeitpunkten Eintragungen jeweils vorgenommen wurden. Im Falle gleichzeitiger Einschätzungen für beide Messpunkte erst nach dem zweiten Praktikum können sich Effekte auch mit dem Faktor sozialer Erwünschtheit erklären. Hinsichtlich des stärker eingeschätzten kontinuierlichen Kompetenzzuwachses in Selbsteinschätzungen im Vergleich zu Fremdeinschätzungen ist wiederum anzumerken, dass Studierende in den seltensten Fällen in beiden Praktika von denselben Lehrpersonen eingeschätzt wurden.
Dass Mentoren die Studierenden insbesondere im ersten Blockpraktikum besser einschätzen als diese sich selbst kann zum einen damit erklärt werden, dass den Studierenden ein Referenzwert für die Einschätzung ihrer Kompetenzen zunächst fehlt und sie damit ihre Kompetenzwahrnehmung auf das für sie noch weitgehend unbekannte Handlungsfeld Schule beziehen. Im zweiten Blockpraktikum fällt die Diskrepanz zwischen Selbst- und Fremdeinschätzung deutlich geringer aus, woraus man folgern kann, dass die Studierenden mittlerweile über eine bessere Orientierung zur Einschätzung ihrer Kompetenzen verfügen. Höhere Einschätzungen durch Mentoren können aber auch dadurch erklärt werden, dass diese immer auch ihre eigene Arbeit mit bewerten und insofern zu optimistischeren Beurteilungen neigen. Angegebene Entwicklungen innerhalb eines Praktikums stellen deshalb einen zu relativierenden Beleg für eine beobachtbare Kompetenzentwicklung dar. Dass allerdings eine Entwicklung über beide Praktika stattfindet (Referenz jeweils die Kompetenzeinschätzungen am Ende der zwei Praktika über beide Erhebungsverfahren), darf als Indiz für die über eine Sozialisation in Praxisbezügen hinausreichende wichtige Funktion gelten, die den Schulpraktischen Studien innerhalb der Lehrerbildung zukommt.

7.5.1 Untersuchungen zur Entwicklung von diagnostischer Kompetenz in der zweiten Phase der Lehrerbildung

Abs (2006) bezieht sich in seiner Untersuchung über die zweite Phase der Lehrerbildung auf die Ergebnisse von PISA 2000, nach denen Lehrer die Fähigkeiten ihrer Schüler oft falsch einschätzen. Förderbedarfe von Schülern werden aus diesem

Grund nur unangemessen erkannt, was wiederum eine defizitäre beratende Unterstützung zur Folge hat (Baumert et al. 2001, S. 99ff.). Die PISA-Ergebnisse zur Kompetenz- und Potentialunterschätzung waren ein gewichtiger Auslöser für die Einführung von Bildungsstandards auf der Basis von Kompetenzmodellen (zunächst auf der Ebene der Kultusministerkonferenz, dann in der Umsetzung einzelner Bundesländer). Daraus ergibt sich die Notwendigkeit neue Formen der Diagnostik in der ersten und zweiten Phase auszubilden, zumal die von der KMK (2004) vorgelegten Standards der Lehrerbildung explizit auf einem diagnostischen Verständnis fußen, welches im Kern einem Fördergedanken und nicht habitualisierter Selektionskriterien verpflichtet ist.

In diesem Zusammenhang fokussiert die Studie von Abs die zweite Phase der Lehrerbildung, weil diese nicht mehr polyvalent angelegt ist, sondern als Vorbereitungsdienst unmittelbar in die Berufstätigkeit führen soll und damit zentrale Bedeutung für die Berufsausübung sowie den weiteren Kompetenzerwerb in der Berufsausübung und der dritten Phase der Lehrerbildung einnimmt (Abs 2006, S. 219). Theoretisch verortet der Autor seine Studie in einem hypothetischen Wirkungsmodell zur Lehrerbildung in der zweiten Phase, das sich stark an das Angebot-Nutzungsmodell von Fend (2000) anlehnt. Zudem verschränkt Abs es mit dem Bildungsproduktionsmodell von Scheerens & Bosker (1997) und richtet es auf die Einbeziehung politischer, organisatorischer und personeller Kontextfaktoren aus.

Seine Studie ist in ein übergreifendes Evaluationsprogramm eingebettet, durch welches alle Studienseminare Hessens mindestens einmal evaluiert werden. Die hier interessierenden Ergebnisse beziehen sich auf die erste Erhebungswelle 2004, in der fünf Studienseminare für das Lehramt an Grund-/Haupt-/Real- und Förderschulen sowie fünf Studienseminare für das Lehramt an Gymnasien aus verschiedenen Regionen des Landes unter jeweiliger Hinzuziehung aller Ausbildner untersucht wurden.[73] Hingegen wurden kooperierende Ausbildungsschulen bzw. deren Mentoren nur in kleineren Stichproben einbezogen. Die Untersuchung der Referendare war als Vollerhebung konzipiert, wobei 82% erreicht wurden, was einer Stichprobengröße von N=1102 Personen entspricht. Die Studierenden wurden am Ende der dreimonatigen Eingangsphase sowie am Ende des Referendariats (in einem Quasi-Längsschnitt) schriftlich befragt und sollten die Arbeit der Ausbildner und Mentoren mit Bezug auf Items zu folgenden Skalen einschätzen (Abs 2006, S. 221):

- Beurteilung der Arbeit von Ausbildnern/Ausbildnerinnen (Beispielitem: „Die Beratungen mit dem Ausbildner/der Ausbildnerin nach Unterrichtsbesuchen sind eine wesentliche Unterstützung für mich“)
- Beurteilung der Arbeit von Mentoren/Mentorinnen (Beispielitem: „Ich kann von meinen Mentoren/Mentorinnen lernen, wie ich guten Unterricht halten kann“)
- Geteiltes Grundverständnis der Lehrerbildner/Lehrerbildnerinnen (Beispielitem: „Die Bewertung der Referendare/Referendarinnen folgt einheitlichen Kriterien“)

[73] Die Stichprobe setzte sich zu 50% aus Referendaren für das Lehramt an Gymnasien zusammen, die andere Hälfte verteilte sich wie folgt auf die anderen Schularten: Grundschule: 27%, Haupt und Realschule: 11% und Förderschule: 12%. Die Referendare befanden sich mehrheitlich in einer Altersspanne von 26-30 Jahren und waren zu 71% weiblich.

Zusätzlich wurden in Anlehnung an Oser (2001, S. 57) Standards formuliert, von denen 23 fachspezifisch über alle Fächer und 34 im Sinne einer übergreifenden Einschätzung der gesamten Ausbildung von den Referendaren erfragt wurden. Das Abfragen eines Standards wurde jeweils mit dem Satz eingeleitet: „In meiner Referendariatsausbildung wurde mir bislang vermittelt…“. Das darauf bezogene Antwortformat rekurriert ebenfalls auf die Studie von Oser (2001) und umfasst fünf Dimensionen standardbezogener Lerngelegenheiten (Theorievermittlung im Seminar, Übung im Seminar, Erfahrung in der Schule, Reflexion mit Ausbildner/in, Reflexion mit Mentor/in). Auch wenn man annehmen kann, dass sich hinter dieser Aufzählung eine Hierarchisierung der Lerngelegenheiten verbirgt, wurde – im Gegensatz zu der Studie von Oser – eine getrennte Ankreuzmöglichkeit präsentiert, um eine Einschränkung von Antwortmöglichkeiten und die damit verbundene Abbildung vermeintlicher Verarbeitungstiefen zu vermeiden. Das zentrale Erkenntnisinteresse zielte schließlich auch auf die Frage, inwieweit sich verschiedene Kontextfaktoren (das Geschlecht, das Fach, der Seminarort, die Seminarart) als unabhängige Variablen auf die selbst eingeschätzten Kompetenzdimensionen auswirken. Im Weiteren liegt die Aufmerksamkeit mit Blick auf die eigene Fragestellung nur auf jenen Standards, die sich fachspezifisch auf die Dimensionen Diagnostik und Unterrichtsplanung beziehen.

Als Ergebnis einer Varianzkomponentenanalyse stellt Abs unterschiedliche Aufklärungen der Varianzen durch die kontrollierten Kontextfaktoren für die Kompetenzdimensionen Diagnostik und Unterrichtsplanung heraus. Während das Geschlecht überhaupt nicht und das jeweilige Fach nur sehr wenig zur Aufklärung beitragen, wird die größte Varianzaufklärung durch die Zuordnung der Kompetenzeinschätzungen zu einer Person geleistet. Dieser Sachverhalt gilt für beide Dimensionen, im größeren Maße aber für die Unterrichtsplanung (38% bzw. 26% geklärte Varianz). Der Seminarort (dichotome Unterscheidung zwischen Grund-, Haupt-, Real- und Förderschule vs. Gymnasium) wirkt nicht auf die Kompetenzeinschätzungen zur Unterrichtsplanung ein, aber durchaus auf die diagnostische Dimension (7% geklärte Varianz). Ein Vergleich der Mittelwerte von Selbsteinschätzungen bestätigt die Ergebnisse der Varianzkomponentenanalyse. Die Einschätzungen zur Unterrichtsplanung unterscheiden sich bezogen auf die zwei Seminararten nicht. Im Bereich der diagnostischen Kompetenz geben Referendare an Studienseminaren für das Lehramt an Grund-, Haupt-, Real- und Förderschulen signifikant höhere Werte als Referendare aus Studienseminaren für das Gymnasium an. Regressionsanalysen, die Daten zur Einschätzung der Lerngelegenheiten bzw. zur Beurteilung der Arbeitsqualität von Ausbildnern und Mentoren auf selbst eingeschätzte Kompetenzdimensionen beziehen, führen für die beiden Seminararten zu teilweise stark voneinander abweichenden Ergebnissen, während das Bedingungsgefüge innerhalb der Seminararten für die Dimensionen Diagnostik und Unterrichtsplanung tendenziell ähnlich ausfällt. Betrachtet man die Aussagen der Referendare aus Grund-, Haupt-, Real- und Förderschulen, so zeigt sich eine signifikante ($p<.01$) Abhängigkeit ihrer Kompetenzeinschätzungen von der Nutzung der Lerngelegenheiten „Reflexion mit dem Mentor/der Mentorin“, „Beurteilung der Ausbildner/Ausbildnerinnen“ und „Beurteilung der Mentoren/Mentorinnen“ (letztaufgeführte Lerngelegenheit zeigt allerdings nur einen Zusammenhang zur

Diagnostik-Dimension, bei angenommener Irrtumswahrscheinlichkeit von 5%). Aussagen der Referendare im Gymnasium zeugen nicht von entsprechenden signifikanten Abhängigkeiten. Vielmehr kann für sie als einzig signifikanter Prädiktor „Erfahrung in der Schule“ herausgestellt werden ($p<.01$ für „Diagnostik“ und $p<.05$ für „Unterrichtsplanung“). Für die potentiellen Lerngelegenheiten „Theorievermittlung im Studienseminar“ und „Übung im Seminar“ liefert die Studie keine nennenswerten Zusammenhänge zu Kompetenzeinschätzungen. Dasselbe gilt für die Beurteilung der Lehrerbildner durch die Studierenden mit Bezug auf ein „geteiltes Grundverständnis“.

Die Befunde zu unterschiedlichen Kompetenzeinschätzungen nach Seminararten eröffnen eine problematische Perspektive auf die Entwicklung diagnostischer Kompetenz von angehenden Gymnasiallehrern. Ihre im Vergleich zu anderen Lehramtsstudierenden geringere Kompetenzeinschätzung in einer frühen Phase des Vorbereitungsdienstes lässt darauf schließen, dass die schulartspezifische Berufssozialisation schon vor der zweiten Phase der Lehrerbildung einsetzt (Abs 2006, S. 229). Der im Verlaufe des Vorbereitungsdienstes größer werdende Einfluss der Seminarart (gemessen an der Verdoppelung der Varianzaufklärung durch die Seminarart für die Dimension Diagnostik) kann dahingehend interpretiert werden, dass die Berufssozialisation Kompetenzdefizite weiter ausbaut und die zweite Phase der Lehrerbildung dieser Entwicklung durch Angebotsformen der Studienseminare nur wenig entgegensetzt.

Die regressionsanalytischen Ergebnisse verweisen darauf, dass die in Studienseminaren umgesetzten didaktischen Settings von den Referendaren offenbar nicht als Lerngelegenheiten angesehen werden. Diese Einschätzung gilt sowohl für die Darstellung von theoretischen Inhalten als auch für praktizierte Übungsformen zur Praxissimulation, was die Frage nach revisionsbedürftigen Arbeitsformen aufwirft. Die Referendare beider Seminararten schreiben zudem der Rolle von Ausbildnern bzw. Mentoren eine unterschiedliche Relevanz hinsichtlich der begleiteten Entwicklung von Reflexionsprozessen zu. In differenten Gewichtungen drücken sich voneinander abweichende Ausbildungskonzeptionen aus, die für die Gymnasialausbildung eine größere Ferne zwischen Seminaren und Ausbildungsschulen beinhalten und den Mentoren an den Schulen stärker eine Koordinierungs- denn eine Beratungsfunktion zuweisen.

Insgesamt kann die Studie von Abs als Beleg dafür gewertet werden, dass Studierende oder Referendare die in der schulischen Praxis verorteten Lernsituationen in einem stärkeren Maße als theoretische Ausbildungssituationen mit der Entwicklung von Kompetenzen attribuieren. Abs (ebd., S. 232) vermutet einen Zusammenhang dieses Phänomens mit dem Frageformat, indem die Erfragung von Kompetenzen nicht mit einer Überprüfung zugrunde liegender Wissensformen verbunden ist und die kompetenzbezogene Selbsteinschätzung deshalb auf Erfahrungen über die eigene Handlungsfähigkeit unter praktischen Rahmenbedingungen des Lehrerhandelns basiert. Kompetenzeinschätzungen im Rahmen von theoretischen und praxisorientierten Lernkontexten speisen sich hier notwendigerweise aus einer Identifikation mit dem Beruf und sind zugleich darauf ausgerichtet. Insofern stellen sie einen Indikator für die subjektive Einschätzung erfolgreichen beruflichen Handelns dar. Ungewiss bleibt, ob diese Einschätzung in

Kategorien externer Beurteilung oder des Lernerfolgs von Schülern Bestätigung findet.

7.5.2 Untersuchungen zum Erwerb unterrichtsrelevanten pädagogischen Wissens

Nölle (2002) diskutiert das Verhältnis von Theorie und Praxis für die Entwicklung professioneller Kompetenz. In diesem Zuge eröffnet die Autorin ein Spannungsfeld von zwei diametral gegenüberstehenden Positionen, die sie mit den Begriffen „Differenzthese“ und „Integrationsthese“ belegt. Differenz steht für die Annahme, dass Theorie und Praxis zwei von einander geschiedenen Bereichen angehören und deshalb nicht voneinander lernen können; Integration verweist auf das Ernstnehmen von verbreiteten Klagen über den mangelnden Praxisbezug herkömmlicher Lehramtsstudien (Bohnsack 2000, S. 56f.). Beide Positionen bejahen die grundsätzliche Notwendigkeit von Praxislernsituationen, ziehen aber unterschiedliche Schlüsse daraus. Die in der Differenzposition markierte Ausrichtung will wissenschaftliches Wissen für Reflexionsprozesse von Praxissituationen aufschließen, betont aber gleichzeitig die systematischen Unterschiede. Damit kommt es letztlich zur Entwertung der eigenen Position, schließlich gibt es nur geringe empirische Evidenz dafür, dass Lehrpersonen ihr Handeln durch pädagogische Theorien fundieren, wenn auch nicht die prinzipielle Möglichkeit einer umsetzungsorientierten Funktion von wissenschaftlichem Wissen für Praxissituationen abgesprochen werden kann. Allerdings scheint es nur wenigen Studierenden zu gelingen sich diese Potentiale zu erschließen.
Nölle intendiert mit ihrer Studie die Untersuchung kognitiver Muster über unterrichtliches Handeln von Studierenden des Lehramtes. In einem Vergleichsgruppendesign werden entsprechende Profile von Studierenden miteinander abgeglichen, die einen eher theorieorientierten Studiengang absolviert haben (reguläres Lehramtsstudium für Grund-, Haupt- und Realschule in Niedersachsen mit zwei Blockpraktika) bzw. die stärker praxisintegrierende Veranstaltungsangebote wahrgenommen haben.[74] Der letzten Gruppe wurde die Möglichkeit gegeben didaktische Konzepte zu entwickeln, diese in schulischen Unterrichtssituationen auszuprobieren und so gesammelte Erfahrungen wiederum im Seminar zu diskutieren. Diese Stichprobe setzt sich zu einem Teil aus schweizerischen Studierenden zusammen, die im Rahmen eines einphasigen Studiengangs nach dem Theoriekonzept des Cognitive Apprenticeship Approach (Collins et al. 1989) ausgebildet wurden.
Methodisch bearbeiteten alle Studierenden im ersten und im Abschlusssemester einen schriftlichen Anreiztext, in dem didaktische Konzepte (offene vs. direkte Instruktion) diskutiert wurden. Die Probanden sollten in einem offenen Antwortformat Stellung beziehen, woran eine softwarebasierte Textauswertung anschloss. Ein zweiter Schritt sah leitfadengestützte Interviews auf der Basis einer videographierten

[74] Diese Klassifizierung gilt nur für die erste Phase der Lehrerbildung in Deutschland. Interessant wäre die Untersuchung von Referendaren am Ende des Vorbereitungsdienstes, um aufzeigen zu können, wie sich die Sequenzierung von Theorie- und Praxisphasen auf die Entwicklung didaktischen Wissens auswirkt.

Unterrichtsstunde vor, die sich Studierende vorher ansahen. Die Auswertung der transkribierten Interviews mittels strukturanalytischer Verfahren sollte Aufschluss über Klassifikationen und deren Verknüpfungen gewähren, die in Richtung zu identifizierender expliziter oder impliziter Problemdefinitionen gedeutet wurden. Konkret wurden die Texte nach den Formen des Theoriebezugs klassifiziert (Nölle 2002, S. 55):

1) Theorie: nicht erkennbar – Texte ohne erkennbare theoretische Zusammenhänge
2) Theorie: additiv – Texte, in denen theoretische Elemente in eher aufzählender Weise vorkommen
3) Theorie: implizit aggregierend – Texte mit erkennbarer thematischer Abfolge, ohne deutliche theoretische Bezüge – „subjektive Theorien"
4) Theorie: klassifizierend – Texte mit theoretischen Elementen; Darstellungen mit Bezug auf diese theoretischen Elemente; deutliche, z.B. hierarchische Zuordnungen

Die ersten beiden Kategorien fanden sich in einer nur so geringen Zahl der Texte, die keine statistische Auswertung zuließ. Auffallend sind Unterschiede für die Kategorien 3) und 4). Der Darstellungsform „implizit aggregierend" konnten 60% der im Rahmen von konventionellen Studiengängen produzierten Texte und lediglich 35,7% (Deutschland) bzw. 37,5% (Schweiz) der Texte von praxisintegrierend studierenden Personen zugeordnet werden. Noch größer fallen Differenzen für die Darstellungsform „klassifizierend" aus. Die Texte der Studierenden konventioneller Studiengänge konnten der Kategorie nur zu 18,3%, die von Personen in integrierenden Ausbildungsformen zu 57,1% (Deutschland) bzw. 50% (Schweiz) zugeordnet werden. Damit ist ein überzeugender Beweis dafür erbracht worden, dass Studiengänge, die Praxis- und Theorieelemente miteinander verknüpfen, eine höhere theoretische Durchdringung unterrichtsrelevanter Wissensformen bei den Studierenden erreichen.
Das Ergebnis ist deshalb von einiger Brisanz, weil es die deutsche, akademisch gewichtete Ausbildungsform in Frage stellt. Nölle (2002, S. 56) bezieht sich erklärend auf das Konzept der „Anchored Instruction" (Cognition and Technology Group at Vanderbilt 1993a, b, 1996). Demnach sind voraussetzend für den Wissenserwerb und den Erwerb professioneller Handlungsmuster spezifische „Anker" erforderlich, um neu zu erwerbende Kompetenzschemata mit bereits vorhandenen Strukturen verknüpfen zu können. Solche Anker werden in Form von Bildern oder episodischen Elementen durch Praxiserfahrungen generiert. Sie werden zugleich durch angeleitete Reflexionsphasen unterstützt.
Die vor dem Hintergrund dieser Konzepte plausiblen Ergebnisse der Studie lassen allerdings offen, weshalb es immerhin einem knappen Fünftel der Studierenden in konventionellen Studiengängen gelingt bzw. etwa der Hälfte der Studierenden in den integrierten Studiengängen nicht gelingt, ein klassifizierendes Niveau im Umgang mit didaktischen Theorien zu entwickeln. Trotz dieser Einschränkungen liefert die Studie einen Beleg dafür, dass Ausbildungskonzeptionen wesentlich die kognitive Schemataentwicklung von Studierenden bestimmen. Für das Theorie-Praxis-Verhältnis schließt Nölle aus den in den Texten gefundenen Strukturen und Theorieverarbeitungsstufen (Nölle 2002, S. 57), „dass theoretische und empirische Erkenntnisse in komplex vernetzter Form verarbeitet und mit episodischen

Elementen angereichert sein müssen, um dauerhaft Eingang in die kognitive Struktur zu finden und zugleich den Transfer in pädagogische Handlungssituationen vorzubereiten“. In dieser Sichtweise gestaltet sich das Theorie-Praxis-Verhältnis als Transformationsverhältnis, in dem theoretische und empirische Wissensformen und reflektierte Erfahrungen wechselseitig Anker füreinander abgeben und als Ergebnis dieses Wechselbezugs eine höhere theoretische Durchdringung didaktischer Wissenselemente und Handlungssituationen erlauben. Gelingende Transformationsprozesse setzen einen hohen Grad jeweiliger Verarbeitungsintensität voraus, womit ein Erklärungsmuster dafür gegeben ist, weshalb auch in den integrierten Studiengängen ca. die Hälfte der Studierenden kein befriedigendes Niveau des Umgangs mit didaktischem Wissen erreicht.
Besorgniserregend ist schließlich das in der Studie vorgestellte Ergebnis, in einer quasi-längsschnittlichen Betrachtung lassen sich für die als „implizit-aggregierend“ kodierten Texte kaum Unterschiede zwischen Studierenden im Erst- und im Abschlusssemester finden. Hier verfehlt die Lehrerbildung offenbar ihre Zielsetzung.

7.6 Fazit

Im Kapitel wurden die großen nationalen und internationalen Schulleistungsstudien vorgestellt, weil sie die dezidiertesten Ergebnisse für das Feld Schule liefern. Wurden die Studien der 1990er Jahre (BIJU, Markus, TIMMS) von einer breiten Öffentlichkeit (und auch der Erziehungswissenschaft bzw. dem Feld Schule) kaum wahrgenommen, kam es Ende 2001 mit der Präsentation der PISA-Ergebnisse zum viel zitierten „PISA-Schock“. Damit wurde der Weg für anschließende Studien (IGLU, DESI, PISA 2003, 2006) bereitet, sich in einen breiteren Fokus der gesamtgesellschaftlichen Diskussion zu stellen, womit sich auch die betreffenden Fachwissenschaften dieser Diskussion nicht mehr entziehen konnten.
Strittig bleibt allerdings, welche Rückschlüsse aus den Leistungsdaten von Schülern gezogen und inwiefern daraus Implikationen für die Notwendigkeit einer Veränderung des Schulsystems abgeleitet werden können. Die Einordnung der Leistungsdaten deutscher Schüler im internationalen Vergleich führt zu der globalen Einschätzung, dass es den Schulsystemen anderer Länder in größerem Ausmaß gelingt, Schülerleistungen zu fördern. Methodisch betrachtet erlauben die deskriptiven Befunde zwar die Bildung von Korrelationen zwischen den Schülerleistungen und den Schulsystemen, eine kausal-attribuierende Interpretation dieser Korrelationen ist allerdings nicht möglich, weil es innerhalb der Schulsysteme zu viele Faktoren gibt, die sich auf die Ergebnisse auswirken können. Bereits Anfang der 1980er hob Fend (1982, S. 289) hervor, dass sich innerhalb der Schulsysteme bzw. innerhalb der Schulformen Unterschiede aufzeigen, die sich nicht mit der Zuordnung zum System oder der Form erklären lassen, sonder nur auf der schul- oder klassenspezifischen Ebene untersucht werden können, weil nicht das System an sich der ausschlaggebende Beeinflussungsfaktor sein muss (Helmke & Weinert 1997, S. 92).
Nimmt man also an, dass sich die Ergebnisse der Studien nicht in erster Linie durch soziokulturelle Faktoren erklären lassen (wofür es im Falle Deutschlands wenig theoretische Evidenz gibt, auch wenn die Bildungsausgaben in Deutschland für den Primarbereich und den Sekundarbereich I unterhalb des Mittelwertes der OECD-

Staaten (OECD 2007, S. 15) angesiedelt sind), liegt es nahe unterrichtliche Prozesse in den Blick zu nehmen und dahingehend zu untersuchen, ob sich hierin Gründe für die Leistungsentwicklung deutscher Schüler finden lassen. In allen genannten Studien wurde dieser theoretische Zusammenhang hergestellt und mit unterschiedlichen Methoden untersucht.
Die zentralen Ergebnisse als Ausgangspunkt für die eigene Untersuchung lauten:

- Deutsche Schüler markieren im internationalen Vergleich nicht die Leistungsspitze, sondern finden sich im mittleren und dem Bereich des unteren Drittels wieder (die Ergebnisse der IGLU-Studie bilden diesbezüglich eine Ausnahme).
- Unterricht findet in Deutschland vergleichsweise gleichförmig statt.
- Dabei dominieren lehrerzentrierte Unterrichtsformen, die auf rekonstruktive und weniger auf konstruktive Lernprozesse bei den Schülern abzielen.
- Im Vergleich der Unterrichtsfächer lassen sich Spezifika in den Verlaufsformen und den lernunterstützenden Notwendigkeiten herausarbeiten.
- Die Lehrpersonen haben Schwierigkeiten, individuelle Leistungspotentiale und Risikofaktoren bei ihren Schülern korrekt zu diagnostizieren.
- Damit fehlen Möglichkeiten individuelle Lernstandserhebungen didaktisch nutzbar zu machen (was in lehrerzentrierten Unterrichtsformen ohnehin nur eingeschränkt möglich ist).
- Die Notengebung ist gemäß der PISA-Studien stark von der sozialen Herkunft der Schüler abhängig.

Für die eigene Fragestellung lassen sich aus diesen zentralen Ergebnissen Beobachtungsdimensionen ableiten, die in ihrer Bedeutung bereits validiert sind und niedrig- bzw. hoch-inferent untersucht werden können (vgl. Kap. 9). Allerdings werden in der vorliegenden Studie nicht die Schülerleistungen erhoben, sondern es erfolgt eine Beschränkung auf eine Untersuchung des unterrichtlichen Handelns (Blömeke 2004, S. 75f.; vgl. Kap. 7.2), weil angenommen werden muss, dass sich die Schülerleistungen nicht allein auf das Unterrichtsverhalten der Praktikanten zurückführen lassen. Videogestützte Beobachtungsverfahren erlauben aber eine Abschätzung darüber, welche Lernpotentiale in den Stunden angelegt werden, welche methodischen Verlaufsformen präferiert werden und ob z.B. eine Berücksichtigung von Lernstandserhebungen für das unterrichtliche Handeln sichtbar wird.
Die internationalen Schulleistungsstudien erhärten zudem die Hypothese, dass sich verschiedene Unterrichtskonzepte unterschiedlich auf Schülerleistungen niederschlagen. Damit werden überhaupt erst die Voraussetzungen geschaffen, um etwaige Veränderungen auf der Unterrichtsebene sinnvoll untersuchen und interpretieren zu können.
Die Untersuchung von Oser (2001), in der eine relative Wirkungslosigkeit der Lehrerbildung in der Schweiz festgestellt wurde, macht auf die Notwendigkeit aufmerksam, Veränderungen für die Lehrerbildung vorzusehen. Sein Konzept der „Verarbeitungstiefe“, als standardbasiertes Portfolio von Theorie, Praxis und reflektierter Praxis wird durch die Ergebnisse von Nölle (2002) tendenziell bestätigt. Kann man aufbauend auf den Ergebnissen von Oser zu einer insgesamt

pessimistischen Einschätzung von Möglichkeiten der Lehrerbildung gelangen, eröffnen die Ansätze von Nölle Wege zur konkreten Verbesserung. In diesem Kontext ist auch das Praxisjahr konzeptionell eingebettet, weshalb eine Untersuchung sinnvoll erscheint.
Die Kompetenzanalysen von Frey (2004, 2006) verdeutlichen, dass einzelne Kompetenzklassen nicht für sich alleine betrachtet werden können, sondern sich wechselseitig beeinflussen. Dies scheint besonders für Niveauausprägungen zu gelten, wobei (übernimmt man die von Frey vorgelegte Kompetenzklassifikation) einseitig hohe Ausprägungen auf nur einer Kompetenzklasse wenig wahrscheinlich sind. Die Ergebnisse verweisen auf eine möglichst umfassende Modellierung unterrichtlicher Handlungsmöglichkeiten von Lehrpersonen, um mögliche Beeinflussungen der Kompetenzdimensionen untereinander aufzeigen zu können. Die Fokussierung der Fachkompetenzen bzw. fachdidaktischer Kompetenzen wird hier nicht geleistet. Insofern kann auch nicht die These von Frey gestützt oder widerlegt werden, die Fachkompetenzklasse würde den wichtigsten Ausgangspunkt für eine Kompetenzentwicklung darstellen.
Die Untersuchung von Abs (2006) zur Entwicklung der diagnostischen Kompetenz in der zweiten Phase der Lehrerbildung in Deutschland, strukturell sehr ähnlich angelegt wie das Praxisjahr, zeigt auf, dass Kompetenzentwicklungsprozesse in Praxisphasen möglich, aber von individuellen Beratungs- und Reflexionsprozessen abhängig sind. Damit wird ein wichtiger Theoriehintergrund für die Interpretation der Daten geliefert und noch einmal deutlich gemacht, dass Praxisbegegnungen nicht per se zu Lerngelegenheiten führen bzw. individuell förderlich genutzt werden. Gleichzeitig wird die Bedeutung der Entwicklung diagnostischer Kompetenz innerhalb der Lehrerbildung hervorgehoben, der auch die vorliegende Arbeit nachgeht.

8 Untersuchungen zur Unterrichtsqualität als Ausgangspunkt für videobasierte Beobachtungsverfahren

Theoriebildung und empirisch gewonnene Ergebnisse zur Unterrichtsqualität wurden bislang nur randständig in die Betrachtungen einbezogen, um sie an dieser Stelle genuin auf den methodischen Teil beziehen zu können und den engen Zusammenhang der eigenen Untersuchungsanlage zur qualitätsbezogenen Unterrichtsforschung auch in der Chronologie der Studie auszudrücken.
Unterricht kann aus unterschiedlichen Perspektiven (s.u.) wahrgenommen werden. Entsprechend existieren verschiedene Konzeptionen zur Beschreibung und Klassifizierung des „guten Unterrichts". Einen Überblick über konzeptionelle Ansätze gibt Ditton (2002). Er unterscheidet folgende Forschungstraditionen bzw. -richtungen:

- Didaktische Modelle
- Forschung zum Schul- bzw. Klassenklima
- Lehr-Lernforschung
- Schulqualitätsforschung

Weiterhin hebt er hervor, dass es zwischen den vier Ansätzen nur wenige wechselseitige Bezugnahmen gibt, obwohl sich durchaus Überschneidungen oder zumindest Berührungspunkte feststellen lassen. Das Verhältnis von didaktischen Modellen und der Lehr-Lernforschung ist bereits ausführlich dargestellt worden (vgl. Kap. 4.3.1) und wird an dieser Stelle nicht erneut aufgriffen.
In der Forschungslinie zur Untersuchung des Schul- bzw. Klassenklimas stehen die „Wahrnehmungen und Beurteilungen von Aspekten des Unterrichts, des Lehrer-Schüler-Verhältnisses und der Lehrperson durch die Schülerinnen und Schüler einer Klasse" (Helmke 2003, S. 33), während Forschungen zur Schulqualität sich gewissermaßen über Umwege der Frage nach der Qualität von Unterricht nähern (Ditton & Krecker 1995). Helmke (2002, S. 266) führt jedoch kritisch an, dass seiner Meinung nach Schulentwicklung im Kern Unterrichtsentwicklung sein müsse. Er hebt hervor, dass die Varianz bezüglich verschiedener Unterrichtsmerkmale zwischen den Klassen einer Schule häufig größer ist als zwischen verschiedenen Schulen. Dieser Sachverhalt bleibt in einer lediglich aggregierten Betrachtung von Daten der verschiedenen Klassen einer Schule verdeckt.
Untersuchungen zur Frage der Auswirkungen von Schulqualität auf die Qualität des Unterrichts betreffen nur eine Teildimension, wenn auch eine wichtige, dieser Forschungsrichtung.[75] Insgesamt wird hier der Zusammenhang zwischen den Ebenen Schule, Unterricht und schulische Kontextfaktoren untersucht. Die Ebene der Schule umfasst Kriterien der Schulentwicklung, der Einbeziehung von Lehrern in Entscheidungen, welche die ganze Schule betreffen oder auch der vom ganzen Kollegium getragenen Leitprinzipien. Die Ebene des Unterrichts zielt auf Fragen der tragenden Richtlinien im Bezug auf Didaktik, Diagnostik, Umgang mit Schülern und übergreifende Ziele des Unterrichts. Die Ebene der schulischen Kontextfaktoren schließlich steht für den Umgang mit Eltern, für Öffentlichkeitsarbeit sowie für die

[75] Mortimore et al. (1988, 1989) konnten z.B. nachweisen, dass Faktoren der Schulqualität, wenn sie verbindlich im Kollegium durchgesetzt werden, die Varianz zwischen den Fächern und dem Lehrverhalten im Hinblick auf etwa Leistungsanforderungen des Unterrichts verringern können.

Kooperation mit zuständigen administrativen und politischen Institutionen sowie anderen außerschulischen Einrichtungen (z.B. Vereine). Mit dieser übergreifenden und multiperspektivischen Herangehensweise ergeben sich methodische Probleme, die Fülle der Daten aufeinander zu beziehen und sinnvoll zu interpretieren.
Der Zusammenhang zwischen Faktoren der Unterrichtsqualität und des Unterrichtsklimas[76] lässt sich spezifischer beschreiben. Eder (2002) diskutiert Auswirkungen des Unterrichts- und Schulklimas auf die Bereiche:

- Leistung,
- Einstellungen zu Schule und Lernen,
- Verhalten und Befinden in der Schule,
- Soziales Verhalten,
- Selbstkonzepte,
- Interessen.

Nach dem Autor kann der Einfluss von günstigen Klimafaktoren auf die genannten Bereiche als konsistent positiv angesehen werden, wobei er insgesamt eher niedrig ausfällt (Eder 1996, 2002). Dieser Befund ist auch theoretisch evident. Das Klima beeinflusst als Mediatorvariable die prozessuale Verknüpfung von schülerbezogenen Person-Merkmalen und schulischen Ergebnisvariablen (Eder 2002, S. 221). Besonders für Leistungsvariablen stellt das Schul- und Unterrichtsklima eine notwendige, aber nicht hinreichende Voraussetzung dar und wirkt sich vor allem längerfristig auf die Herausbildung von positiven Einstellungen, Dispositionen und Selbstattribuierungen (z.B. Selbstwirksamkeit, Lernmotivation) aus (Fend 1977, 1998, 2000).

Generell bestehen zur Untersuchung von Qualitätsmerkmalen des Unterrichts verschiedene Annäherungen. Weinert, Schrader & Helmke (1989, S. 899) definieren Unterrichtsqualität im Hinblick auf Lernerfolg als „jedes stabile Muster von Instruktionsverhalten, das als Ganzes oder durch einzelne Komponenten die substanzielle Vorhersage und/oder Erklärung von Schulleistung erlaubt“. Wellenreuther (2004, S. 347ff.) wirft dieser Position eine verengende Sicht sowohl auf das Instruktionsverhalten von Lehrpersonen als auch hinsichtlich der alleinigen Berücksichtigung von Schülerleistungen vor. Die Einbeziehung von übergeordneten Erziehungszielen sowie von moralischen Standards wurde schon von Oser, Dick & Patry (1992) für die empirische Lehr-Lernforschung eingefordert.
In einer grundsatztheoretischen Herangehensweise kann man, ähnlich wie beim Kompetenzbegriff, beim Qualitätsbegriff zwischen verschiedenen Bedeutungszuschreibungen unterscheiden. Weisen fundierte Begriffsherleitungen in beiden Fällen zunächst wertneutrale Bedeutungsverständnisse aus, was meint, dass sowohl Kompetenz als auch Qualität positiv wie auch negativ ausgeprägt sein können, hat sich im allgemeinen und zumeist auch wissenschaftlichen Sprachgebrauch jeweils eine positiv konnotierte Begriffsverwendung im Sinne von

[76] Eder (2002, S. 213) hebt hervor, dass es sich bei dem Konstrukt Klima um eine auch alltagssprachlich geläufige Metapher zur Beschreibung der Qualität sozialer Verhältnisse handelt. Bezogen auf das komplexe System Schule ergeben sich damit verschiedene Facetten, die oftmals schwer voneinander abzugrenzen sind.

zugeschriebenen normativen Gütemerkmalen durchgesetzt. Analog zum Kompetenzbegriff erklärt sich die Begriffskarriere des Qualitätsbegriffs u.a. gerade mit diesen oftmals nicht präzise vorgenommenen Bedeutungsdifferenzen. Zusätzlich zu den grundsätzlich unterschiedenen Begriffsdimensionen differenziert Terhart (2000) ein Qualitätsverständnis nach normativen, empirischen und operativen Aspekten. Unter normativen Aspekten versteht er die beabsichtigten Wirkungen von Schule und Unterricht im Hinblick auf übergeordnete Erziehungsziele sowie damit verbundene bildungspolitische Intentionen. In empirischer Perspektive diskutiert er das Verhältnis von Input und Output und die Frage, wodurch sich Merkmale eines wirksamen von denen eines unwirksamen Unterrichts unterscheiden lassen. Operative Bemühungen und Bereiche, die schließlich bessere Effekte des Unterrichts in Aussicht stellen, gehen wiederum aus den empirischen Analysen hervor. Heid (2000) relativiert den Qualitätsbegriff dahingehend, dass sich Qualität nicht aus sich selbst heraus ergebe, sondern jeweils Bestimmungen in Abhängigkeit von normativen Beobachterpositionen unterliegt. Zusammenfassend entnimmt Helmke (2003, S. 41) der Diskussion um Merkmalsbereiche von Unterrichtsqualität im Kern zwei Fragestellungen:

- Welches sind die für guten Unterricht erforderlichen Kompetenzen?
- Welche Merkmale sind für die Beurteilung der Unterrichtsqualität relevant?

Damit sind die Hauptkomponenten der Angebotsseite seines Angebot-Nutzungs-Modells herausgestellt (vgl. Kap. 4.2), nämlich spezifische Lehrerkompetenzen, die die Unterrichtsqualität beeinflussen. Art und Umfang dieses Einflusses sind allerdings bislang hypothetischer Natur, weil man zwar Merkmals- und Handlungsbereiche des Unterrichts im Hinblick auf korrelative Potentialvariablen untersuchen kann, der Rückschluss auf dahinter liegende bzw. dafür benötigte Kompetenzen und Expertiseausprägungen aber lediglich ausschnitthaft möglich ist. Im Kontext dieser Problematik liefert Helmke (2003, S. 43) folgende, ein wenig tautologisch anmutende Definition:

„Der Kern der Unterrichtsqualität im engeren Sinne sind diejenigen Prinzipien und Merkmale, die für den Unterrichtserfolg ausschlaggebend sind …“

Eine erweiterte Definition für den Begriff Unterrichtsqualität nimmt Einsiedler (2002, S. 195) vor:

„In Erweiterung der o.g. Begriffsbestimmung kann man Unterricht also definieren als Bündel von Unterrichtsmerkmalen, die sich (als ‚Bedingungsseite’ oder Prozessqualität) auf Unterrichts- und Erziehungsziele (‚Kriterienseite’ oder Produktqualität) positiv auswirken, wobei die Kriterienseite überwiegend von normativen Festlegungen bestimmt ist und der Zusammenhang von Unterrichtsmerkmalen und Zielerreichung von empirischen Aussagen geleitet ist.“

Forschungen zur Unterrichtsqualität können demgemäß entweder auf die Prozess- oder/und Produktqualität fokussieren. In der Konzentration auf die Prozessseite ist die Aufmerksamkeit auf eine methodenorientierte Bestimmung der Qualität von unterrichtlichen Lehr-Lernprozessen gerichtet. Im Gegensatz dazu zielt eine

Orientierung an der Produktqualität auf die Untersuchung der unterrichtlichen Wirkungen, zentral auf die Erreichung von Bildungszielen oder Kompetenzen bzw. auf die Motivationslagen der Lernenden. Einen dritten Ansatz nennt Helmke (2006) mit der sog. „Personorientierung". Hier geht es um die Identifikation von Schlüsselkompetenzen und Orientierungen von Lehrpersonen.
Erschwerend zur Bestimmung von Unterrichtsqualität kommt mit Blick auf die theoretisch unterschiedenen Ansätze hinzu, dass ein nach normativen Prozesskriterien positiv einzuschätzender Unterricht hinsichtlich seiner Wirkungen, z.B. der Lernerfolge von Schülern, dennoch negativ beurteilt werden kann. Umgekehrt erzielt ein auf der Prozessebene kritisch zu beurteilender Unterricht möglicherweise durchaus positive Wirkungen. Es wird deutlich, dass Beziehungen zwischen Einzelmerkmalen der Prozessqualität und auszumachenden Wirkungen bzw. Produkten als probabilistische (und oft schwache), indirekte und oftmals nicht lineare Verhältniskonstruktionen zu beschreiben sind (Helmke 2006). Zudem stehen die intendierten Produkte des Unterrichts z.T in einem dynamischen Komplementärverhältnis zueinander. Trotz dieser Einschränkungen kann man festhalten, dass sich Unterrichtserfolg, gemessen an kriteriellen Wirkungen, auf verschiedene Arten und Weisen, aber nicht beliebig realisieren lässt.

Über die Unterscheidung von Prozess-, Produkt- oder Personorientierung hinaus ist die Beurteilung von Unterrichtsqualität abhängig von der jeweiligen Perspektive, die sich in akteursspezifischen Beurteilungen manifestiert (Clausen 2002). Selbsteinschätzungen in Form von Retrospektiven durch Lehrpersonen werden anders ausfallen als Beurteilungen von betroffenen Schülern, auf die der Unterricht gerichtet ist. Während sich die Schülerperspektive häufig durch affektive Beurteilungstendenzen auszeichnet (Clausen 2002, S. 188) und deshalb Verzerrungen in der Einschätzung objektiver Qualitätsdimensionen enthalten kann, ist ebenso bei Lehrern davon auszugehen, dass Qualitätsaussagen über ihr eigenes Handeln nicht rein objektiven Maßstäben folgen. Clausen hebt in seiner Studie allerdings insgesamt hervor, dass die Aussagen von Schülern oder Lehrern für bestimmte Fragestellungen optimale Validitäten aufweisen. Dies gilt besonders dann, wenn die einzuschätzenden Merkmale nur selten auftreten oder schwer beobachtbar sind (Kunter 2005, S. 198). In diesem Fall erscheint es günstig auf die Perspektive betroffener Akteure zurückzugreifen, um so über längere Zeiträume sich erstreckende Unterrichtsprozesse mit berücksichtigen zu können.[77]

[77] Mit diesem Argument wird eine Einschränkung bezüglich der Durchführung von Videostudien deutlich. Der zu betreibende Aufwand für Erhebungen und Analysen macht es fast unmöglich den Unterricht von Lehrpersonen über einen längeren Zeitraum kontinuierlich zu verfolgen, bedenkt man zudem, wie begrenzte Mittel für Forschungszwecke den Hochschulen in der Regel zur Verfügung stehen. Müller, Eichler & Blömeke kritisieren in diesem Zusammenhang die Handlungspraxis der DFG, hinsichtlich der Beantragung von Drittmitteln für die Durchführung von Videostudien hohe Anforderungen an die bereits bestehende Infrastruktur zu stellen, die praktisch nur noch von den großen Instituten (DIPF, IPN, MPIB) eingelöst werden können. Damit zeichnet sich für die Durchführbarkeit groß angelegter Videostudien eine ähnliche Entwicklung ab, wie sie bereits bei den internationalen Schulleistungsvergleichen beobachtet werden konnte, nämlich eine Konzentration wichtiger empirischer Studien auf wenige das Feld dominierende Vertreter und Institutionen in Deutschland. Ein Ausweg in Richtung der Forschungspartizipation für weniger ausgestattete Institutionen liegt evtl. in Methoden der teilnehmenden Beobachtung, die deutlich ressourcenschonender angelegt sind als Videostudien (Kucharz & Wagener 2006; Friebertshäuser 1997).

Die wissenschaftliche Perspektive auf Unterrichtsqualität wird wiederum eine andere sein als die der betroffenen Lehrer und Schüler. Clausen (2002) konnte aufzeigen, dass die methodischen Beobachtungen durch wissenschaftliche Experten von einer guten prädiktiven Validität, speziell für didaktische Dimensionen der Unterrichtsqualität zeugen und diese insgesamt zu prägnanten Urteilen neigen, womit sich Differenzen in den Unterrichtsmustern verschiedener Lehrpersonen aufzeigen lassen (Kunter 2005, S. 198). Die nachweisliche Prägnanz der Urteile aus Beobachterperspektive lässt sich zum einen mit der fehlenden Eigenbetroffenheit, zum anderen mit der – im Vergleich zu involvierten Schülern und Lehrern – insgesamt deutlich geringeren und auf das Untersuchungsmaterial beschränkten Beobachtungszeit erklären, die auch bei hoch-inferenten Beurteilungen ein Rekurrieren auf prägnante Unterrichtsteile verlangt. Eine wissenschaftliche Annäherung erfordert eine kriteriumsgeleitete und transparente Beurteilung von Unterrichtsqualität. Maßzahlen der intersubjektiven Überprüfung kriteriumsgeleiteter Einschätzungen dienen dabei der validitätsbezogenen Absicherung getroffener Entscheidungen (Clausen 2002).
In überwiegenden Fällen verfolgen empirische Studien zur Erforschung von Unterrichtsqualität einen produktorientierten Ansatz (prototypisch seien hier TIMMS und PISA genannt). Sie ziehen ausgehend von untersuchten Produkten (im Regelfall Lernleistungen) Rückschlüsse auf die Prozessqualität des Unterrichts oder auch die Kompetenzen von Lehrpersonen bzw. die Qualität der Lehrerausbildung. Eine der ersten umfassenden und internationalen Studien zur Untersuchung von Unterrichtsqualität ist die Classroom Environment Study der IEA. Sie wurde von 1978 bis 1989 in neun Ländern durchgeführt (Anderson, Ryan & Shapiro 1989). Deutschland hat sich in Form der von Weinert & Helmke geleiteten Münchner Hauptschulstudie beteiligt. Die Studie fokussierte Auswirkungen von Merkmalsbereichen der Unterrichtsqualität auf das Lernverhalten der Schüler. Neben der kognitiven Leistungsmessung der Schüler durch Intelligenz- und Leistungstests im Fach Mathematik wurden motivationale Faktoren, Spezifika des Elternhauses und schulische Kontextfaktoren kontrolliert (Helmke, Schneider & Weinert 1986; Helmke & Schrader 2001). Im Ergebnis konnte die Expertise von Lehrpersonen als wichtige Einflussvariable für die Entwicklung von Mathematikleistungen identifiziert werden (Weinert, Schrader & Helmke 1990; Helmke & Weinert 1997). Faktorenanalytisch ließen sich dabei vier unabhängige, aber integriert nutzbare Komponenten des professionellen pädagogischen Wissens voneinander unterscheiden:[78]

- Klassenführung
- Diagnosekompetenz
- Didaktische Kompetenz
- Sachwissen

Eine solche Dimensionierung des professionellen pädagogischen Wissens überzeugt auch deshalb, weil sie einer faktorenanalytischen Überprüfung standgehalten hat. Sie wird in der eigenen Studie als übergreifender Orientierungsrahmen für die

[78] Liegt in dieser Perspektive zwar die Betonung auf den Expertisemerkmalen der Lehrpersonen, stellen die Autoren aber zusätzlich die Bedeutung des Vorkenntnisniveaus bzw. der Leistungshomogenität vs. Leistungsheterogenität der untersuchten Klassen heraus.

Entwicklung hoch-inferenter Ratingdimensionen übernommen. Die faktorenanalytische Differenzierbarkeit der Dimensionen verweist jedoch auf ihre je nach Untersuchungsanlage notwendig werdende Spezifizierung. Disziplinär viel versprechend erscheinen die Dimensionen mit Bezug auf ihren möglichen fächerübergreifenden Einsatz. Wenn auch die fachdidaktische Forschung in Deutschland erst in den Anfängen steht, gibt sie Anhaltspunkte für die darauf aufbauende Entwicklung übergreifender Dimensionen. Ein entsprechendes integratives Modell legt Helmke (2006) vor, indem er zunächst die Unterrichtsqualitätsdimensionen der TIMMS-Studie für das Fach Mathematik den Unterrichtsqualitätsdimensionen der DESI-Studie für das Fach Englisch gegenüber stellt. Für die TIMMS-Studie sind die Dimensionen zu nennen:

- Kognitive Aktivierung
- Zeitnutzung, Klassenführung
- Schülerorientierung

Für die Desi-Studie führt Helmke die Dimensionen auf:

- Prozessqualität (Positive Fehlerkultur, Unterrichtsklima, thematische Motivierung, Schülerorientierung, Unterstützung, Verständlichkeit)
- Anspruchsniveau (Wichtigkeit von Kommunikation und Korrektheit, instrumentelle Motivierung)
- Integration der Hausaufgaben
- Klassenführung (Management, Störungsfreiheit)
- Tempodruck
- Verwendung von Englisch als Unterrichtssprache
- Strukturiertheit

Aus den Qualitätsdimensionen beider Studien destilliert der Autor die sog. „Big Ten“ als fächerübergreifende Merkmale erfolgreichen Unterrichts:

- Effiziente Klassenführung und Zeitnutzung
- Lernförderliches Unterrichtsklima
- Vielfältige Motivierung
- Klarheit, Verständlichkeit
- Wirkungs- und Kompetenzorientierung
- Schülerorientierung, Unterstützung
- Förderung aktiven, selbstgesteuerten Lernens
- Angemessene Variation von Methoden und Sozialformen
- Konsolidierung, Sicherung, Intelligentes Üben
- Individualisierung und Differenzierung: Umgang mit heterogenen Lernvoraussetzungen

Betrachtet man die zehn von Helmke als Ergebnis großer empirischer Studien herausgestellten Dimensionen, so fällt eine hohe Übereinstimmung zu den von Meyer (2004) formulierten „zehn Merkmalen guten Unterrichts“ auf (z.B. Klarheit, lernförderliches Klima, Anteile echter Lernzeit). Im Unterschied zu Helmke nennt

Meyer „sinnstiftendes Kommunizieren“, „transparente Leistungserwartungen“ und eine „vorbereitete Umgebung“ als zusätzliche Merkmale. Auch wenn die zehn Merkmale seiner Aussage nach weitgehend auf empirischen Daten und Erfahrungen beruhen (und vermutlich deshalb mit den Dimensionen von Helmke fast identisch sind), spricht sich Meyer dafür aus, die Kriterien nicht allein aus den Ergebnissen der empirischen Unterrichtsforschung abzuleiten, sondern eine bildungstheoretisch-normative Orientierung beizubehalten (ebd.). Vor dem Hintergrund dieser Position kann man in einem nochmaligen Vergleich der Dimensionen von Helmke und Weinert festhalten, dass trotz der vergleichsweise geringen Differenzen eine grundlegende Unterscheidung in der Herangehensweise feststellbar ist. Während Meyer längerfristige Prozesse der Unterrichtsplanung und -durchführung mitdenkt, insofern er eine zeitliche Entwicklungsdimension als Grundlage der didaktischen Vermittlung berücksichtigt, entwirft Helmke Dimensionen, die einer empirischen Überprüfung (über verschiedene Methoden) standgehalten haben. Dabei sind Verkürzungen, die sich aus einer summativen Forschungsperspektive ergeben, nicht gänzlich auszuschließen.

Die von Helmke genannten Dimensionen bilden die Grundlage für die Entwicklung von hoch-inferenten Ratingverfahren für die eigene Studie (vgl. Kap. 10.5). Zusätzlich wurden die von Meyer genannten Merkmale „transparente Leistungserwartungen“ und „vorbereitete Umgebung“ mitberücksichtigt. Das Merkmal „sinnstiftende Kommunkation“ konnte nicht verwandt werden (womit die prinzipielle Wichtigkeit des Merkmals nicht eingeschränkt werden soll), weil es als Grundlage für Beobachtungsverfahren zu wenig handhabbar erschien.

9 Anlage der Videostudie zur längsschnittlichen Untersuchung unterrichtsrelevanter Kompetenzen in verlängerten Praxisphasen

9.1 Fragestellungen

Die vorliegende Studie ist inhaltlich eingebettet in umfassende Untersuchungen zur Wirksamkeit der Lehrerbildung (Oser & Oelkers 2001), der Modellierung handlungsrelevanter Kompetenzprofile (Frey 2004, 2006) bzw. der Entwicklung professioneller Kompetenzen von Studierenden des Lehrerberufes (Baumert & Kunter 2006). Spezifischer geht es um die Auswirkungen der Praxisphasen innerhalb der Lehrerbildung auf die Kompetenzentwicklung von Studierenden.
Wird auf einer politischen und organisatorischen Ebene die Bedeutung von Praxisphasen innerhalb der Lehrerbildung nicht in Frage gestellt (Oelkers 1999), finden sich in der erziehungswissenschaftlichen Forschung populäre Diskussionsstränge, in denen die Relation zwischen wissenschaftlich gewonnenem Wissen und einer daraus ableitbaren unmittelbaren Handlungsrelevanz für die Praxis in Frage gestellt wird (Dewe 1998) oder die Einschätzungen von Studierenden bezüglich der Bedeutung von Praxisphasen für ihre Lernentwicklung problematisiert werden (Hascher 2005, 2006). Die theoretische Fragestellung, welche Handlungsorientierungs- oder Reflexionsfunktion wissenschaftlich gewonnenem Wissen in Praxisbezügen zukommen kann, wird für die Entwicklung der Instrumente und die Erhebungsphase nicht weiter verfolgt. Für die abschließende Interpretation und Einordnung der Ergebnisse kann aber auf diese Diskussionsstränge zurückgegriffen werden.
In der grundsätzlichen Anlage der Studie wird die Teilnahme am Praxisjahr als unabhängige Variable („treatment“) modelliert, die als Ganzes auf die Kompetenzentwicklung der Studierenden als zu untersuchende abhängige Variable einwirkt. Dabei wird bewusst der Nachteil akzeptiert, dass innerhalb des Modellversuches verschiedene Einflussgrößen auf die Studierenden einwirken (z.B. Peergroup, Verhältnis zu den Mentoren oder Mitarbeitern des betreuenden Seminars). Eine differenzierende Analyse der möglichen Einflussvariablen leistet diese Dissertation nicht, weil die These vertreten wird, dass die Begegnung mit schulischen Praxiskontexten – für fast alle Studierende der erste Berufskontakt – den gravierensten Einflussfaktor darstellt. Im Vergleich dazu fungieren andere Einflussfaktoren lediglich als Mediatorvariablen, weshalb ihre Vernachlässigung für mich vertretbar erscheint. Die Untersuchungen von Hascher (2005, 2006) erlauben darüber hinaus die Auffassung, dass Studierende ihre ersten Praxisbegegnungen als existenziell erleben (diese deshalb überschätzen) und begleitende Erfahrungen diesen unterordnen, wobei man andererseits Zweifel darüber anmelden kann, inwieweit die Studierenden in der Lage sind, ihre Praxisbegegnungen objektiv als Lerngelegenheiten einzuschätzen und beurteilen können, welche Lerngelegenheiten sich förderlich oder eher hemmend auswirken. Ähnliche Kritikpunkte einer fehlenden kriteriellen Objektivität und Validität der diesbezüglichen Aussagen lassen sich für die anderen am Modellversuch beteiligten Akteure anführen, womit ein weiteres Argument geliefert ist, das Praxisjahr als ganzes zu untersuchen. Berücksichtigt man die Ergebnisse von Hascher, aus denen eine grundsätzliche Überschätzung von

Praxiserfahrungen ableitbar ist, so kann man zu dem Ergebnis kommen, dass verlängerte Praxisphasen innerhalb des Studiums (Modellversuch) oder im Rahmen der zweiten Phase der Lehrerausbildung (Vorbereitungsdienst) wesentlich der Berufssozialisation dienen. Im Rahmen dieser Berufssozialisation soll untersucht werden:

(1) Lassen sich Entwicklungsprozesse im unterrichtlichen Verhalten der Studierenden aufzeigen?

Insofern die Befundlage zur Untersuchung von Praxislerngelegenheiten in ihren Auswirkungen auf Kompetenzentwicklungsprozessen bei Studierenden sich heterogen darstellt (vgl. Kap. 6), wird die Frage absichtlich ergebnisoffen formuliert, da keine eigenen sinnvoll formulierbaren Hypothesen vorliegen. In der ergebnisoffenen Fragestellung ist die Möglichkeit angelegt, dass sich keine linearen Entwicklungen feststellen lassen. Besonders im Hinblick auf Kompetenzentwicklungsprozesse bestehen Unklarheiten, inwieweit diese nicht lineare, sondern kurvilineare oder sogar zirkuläre Funktionen aufweisen (Blömeke 2007, S. 22). Insgesamt kann man allerdings annehmen, dass verlängerte Praxisphasen einen Einfluss auf das unterrichtliche Verhalten aufweisen. Als angenommener wichtigster Faktor wird damit die Verweildauer in Praxislerngelegenheiten über drei Messzeitpunkte (vor dem Praxisjahr, zu einem frühen Zeitpunkt im Praxisjahr, zu einem späten Zeitpunkt im Praxisjahr) modelliert. Damit sind die zweite und dritte Forschungsfrage aufgeworfen:

(2) Lassen sich zwischen den drei Messzeitpunkten signifikante Veränderungen aufzeigen? bzw.

(3) In welche Richtung entwickeln sich diese Veränderungen (so sie denn festgestellt werden können)?

Methodisch geht es darum, in Form videogestützter Beobachtungsverfahren zu untersuchen, welche Auswirkungen verlängerte Praxisphasen auf das unterrichtliche Handeln der Studierenden haben. Beim Design der Studie handelt es sich um eine explorative, längsschnittlich angelegte quasiexperimentelle Interventionsstudie mit Pre- und Posttest. Eine Randomisierung oder Parallelisierung der Stichprobe war aus organisatorischen Gründen und wg. der Größe der Stichprobe nicht möglich. Um die Reliabilität der Veränderungsmaße zu erhöhen, wurden zwei Messzeitpunkte in das Praxisjahr gelegt. Prinzipiell wäre es aus methodischen Gründen sinnvoll die Anzahl der Messzeitpunkte zu erhöhen, um die Reliabilität der Differenzmaße zu verbessern. Diese Vorgehensweise wurde aber nicht gewählt, weil damit nicht zu kontrollierende Interventionseffekte verbunden sein können und der damit verbundene Aufwand nicht geleistet werden konnte.
Absichtlich wird keine hypothesenüberprüfende, sondern eine explorative Vorgehensweise gewählt, weil keine Aussagen zu einem spezifisch gerichteten Interventionseffekt möglich waren. Die Analyse von Veränderungen zählt sicherlich zu den schwierigsten Aufgabenstellungen innerhalb der Sozialwissenschaften (Bortz & Döring 2006, S. 547ff.), da sich auf jeweilige Instrumente zurückführbare

Messfehler über die Messzeitpunkte summieren, womit sich die Validität der Differenzwerte verringern kann (ebd.). Hohe Reliabilitätswerte der Beobachtungsinstrumente entkräften diese Problematik.
Im Unterschied zum Forschungsdesign des Gesamtprojektes (Dörr & Küster 2006; Baer et al. 2007) wird hier nur auf das unterrichtliche Verhalten abgehoben, welches von der didaktischen Kompetenzentwicklung abgegrenzt werden soll. In diesem Sinne wird nur das Verhalten in Unterrichtssituationen selbst, nicht jedoch das vorangegangene Planungsverhalten oder anschließende Reflexionen über den Unterrichtsverlauf untersucht. Zur Annäherung an die Forschungsfragen wurden Beobachtungsinstrumente konstruiert, die sich an bereits bestehende Instrumente anlehnen und sich in niedrig- und hoch-inferente Zugänge unterscheiden lassen (Clausen, Reusser & Klieme 2003; Seidel 2003; Kocher & Wyss 2006).
Mit niedrig-inferenten Kodierverfahren sollen basale unterrichtsmethodische Verläufe erfasst werden. Die Kodierung der Videos erfolgt mit der am IPN entwickelten Software „Videograph“ (Rimmele 2006), die statistische Weiterverarbeitung der Daten mit Tabellenkalkulationsprogrammen (Excel, SPSS).
Hoch-inferente Beobachtungsverfahren sollen eine Einschätzung der Qualität der videographierten Unterrichtsstunden erlauben. Die Fragestellungen der Studie werden zwecks umfassender Analysemöglichkeiten auf beide methodischen Zugänge bezogen. Methodisch interessant sind dabei die vierte und fünfte Fragestellung:

(4) Lassen sich niedrig- und hoch-inferente Beobachtungszugänge methodisch und inhaltlich zur Validierung aufeinander beziehen? und

(5) Weisen die Ergebnisse der beiden Instrumente in eine ähnliche Entwicklungsrichtung (so denn überhaupt Entwicklungsdynamiken festgestellt werden können)?

Die Ergebnisse der Studie können zur Generierung von Hypothesen und deren Überprüfung in experimentellen Designs dienen.

9.2 Design

Das Design der vorliegenden Arbeit ist eingebettet in ein umfassendes Forschungsvorhaben zu einem Praxisjahr innerhalb der Lehramtsausbildung (Dörr & Küster 2006; Barthold, Küster & Müller 2006). In diesem Gesamtdesign des Modellversuches werden neben der hier vorgestellten Beobachtungsmethodik auch verschiedene schriftliche Befragungen und Interviewtechniken eingesetzt (vgl. Müller i.V.). Die schriftlichen Befragungen bestehen aus Selbsteinschätzungen zur Kompetenzentwicklung, persönlichkeitsdiagnostischen Verfahren und Selbstwirksamkeitsuntersuchungen. Zusätzlich werden den Studierenden Fragestellungen zur Unterrichtsplanung vorgelegt, sog. „Vignettentests“ (vgl. Baer et al. 2007), um die didaktische und diagnostische Planungskompetenz zu untersuchen, und Interviews durchgeführt, mit dem Ziel einer Rekonstruktion der Binnenperspektive der Studierenden während der Teilnahme am Praxisjahr.

Die hier vorliegende Arbeit befasst sich im empirischen Teil mit einer längsschnittlich angelegten Videostudie über drei Messzeitpunkte, in die niedrig- und hoch-inferente Beobachtungsverfahren integriert sind. Zum ersten Messzeitpunkt wurden die Studierenden, die ihre freiwillige Teilnahme am Modellversuch zugesagt haben, jeweils an einem Unterrichtstermin innerhalb des ersten allgemeinen Tagespraktikums gefilmt. Im Regelfall absolvieren die Studierenden das Tagespraktikum im zweiten Semester und erhalten innerhalb einer Gruppe von sechs Studierenden einmal in der Woche abwechselnd die Gelegenheit eine Unterrichtseinheit zu planen und alleine oder im Tandem durchzuführen. Die Planung der Unterrichtseinheiten erfolgt meistens in Absprache mit den Ausbildungslehrern der jeweiligen Schulen und gemeinsam mit Lehrenden der Hochschule, welche die Praktikumsgruppen betreuen.
Um die Teilnahme am Praxisjahr als Treatmenteffekt zu untersuchen, wurden zwei weitere Messzeitpunkte zu Beginn (2. Messpunkt) bzw. zum Ende (3. Messzeitpunkt) des Praxisjahres angelegt. Auch hier wurden durchgeführte Unterrichtseinheiten der Studierenden mit zwei Kameras videographiert. In der theoretischen Anlage wird die Teilnahme am Praxisjahr in Gänze als unabhängige Variable modelliert, weil andere denkbare Einflussfaktoren, z.B. Vorwissen und Erfahrungen der Studierenden, über die gewählte Methodik nicht untersucht werden können. Es besteht aber die Annahme, dass die Untersuchungsergebnisse aus dem Tagespraktikum als Eingangserhebung dienen, von der aus eine Beurteilung der Entwicklung innerhalb des Praxisjahres erfolgen kann. Methodisch zentral ist zudem eine Integration von niedrig- und hoch-inferenten Beobachtungsverfahren, deren Ergebnisse wechselseitig aufeinander bezogen werden können.

9.3 Beschreibung der Stichprobe

Die Stichprobe setzt sich aus elf Studierenden zusammen, die das Praxisjahr[79] vom Anfang bis zum Ende absolviert haben. Die eigentlich für den Modellversuch vorgesehenen 16 Praktikantenplätze konnten nicht im vollen Umfang vergeben werden, da einige Studierende vor Beginn des Schuljahres von Ihrer Zusage zurückgetreten sind und kurzfristig für sie kein Ersatz gefunden wurde. Konnte die Kohorte am Anfang mit zwölf Studierenden beginnen, so hat ein Studierender zu einem relativ späten Zeitpunkt das Praxisjahr noch abgebrochen. Somit lag für diesen Studierenden kein vollständiger Datensatz vor, weshalb seine Daten auch insgesamt nicht berücksichtigt wurden.
Die verbliebenen elf Studierenden sind zu 64% weiblich und zu 36% männlich. Die gleichen Prozentverteilungen ergeben sich für den Beginn des Praxisjahres im dritten bzw. vierten Semester. 73 % der Studierenden beginnen das Praxisjahr im dritten Semester, 27 % im vierten.
Der Altersdurchschnitt beträgt 22 Jahre (Min: 20 / Max: 28). Voraussetzung für die Teilnahme war die Absolvierung des ersten allgemeinen Tagespraktikums, welches dementsprechend von allen Teilnehmern belegt wurde. Die Prüfungsordnung der Pädagogischen Hochschulen in Baden-Württemberg sieht fachwissenschaftlich eine

[79] Zur Beschreibung der spezifischen Modalitäten des Modellversuches s. Anhang.

Orientierung an drei Fächern vor (Hauptfach, Leitfach, affines Fach), die mit unterschiedlicher Intensität, gemessen an den Semesterwochenstunden, studiert werden müssen. Dabei greift die Einschränkung, dass mindestens in einem dieser Fächer Mathematik oder Deutsch gewählt werden muss. Auffallend an der Stichprobe ist, dass das Fach Mathematik deutlich unterrepräsentiert ist. Als Hauptfach hat keiner der Studierenden das Fach Mathematik gewählt, jeweils nur ein Studierender im Leitfach bzw. affinen Fach. Dagegen haben 45% das Fach Deutsch als Hauptfach, 27% als Leitfach gewählt. Zusammengefasst ergibt sich damit eine deutliche Präferenz des Faches Deutsch, woraus allerdings keine Schlussfolgerungen für die eigene Fragestellung gezogen werden kann, zumal diese nicht fachwissenschaftlich oder -didaktisch ausgerichtet ist. Bezogen auf die anderen Fächer lassen sich keine eindeutigen Präferenzen aufdecken, vielmehr findet man eine Streuung der möglichen Fächerkombinationen.
Auffallend an der Stichprobe ist zudem der signifikante Vorzug der Studienausrichtung „Lehramt an Grundschulen“ gegenüber der Hauptschule durch die Studierenden.
Innerhalb der Organisation des Modellversuches konnte dieser Präferenz allerdings nicht nachgegeben werden, weil die Studierenden möglichst gleichmäßig an die zwei beteiligten Schulen, eine Grundschule und eine Hauptschule, aufgeteilt werden sollten. Diese Vorgehensweise entspricht auch der Einstellungspraxis der Schulen, die sehr nachfrageorientiert ist, womit auch nicht immer auf den Studienschwerpunkt Rücksicht genommen werden kann, sondern möglichst schnell eine Nachfrage bedient werden muss.
Auch bei den Videoaufnahmen wurde der Studienschwerpunkt nicht berücksichtigt, sondern vor allem Wert darauf gelegt, dass die Aufnahmen für alle Studierenden zu einem annähernd gleichen Zeitpunkt erfolgten, um Auswirkungen durch eine kürzere bzw. längere Verweildauer im Praxisfeld ausschließen zu können.
In der Konzeption der Studie war vorgesehen, Daten über die Teilnehmer des Modellversuches an drei Messpunkten (im ersten allgemeinen Tagespraktikum, zu einem frühen bzw. späten Zeitpunkt im Praxisjahr) in Form von Videoaufnahmen zu erheben. Die Studierenden wurden frühzeitig ausgewählt, um sie im Semester vor Beginn des Praxisjahres bereits im Tagespraktikum aufnehmen zu können.

10 Entwicklung von videobasierten Beobachtungsverfahren

Nicht zuletzt auf Grund von technischen Entwicklungen und dadurch sich erweiternden Erhebungs- und Auswertungsmöglichkeiten haben sich Videostudien in der empirischen Unterrichtsforschung, neben traditionell verwendeten Methoden (Fragebogen, Interviews, teilnehmende Beobachtungen im Klassenzimmer), als probate Methode durchgesetzt. Waren noch vor 30 Jahren hoch technisierte, immobile Mitschauanlagen Voraussetzung für die systematische Unterrichtsbeobachtung, kann heute dank mobiler, digitaler Aufzeichnungsgeräte Unterricht potentiell in jeder Klasse und unabhängig von Standort und Infrastruktur der Schule aufgezeichnet werden. Die Bedienung der Kameras hat sich mittlerweile deutlich vereinfacht und ist auch Laien innerhalb kurzer Zeit zugänglich. Auch die Anschaffungskosten für das Equipment haben sich im Laufe der Zeit verringert.
Damit sind wichtige Voraussetzungen geschaffen worden, Unterricht ökologisch valide aufzuzeichnen und mögliche Interventionseffekte durch das Aufzeichnungssetting gering zu halten, was einen Fortschritt gegenüber traditionellen Quasi-„Laborstudien“ bedeutet. Die technische Entwicklung im digitalen Bereich erlaubt es zudem Aufnahmen in hochwertiger optischer Qualität zu erreichen. Auch garantieren Tonaufzeichnungen über die Koppelung von Funk- und Richtmikrofonen ein Niveau, das Unterrichtsaufnahmen (meistens) ohne störende Hintergrundgeräusche ermöglicht.[80]
Hinsichtlich der Auswertung stellt die Tonaufzeichnung bei Videostudien die größere Herausforderung dar. Sie lässt sich damit erklären, dass unterrichtliche verbale Kommunikation hochkomplex ist und für die Analyse, in Abhängigkeit der Fragestellungen, einen großen Erkenntniswert verspricht. Digitale Aufzeichnungsverfahren mit ihrer hohen Kompatibilität zu Computersystemen und Softwarelösungen erleichtern die Datenverwaltung und Auswertung. Damit können Videostudien auch bei größeren Fallzahlen eingesetzt werden und ermöglichen die Generierung von Video-Survey-Studien, wenn dem auch durch zeitlich aufwendige Erhebungen und Auswertungen Grenzen gesetzt sind (Stigler, Gallimore & Hiebert 2000).
Die ersten umfangreicheren Videostudien wurden im Rahmen der TIMMS 1995 Videostudie durchgeführt (Stigler et al. 1999). Hier konnte über videographierte Unterrichtsstunden ein repräsentativer Querschnitt von der Unterrichtsrealität der beteiligten Länder abgebildet werden, womit auch Rückschlüsse auf die jeweiligen Unterrichtskulturen möglich wurden (Waldis et al. 2006). Es kam zur Generierung von qualitativ neuen Aussagen, die durch zuvor eingesetzte Verfahren der empirischen Unterrichtsforschung, die in der Regel auf die Reduktion von Komplexität gerichtet waren und sich zumeist an indirekt gewonnenen Daten orientierten (Blömeke, Eichler & Müller 2003), nicht gewonnen werden konnten. Videostudien erlauben es dagegen, das unterrichtliche Geschehen selbst in den Blick zu nehmen (Stadler 2003). Insofern entsprechen videobasierte Aufzeichnungen

[80] Aus eigener Erfahrung kann ich allerdings sagen, dass Funkmikrofone teilweise sehr empfindlich reagieren und die Tonqualität von zwei nacheinander aufgenommenen Unterrichtsstunden deutlich voneinander abweichen kann, ohne dass Gründe offensichtlich sind. Videoaufnahmen mit Richtmikrofonen sind diesbezüglich weniger fehlersensitiv und liefern insgesamt eine akzeptable Tonqualität, ermöglichen es aber nicht jedes Wort zu verstehen.

keiner Momentaufnahme, sondern berücksichtigen die zeitliche Dimension des Untersuchungsausschnittes im vollen Umfang. Reduktion und Abstraktion fallen im Vergleich zu anderen methodischen Verfahren deutlich geringer aus, weil etwa im Unterschied zu teilnehmenden Beobachtungen keine Übersetzung in den hoch abstrakten Code der Sprache notwendig ist (Kommer & Biermann 2005). Fehlende Reduktions- und Abstraktionsschritte im Erhebungsrahmen sind gleichwohl verantwortlich für den hohen Aufwand bei der Analyse von Unterrichtsvideos, insofern themen- bzw. kategorienfokussierende und theorieverknüpfende Prozesse erst im Nachhinein geleistet werden. In der Darstellung bewegter Vorgänge, wie z.B. Handlungs- und Ereignisfolgen, können Entwicklungen und Veränderungen dokumentiert werden, womit sich Videos besonders zum Nachvollzug prozeduraler Wissensaneignung eignen (Ballstaedt 1997, S. 255). Neben der Auswertung größerer Fallzahlen in Form von Survey-Studien liefern sie auch die Grundlage für Fallstudien, die etwa in der fachdidaktischen Forschung überhaupt erst einmal Vorstellungen und Ansatzpunkte über jeweils vorfindbare Spezifika aufschließen (Fischer 2006). So beschreiben Waldis et al. (2006, S. 156) die Ausgangslage ihrer Videostudie im Fach Geschichte dahingehend, dass nur wenige Informationen zur Unterrichtspraxis vorliegen und dementsprechend eine solide Datenbasis fehlt.
Neben beschriebenen forschungsmethodischen Ansätzen dienen Unterrichtsvideos als didaktisches Medium, das als Anschauungsmaterial bzw. Reflexionsfolie dient oder Beispiele für gelungenen bzw. misslungenen Unterricht aufzeigt (Krammer & Reusser 2006; Stadler 2003).

10.1 Möglichkeiten und Grenzen von Videostudien

Videostudien bieten den Vorteil, dass das Unterrichtsgeschehen mit Hilfe der Videos von mehreren Beobachtern zeitlich unabhängig voneinander (und vom Zeitpunkt der Aufnahme) untersucht werden kann und gewähren so einen mehrperspektivischen Zugang. Dabei können Experten verschiedener Fachgebiete und mit unterschiedlichen Fragestellungen eingesetzt werden. Deren wechselseitiger Austausch von Ergebnissen und episodischen Beispielen trägt wesentlich zur Erhöhung der Ergebnisvalidität bei.[81] Weiterhin können Unterrichtsvideos potentiell wiederholt untersucht und somit für Sekundäranalysen verwendet werden, zudem haben sie auch eine zeitgeschichtliche Bedeutung. Trotzdem stellt die jeweilige Unterrichtsaufnahme immer nur einen Ausschnitt der Unterrichtsrealität dar. So sollte man die Erkenntnisreichweite von Unterrichtsvideos, die keinen Einblick in die Köpfe der Akteure gewähren, nicht überschätzen. Videobezogene Unterrichtsanalysen bilden letztlich nicht die Realität des Unterrichts an sich ab, sondern basieren immer auf indikatorengestützten Untersuchungsanlagen. Die Indikatoren sind vom

[81] Unterrichtsvideos stellen zudem eine Transferstelle zur Verknüpfung von Theorie und Praxis dar. Sie liefern wissenschaftliche Daten, die zu einem besseren Verständnis der komplexen Unterrichtspraxis verhelfen. Schulpraktikern bieten sie die Gelegenheit über das selbst- und fremdgestaltete Unterrichtsgeschehen zu reflektieren und über einen damit zusammenhängenden Austausch eine Sprache zu entwickeln, die als Grundlage für eine professionelle Kommunikation angesehen werden kann (Stadler 2003). Annäherungen zwischen Wissenschaft und Praxis werden möglich, insofern die von beiden Seiten verbreitet akzeptierte Unabhängigkeit ihrer Anstrengungen nicht länger aufrecht zu erhalten ist.

Abstraktionsniveau unterscheidbar, erlauben aber lediglich eine mittelbare Untersuchung des Unterrichts.
Mit Hilfe von Videoanalysen kann gleichwohl die Komplexität und Prozesshaftigkeit von Unterricht besser erfasst werden als das mit anderen Methoden (z.B. schriftlichen Befragungen) möglich ist. Sie eignen sich damit auch für interkulturelle Vergleiche (Lipowsky et al. 2003; Klieme & Reusser 2003; Pauli & Reusser 2003). Potentiell stellen sie ein Gegengewicht zur bekannten Schwäche pädagogischer Praxis dar, welche in der Tendenz darin besteht an Hand einzelner Beobachtungen zu einem vorschnellen Qualitätsurteil zu gelangen (Waldis et al. 2006, S. 161). Die Analyse von Unterrichtsvideos ermöglicht hingegen eine „Detailsicht" sowie ein prozessorientiertes Verständnis von Unterricht (Wild 2003).
Die Aufnahme von Unterrichtsstunden ist zunächst nicht an bestimmte Theorien oder Forschungsgruppen gebunden und kann weitgehend hypothesenfrei[82] erfolgen. D.h. inhaltliche Fragestellungen und Auswertungsstrategien müssen nicht zwingend im Vorfeld expliziert werden, womit eine Offenheit gegenüber noch unbekannten Sachverhalten aufrechterhalten wird (prinzipiell ist es aber vorteilhaft grundlegende Fragen schon im Vorfeld theoretisch zu klären, um besonders interessierende Unterrichtsaspekte bereits in der Datenerhebungsphase fokussieren zu können). Damit besteht die Möglichkeit induktiv, im Prozess der Auswertung, die Analyseformen zu adaptieren.
Grenzen setzen allerdings Bild- und Tonaufzeichnungsverfahren. Forschungsökonomisch sind der Einsatz beliebig vieler Kameras und Mikrofone bzw. die Auswertung so gewonnener Daten kaum vorstellbar. Die Auswahl der Anzahl von Kameras und Mikrofonen stellt somit eine Einschränkung dar, wonach immer nur ein Teilbereich der Unterrichtsrealität abgebildet bzw. analysiert werden kann. Wie oben angedeutet gilt diese Einschränkung besonders für Tonaufzeichnungen, da nicht jeder Schüler mit einem eigenen Funkmikrofon ausgerüstet werden kann. Die Auswahl der Kameraperspektive oder der Einsatz von Zoomfaktoren wirken sich ebenfalls einschränkend auf die Untersuchungsperspektive aus.

Betrachtet man den Datenreichtum von Videostudien als Untersuchungsvorteil, beinhalten sie zugleich den Nachteil, quasi in der Fülle der Daten zu „ertrinken". Damit werden hohe methodische Anforderungen an die Forscher und gewählte Auswertungsstrategien gestellt (Seidel 2003b, S. 103). Als problematisch ist weiterhin der hohe zeitliche und inhaltliche Aufwand in allen Phasen von Videostudien einzuschätzen, angefangen mit den Schwierigkeiten im Zuge der Rekrutierung von Stichproben. In Deutschland herrscht eine Kultur des „geschlossenen Klassenzimmers", die die Suche nach Personen, die freiwillig bereit sind sich filmen zu lassen, schwierig gestaltet. In der vorliegenden Studie wurde das Problem dadurch gelöst, dass Teilnehmende am Praxisjahr im Vorfeld Informationen über das Forschungsvorhaben erhielten und wiederum deren Teilnahme an der Unterrichtsvideographie verpflichtend gemacht wurde.

[82] Dieses Argument ist in einem forschungspragmatischen Sinne zu verstehen. Wissenschaftstheoretische Positionen, in denen Fragen grundsätzlicher Erkenntnismöglichkeiten über hermeneutische Verfahren diskutiert werden, bleiben davon unberührt (vgl. z.B. Popper 1935/2005, 2000). Hall (2000, zit. n. Petko et al. 2003) stellt die These auf, dass auch Bild- und Tonaufnahmen immer nach Selektionskriterien erfolgen, wonach die scheinbar neutralen Daten durchaus theoriegeladen sind.

Die eigentliche Erhebungsphase, Digitalisierung der Videos, ggf. Transkription, Aufbereitung und Auswertung der Daten sind ebenfalls mit einem hohen zeitlichen Ressourcenaufwand verbunden. In der eigenen Studie wurde auf eine Transkription der Unterrichtsaufzeichnungen verzichtet. Diesbezügliche vergleichende Analysen versprachen keinen Mehrwert, um den erheblichen Aufwand rechtfertigen zu können.
Ein weiterer problematischer Punkt betrifft die Frage der Anonymisierung. Die Unterrichtsvideos als Rohmaterial können nicht anonymisiert werden. Die aus Film und Fernsehen bekannten „Gesichtsbalken", die ein Erkennen jeweiliger Personen verhindern sollen, finden deshalb in einem frühen Auswertungsstadium keinen Einsatz, weil auf solche Weise eine veränderte Informationsdichte der Videos und also der Untersuchungsergebnisse gegeben wäre. Der damit verbundene Aufwand steht zudem in keinem Verhältnis zum Ertrag. Anonymisierte Transkriptionen bilden das komplexe Unterrichtsgeschehen schließlich nur unzureichend ab. Unterrichtliche Kommunikation läuft nicht allein auf der verbalen Ebene, sondern bedient sich auch anderer Medien, wie z.B. Tafel, Tageslichtprojektor, Bildmedien oder Aufgabenblätter. In diesem Zusammenhang können Aspekte der Mimik, Gestik oder Prosodie kaum in vollem Maße und nur aufwendig transkribiert werden (Petko et al. 2003), weshalb Auswerter jederzeit Zuordnungen zu einzelnen Personen vornehmen mögen. Diese Einschränkungen befördern nicht eben die Bereitschaft zur Teilnahme an Videostudien. Der Einsatz von computer- und softwarebasierten Auswertungsverfahren bietet allerdings wesentliche Hilfsfunktionen, um die Anonymisierung gewonnener Daten zu erleichtern und Zuordnungsmöglichkeiten zu den Rohdaten zu verstellen.
Skizzierte Anonymisierungsschwierigkeiten verweisen auch auf den wichtigen Punkt des Datenschutzes[83]. Unterrichtsaufzeichnungen von lernenden (und manchmal auch scheiternden Kindern) berühren sensible Wahrnehmungspunkte von Eltern. Diese sind deshalb – was wiederum zu einer forschungsökonomischen Erhöhung des Aufwandes führt – über ihr Einverständnis zu befragen. In der eigenen Studie fehlte im Maximalfall bei ein bis zwei Kindern der Klasse eine Einverständniserklärung der Eltern für die Unterrichtsaufzeichnung. Davon betroffene Schüler konnten in den meisten Fällen für den Zeitraum der Videoaufnahme in Parallelklassen unterrichtet werden oder bekamen einen eigenen Arbeitsauftrag. In nur einem Fall war dies nicht möglich, so dass der Schüler auf einen Sitzplatz positioniert werden musste, der nicht im Blickwinkel der Kameras lag. Mit der nur niedrigen Verweigerungsquote kann angenommen werden, dass es mit dem Fehlen der wenigen Schüler nicht zu gravierenden Veränderungen der Klassenzusammensetzung und -dynamik kam.
In diesem Zusammenhang stellt sich auch die Frage, inwieweit Videoaufnahmen eine Störung des Unterrichtsverlaufes bedingen, die ökologische Validität also nicht mehr gegeben ist. Besonders bei Studierenden, die sich noch am Anfang ihrer

[83] Damit sind rechtliche Fragen angesprochen, die der Laie nur schwer überschaut. In dieser Perspektive kann die Aufnahme von Schülern und Lehrern als Eingriff in die Privatsphäre angesehen werden. Deshalb haben alle betroffenen Personen vor der Aufnahme schriftlich ihr Einverständnis abzugeben. Neben der rechtlichen Absicherung ist die missbräuchliche Verwendung der Unterrichtsfilme durch unbeteiligte Dritte zu verhindern. In der vorliegenden Studie wurde betroffenen Personen zugesichert, dass Aufnahmen nur zu wissenschaftlichen Zwecken Verwendung finden würden, womit eine Nutzung als didaktisches Medium nicht inbegriffen war. Ohne diese Zusicherung wären die Aufnahmen nicht durchsetzbar gewesen.

Ausbildung befinden, besteht die Befürchtung, dass sie ihren Unterricht angesichts der Filmsituation anders als im Normalfall gestalten. Blömeke, Müller & Eichler (2005, S. 240) kommen in Ihrer Studie jedoch zu dem Ergebnis, dass handlungssteuernde Kognitionen verschiedener Reichweite (Skripts bzw. subjektive Theorien) für das Lehrerhandeln eine solche Stabilität aufweisen, dass auch „Vorführstunden" nur unwesentlich davon abweichend gestaltet werden können. Demgemäß wird für die eigene Studie angenommen, dass speziell angehende Lehrer noch nicht über ein Handlungsrepertoire verfügen, das grundlegende Abweichungen von ihren „normalen" Unterrichtsgewohnheiten erlauben würde.
Prinzipiell ist anzunehmen, dass die bloße Anwesenheit einer Kamera weder eine herausragende Verbesserung noch Verschlechterung des Unterrichts mit sich bringt (Extrembeispiele sind prinzipiell denkbar, konnten aber in den eigenen Aufnahmen nicht beobachtet werden). Von den Lehrpersonen außergewöhnliche an den Tag gelegte Verhaltensweisen tragen auch das Risiko in sich, ein relativ hohes Fehlerpotential aufzuweisen oder bei Schülern zu Überraschungseffekten zu führen, was an entsprechenden Kommentaren abzulesen wäre (Petko et al. 2003, S. 270). Anzunehmen ist, dass die gefilmten Stunden, zumindest hinsichtlich der Unterrichtsplanung, aus Sicht der Studierenden „Best Practice-Beispiele" darstellen. Dieser Fragestellung ging die eigene Untersuchung nicht systematisch nach. Im Anschluss an die Aufnahmen wurden die Studierenden aber mündlich befragt, ob in der Stunde etwas Ungewöhnliches geschehen oder ob die Stunde stark von der Planung abgewichen sei. Diese Fragen wurden in keinem Fall bejaht, was die Annahme zulässt, dass keine oder nur geringe Einflussnahmen des Filmsettings auf die Unterrichtsgestaltung vorlagen. Bestätigend ließen sich auch in den Aufnahmen keine Abnormalitäten beobachten.
Einflusseffekte durch die Kamera, Invasivität genannt, werden vor allem für Schüler angenommen. Jedoch ließen auch sie sich in den eigenen Unterrichtsaufnahmen in nur geringem Maße beobachten.[84] Sie zeigen sich am ehesten in schwachen Disziplinierungseffekten. Diese Annahme wird durch die Befragungen im Rahmen der TIMMS-Videostudie untermauert, wonach nur wenige Lehrer eine bedeutsame Beeinflussung durch die Videoaufnahme wahrgenommen haben. In der Tendenz wurde hingegen die Auffassung vertreten, die videographierte Unterrichtsstunde weicht nicht deutlich vom „Normalfall" ab (Seidel et al. 2003b).
Neben dem technischen Equipment kann das Verhalten der Kameraleute für Invasivitätseffekte sorgen. Im eigenen Vorgehen wurden die Kamerapersonen im Vorfeld angewiesen, sich im Moment des Filmens möglichst zurückhaltend zu geben und nicht in Interaktionen mit Schülern oder Studierenden einzutreten. Nonverbale Reaktionen auf den Unterrichtsverlauf oder das Verhalten einzelner Schüler wie auch gezielt gesuchte Blickkontakte seitens der Filmenden sollten damit ebenso unterbunden werden. In der Konzentration auf die Durchführung von Aufnahmen und die Wahrnehmung der Unterrichtsrealität vor allem durch das Kameradisplay war zugleich intendiert, das Interesse der Schüler für die Aufnahmen möglichst gering zu

[84] Einflussfaktoren durch Videoaufzeichnungen auf das unterrichtliche Geschehen, so die Diskussion in der Literatur, werden häufig unterschätzt. In diesem Zusammenhang hat es mich immer wieder überrascht, warum oftmals „gestandene" Lehrpersonen im Vorfeld offenbar der Meinung sind, sie müssten die ihnen anvertrauten Schüler vor Videoaufnahmen schützen, während die Schüler in den meisten Fällen entweder mit Desinteresse oder eher mit freudiger Erwartung auf die Ankündigung von Aufzeichnungen reagieren, wie dies in vielen Flurgesprächen in den Schulen deutlich wurde.

halten. Das scheinbar gleichgültige Verhalten der Kamerapersonen hat zwar das Störungspotential der Videoaufnahmen minimiert, wurde aber z.T. von anwesenden Mentoren als negativ, weil als unnahbar wahrgenommen. Solche Reaktionen antizipierte die ursprüngliche Anlage der Studie nicht, weil es nahe liegend schien, dass ein verhaltenes Auftreten eher positiv im Sinne eines unbeeinträchtigten Unterrichtsverlaufes aufgefasst würde. Von dieser Ursprungsannahme abweichende Erfahrungen im Untersuchungsfeld führten zu der Strategie, alle dem Unterricht beiwohnende Personen über die Gründe eines möglichst neutralen Beobachtungsverhaltens aufzuklären. Einwirkungen von Aufnahmen auf das Unterrichtsgeschehen konnten allerdings nicht immer vermieden werden. Da die Aufnahmen im angestammten Klassenzimmer der jeweiligen Klasse stattfanden und einige Aufnahmen aus organisatorischen Gründen in direkter Aufeinanderfolge durchgeführt wurden, reichte eine fünfminütige Pause zwischen den Unterrichtsstunden nicht immer aus, um für die nächste Aufnahme startbereit zu sein. Der Beginn der Stunde wurde dadurch manchmal verzögert, weshalb die videographierten Unterrichtsstunden teilweise kürzer als die fünfundvierzigminütige Standardlänge ausfallen.

10.2 Problemfelder methodischer Beobachtungsverfahren

Methodische Beobachtungsverfahren basieren auf zugrunde gelegten transparent gemachten Kategorien und Kriterien, zu deren intersubjektiver Überprüfbarkeit sie wiederum beitragen. Angestrebt wird eine objektivierte Beobachtung, deren Ergebnisse den psychometrischen Gütekriterien von Objektivität, Validität und Reliabilität standhalten. Die bereits angesprochene Informationsfülle von Unterrichtsvideos erfordert es, dimensionale Einteilungen als Selektionskriterien einzuführen, um die Untersuchung von Filmmaterial handhabbar zu machen (vgl. Kap. 10.4/10.5). Trotz solcher methodischen Anstrengungen zur Generierung von objektivierten Beobachtungsdaten geht weiterhin ein forschungsmethodologisches Problem von der in der menschlichen Informationsverarbeitung angelegten Subjektivität aus. Unterschiede von individuellen Einschätzungen der Unterrichtsmerkmale ergeben sich z.B. aus personenspezifischen biographischen Erfahrungshintergründen oder einem domänenspezifischen Kenntnisstand.
Neben freien Beobachtungsformen, die auf einer beschreibenden Ebene die Analyse unterrichtlicher Merkmale erlauben, haben sich in der videobasierten Unterrichtsforschung zwei methodische Zugänge durchgesetzt, die sich bezüglich ihres Inferenzniveaus unterscheiden lassen. Inferenz meint dabei den Abstraktionsgrad der für die jeweiligen Beobachtungsverfahren notwendigen Schlussfolgerungen, die der Beobachter leistet (Barthold et al. 2006, S. 386f.). Zugegebenermaßen ist der Inferenzbegriff eher unscharf, weil man annehmen kann, dass sich Schlussfolgerungsprozesse individuell nach Abstraktionsgrad und Differenzierungsgrad unterscheiden. Die gleich bleibende Bedingung besteht nur jeweils in der Intention, auf der Basis von Beobachtungen Unterrichtsmerkmale rekonstruieren oder einschätzen zu wollen. In der Forschungsliteratur haben sich dennoch die Begriffe niedrig- bis mittel-inferenter und hoch-inferenter

Beobachtungen durchgesetzt (Seidel et al. 2003b; Clausen, Reusser & Klieme 2003), weshalb sie auch in dieser Studie Verwendung finden.
Besonders für hoch-inferente Beobachtungsverfahren, die qualitative Aussagen über Merkmalsbereiche von Unterrichtsqualität liefern, berichten Krammer & Reusser (2005, S. 43) über verschiedene systematische Fehlerquellen, die zu verzerrten Einschätzungen führen. Hoch-inferente Ratingverfahren, die sich auf die gesamte Unterrichtsstunde beziehen, müssen etwa nach Möglichkeit alle aus der Stunde ersichtlichen Informationen berücksichtigen. Damit werden hohe Anforderungen an die Gedächtnisleistungen der Beobachter gestellt. Ein systematischer Fehler liegt z.B. vor, wenn bereits am Anfang der Stunde ein spezifisches Verhalten beobachtet wird, das prägenden Einfluss auf die weitere Einschätzung des Stundenverlaufs ausübt („Primacy-Effekt“). Eine objektive Einschätzung der gesamten Stunde ist verstellt, wenn auf ein Eingangsereignis folgende Unterrichtsereignisse nur noch einseitig positiv oder negativ als Ergebnis etwa der Stundeneröffnung attribuiert werden. Ein ähnlicher systematischer Fehler bezieht sich auf die Bedeutungszumessung, die einzelne Beobachtungsaspekte im Vergleich zu anderen erhalten. Im Hinblick auf z.B. didaktische Verhaltensweisen von Lehrpersonen werden bestimmte Merkmale gemäß der eigenen „Werte-Wichtigkeits-Hierarchie“ als derart wichtig und bedeutend eingeschätzt (positiv oder negativ), dass es zur Überlagerung anderer Unterrichtsmerkmale kommt („Halo-Effekt“). Übertriebene Bedeutsamkeitszuschreibungen können sich weiterhin auf Merkmalsbereiche beziehen, die in der gefilmten Unterrichtsstunde gar nicht beobachtbar sind und deren Fehlen aus individueller Beobachtersicht zu einer signifikant besseren, meistens aber schlechteren Einschätzung führt. Erfolgen Einschätzungen ausschließlich vor dem eigenen Erfahrungshintergrund („Wie hätte ich die Stunde gehalten“), kann man nicht mehr von einem methodisch geleiteten Beobachtungsverfahren sprechen.
Beide systematische Fehlergruppen verdeutlichen die Notwendigkeit diskursiver Prozesse im Rahmen von hoch-inferenten Unterrichtseinschätzungen, die funktional als Korrektiv fungieren und dazu beitragen, biographisch verankerte Bilder über z.B. „guten Unterricht“ zu relativieren und über Forschergruppen zu nivellieren. Hoch-inferente Unterrichtsbeobachtungen bedürfen deshalb aufwendiger Beobachtungstrainings, um genannte Einseitigkeiten individueller Beobachtung möglichst abzuwenden.
Eine weitere systematische Fehlergruppe lässt sich durch stereotype und zu verzerrten Beobachtungen führende Vorurteile gegenüber bestimmten Menschengruppen (z.B. hinsichtlich Geschlecht, Alter, Ethnie etc.) erklären. Zu dieser Fehlergruppe zählen individuelle Einschätzungen, die durch Sympathie- oder Antipathiebekundungen geprägt sind. Auch Projektionen, die in Einschätzungen in Form wahrgenommener eigener Defizitzuschreibungen oder Bedürfnislagen aufscheinen, nun aber dem zu beobachtenden Probanden zugewiesen werden, bewirken Verzerrungseffekte und sind deshalb auszuschalten. In der Forschungspraxis ist dies nur schwierig zu leisten, weil eine außergewöhnlich vertrauensvolle Interaktion vorauszusetzen wäre, die nur selten vorfindbar ist und prinzipiell auch nicht zentral für Forschungsprojekte ist.
Insgesamt verweisen die systematischen Fehlergruppen auf eine mit aufsteigendem Inferenzniveau hohe Fehlersensibilität videobasierter Beobachtungsmethoden, der

man, wenn überhaupt in Gänze, nur durch aufwendige diskursive Unterrichtsanalysen beikommen kann. Für die Beurteilung von unterrichtlichen Handlungen angehender Lehrer gilt dies im besonderen Maße, weil verschiedene Hintergrundkonzepte zur Einschätzung des „guten Unterrichts“ miteinander konkurrieren (vgl. Kap. 8). Unterricht lässt sich nicht an sich als gut oder schlecht einschätzen, auch wenn teilweise der Eindruck entsteht, dass derartige Beurteilungsmuster in der deutschen Lehrerbildung der ersten und zweiten Phase vorherrschen. Hingegen sind methodische Beobachtungen – auch im hoch-inferenten Bereich – zwingend auf indikatorengestützte Manuale angewiesen, um eine interindividuelle Überprüfung und Abgleichung gewährleisten zu können. Mit steigendem Inferenzniveau wächst die Schwierigkeit, entsprechende Indikatoren zu benennen. Altrichter & Posch (1998, S. 99) legen einen Vorschlag zur Stufung von Situationsinterpretationen vor. Dieser reicht von der konkreten Situationsbeschreibung über die Interpretation kulturell geteilter Bedeutung bis hin zur subjektiven Deutung der Situation. Krammer & Reusser (2005, S. 43) illustrieren diese Stufung an einem Beispiel:

1. Eindeutig beobachtbare Situationsmerkmale: „Die Lehrerin sagt zum Schüler: Hans, deine Leistung ist katastrophal.“
2. Kulturell geteilte Bedeutung: „Die Lehrerin tadelt Hans.“
3. Subjektive Bedeutung für einen bestimmten Hörer: „Die Äußerung der Lehrerin ist wenig einfühlsam.“

Mit ansteigender Stufung sinkt die Wahrscheinlichkeit, dass verschiedene Beobachter zu demselben Ergebnis gelangen.
Dieser kritischen Einordnung hoch-inferenter Beobachtungsverfahren steht jedoch das darin enthaltende Potential von fundierten Experteneinschätzungen entgegen. Eine reine Orientierung an indikatorengestützten Manualen würde dem Charakter hoch-inferenter Einschätzungen widersprechen und den Status von Experteneinschätzungen in Frage stellen. Statt der expertenhaften Beurteilung von Beobachtungssituationen wäre die These von der manualorientierten Zuordnungstätigkeit aufzustellen. Methodologisch stellt sich die Frage, inwieweit es sinnvoll ist, Einschätzungen – im extremen Fall – auf den kleinsten gemeinsamen Nenner zu nivellieren und damit die Varianz verschiedener Wahrnehmungen, die gleichwohl nur einen kleinen Teil der komplexen Unterrichtsrealität abbildet, weitgehend zu reduzieren.
In dieser Problematisierung der Validität von Expertenbeurteilungen spiegelt sich eine grundsätzliche epistemologische Frage, nämlich ob die Beobachtungen ein möglichst genaues Abbild der Realität an sich oder ihrer Repräsentationen liefern sollen (Clausen 2002, S. 78). Zielt man ersteres an, so müssen Abweichungen zwischen den Einschätzungen der Experten notwendigerweise als Fehlervarianz aufgefasst werden, während eine mehrperspektivische Varianz in der zweiten Herangehensweise als besondere Qualität aufscheint. Schnittstellen zwischen diesen beiden Perspektiven lassen sich aufzeigen, wenn in der Entwicklung der Manuale und in diskursiven Beobachtungsprozessen an Hand von Videobeispielen versucht wird, einen konstruktiven Verständigungsprozess über Varianzen zu erreichen.

In der vorliegenden Studie wurde der „konservative“ Weg beschritten, indem über den Abgleich von bestimmenden Beurteilungsdimensionen sowie die Berechnung von Übereinstimmungsmaßen eine möglichst hohe Beurteilerübereinstimmung ausgewiesen wird, um die Objektivität und Validität der Einschätzungen zu stützen. Für die Zukunft hoch-inferenter Videoanalysen erscheint es angebrachter, allerdings nicht unbedingt ressourcenschonender, die Varianz zwischen Experten im Sinne des jedoch wenig aussichtsreichen Versuchs einzufangen, die unterrichtliche Komplexität in individuellen Perspektiven abzubilden. Viel versprechender scheint das Vorgehen zu sein, auftretende Varianzen über das Mitteln möglichst vieler Beobachter ernst zu nehmen und über diese Mittelwerte die Reliabilität zu erhöhen (Wirtz & Caspar 2002). Aus der Warte quantitativer Forschung ist damit kein optimaler Weg eingeschlagen, forschungsethisch erscheint er mir aber plausibler als der zwanghafte Versuch einer Reduzierung von Varianz, der daran zweifeln lässt, ob Einschätzungen noch das Prädikat „hoch-inferent“ verdienen.
Einen Ausweg aus diesem Dilemma verspricht der verstärkte Einsatz niedrig bis mittel-inferenter Dimensionen. Im Vergleich niedrig- bis mittel-inferenter mit hoch-inferenten Kodierungen kann jedoch ein „Übereinstimmungs-Bedeutsamkeits-Dilemma“ (Clausen, Reusser & Klieme, S. 125) ausgemacht werden, d.h. niedrig- bis mittel-inferente Kodierungen weisen, gemessen an Reliabilitätsmaßen, oftmals eine hohe Übereinstimmung auf, während dies bei hoch-inferenten Kodierungen nur mit großem Aufwand und unter genannten Schwierigkeiten möglich ist. Beschränkungen niedrig- bis mittel-inferenter Kodierungen liegen wiederum in Aussagen, die sich lediglich auf den zeitlichen Verlauf oder die Auftretenshäufigkeit bestimmter Unterrichtsgeschehnisse oder Handlungen beziehen; insofern lässt sich dieser methodische Zugang einem eindeutig quantitativen Forschungsansatz zuordnen (Barthold et al. 2006, S. 387). Die damit gewonnenen Daten liefern aber selten für sich genommen eine schlüssige Interpretationsbasis. Sie erfordern konstitutiv eine theoriebasierte Verortung, womit sie qualitativ als weniger bedeutsam als hoch-inferente Kodierungen gelten. So ist etwa der häufig angeführte hohe Gesprächsanteil von Lehrpersonen im Unterricht für sich genommen kein Indikator für Qualitätsaussagen. Erst vor dem Hintergrund theoretischer Zielsetzungen, z.B. der aktiven Beteiligung von Schülern an der Unterrichtskommunikation als fachabhängige Zielsetzung für die Initiierung von Lernprozessen, können entsprechende Ergebnisse kritisch bewertet und hinterfragt werden. Für hoch-inferente Kodierungen nimmt man dagegen per se an, dass die in ihnen modellierten Dimensionen Bedeutsamkeit aufweisen, weil dahinter im Vorfeld immer schon geklärte abstrakte und theoriegeladene Konstrukte stehen. Den Beobachtern fällt dann die Herausforderung zu, ihre Einschätzungen möglichst im Sinne objektiv begründeter und intersubjektiv überprüfbarer Ergebnisse vorzunehmen. Kurz gesagt genügen niedrig- bis mittel-inferente Kodierungen dem Vorteil, dass sie reliabel, aber nicht unbedingt valide sind, während hoch-inferente Kodierungen per se als valide gelten, aber möglicherweise eine nur ungenügende Reliabilität aufweisen.

10.3 Unterrichtsvideos als didaktisches Medium

Die Professionalität des Lehrerberufes wird im Hinblick auf kommunikative und reflexive Kompetenzen allgemein als niedrig eingeschätzt (Stadler 2003, S. 175). Versteht man die Kommunikation zwischen Lehrkräften und Erziehungswissenschaftlern als Möglichkeit Professionalisierungsprozesse voranzutreiben, können Unterrichtsvideos diesbezüglich eine wichtige Rolle spielen. Sie eröffnen kommunikative Aushandlungs- und Interpretationsprozesse zwischen den Berufsgruppen über das Kerngeschäft der Lehrer, das Unterrichten. Dies geschieht allerdings nicht voraussetzungslos, sondern nur unter der Bedingung gelingender Verständigung und also einer gemeinsamen Sprache. Unterrichtsvideos können dafür einen besonderen Mehrwert entwickeln, weil sie ihre Informationen auf verschiedenen Symbolsystemen übertragen und damit Möglichkeiten einer hohen emotionalen und kognitiven Aktivierung in sich tragen (Weidenmann 2001). Über die Koppelung der verschiedenen Symbolsysteme gelingt zudem die Abbildung des Prozesscharakters von Unterricht in hoher Anschaulichkeit, Informationsdichte und Realitätsnähe besser, als das andere Medien (z.B. Tonband) ermöglichen (Krammer & Reusser 2005, S. 36). Damit stehen Videos nicht nur für ein aussichtsreiches Forschungsinstrument, sondern darüber hinaus für ein Anschauungsinstrument im Rahmen der Lehrerbildung.

Als didaktisches Medium für die Lehrerbildung können die zwei Grundfunktionen „Modell“ oder „Reflexionsfolie“ unterschieden werden. In der Modell- oder Illustrierungsfunktion stellen Unterrichtsvideos didaktische Handlungsformen als Ziele der Lehrerbildung anschaulich dar. Im Vergleich zu den Praktika bestehen damit Lernsettings, die den Studierenden die Diskussion von Praxiserfahrungen erlauben, ohne selbst in einer Drucksituation zu stehen.[85] Die Funktion nachträglicher Reflexion ist eine andere als die von erlebter Praxis, allerdings können damit Annäherungsprozesse an Praxissituationen geleistet werden. Die Komplexität unterrichtlicher Prozesse ist anhand von Filmmaterial besser nachvollziehbar als das ansonsten in Gesprächssituationen gelingt.

Damit ist bereits der zweite Funktionsaspekt des didaktischen Einsatzes von Unterrichtsvideos angesprochen, nämlich die Schaffung von Reflexionsgelegenheiten vor dem Hintergrund authentischer Unterrichtssituationen. Man kann die These aufstellen, dass ungeschulte Beobachter in der Betrachtung von Unterrichtsvideos zuerst auf ihr eigenes Vorwissen rekurrieren und damit nur Handlungsmuster ausschnitthaft wahrnehmen, die ihrem eigenen Unterrichtsverständnis entsprechen oder in besonderer Weise zuwider laufen. Danach beziehen sich Studierende mit wenigen Unterrichtserfahrungen eher auf Erlebnisse aus ihrer Schülerperspektive, während sich Lehrpersonen mit

[85] Mühlhausen (2005, S. 19) nimmt an, dass die durch Praktika ermöglichten Praxiserfahrungen in der Lehrbildung häufig unter ungünstigen Bedingungen initiiert werden. Danach verhindert die emotionale Involviertheit der Studierenden, dass sie eine kritische Distanz zu ihrem eigenen Handeln aufbauen. Als weiteren Negativfaktor macht Mühlhausen ungünstige gruppendynamische Konstellationen in den Praktikumsgruppen aus, die sich kontraproduktiv auf diskursive Prozesse auswirken, wobei zugleich ein ständig herrschender Zeit-Handlungsdruck eine vertiefte Beschäftigung mit gesammelten Erfahrungen verhindert. Schließlich würden Unterrichtsbesprechungen auf vagen Beobachtungseindrücken basieren, die häufig nicht methodisch geleitet generiert worden sind. Vor dem Hintergrund dieser Kritikpunkte lassen sich gute Argumente für einen verstärkten Einsatz von Videos in der Lehrerbildung finden.

ansteigender Unterrichtspraxis zunehmend auf subjektive Theorien stützen. In der wiederholten Betrachtung von Unterrichtsszenen können nun jedoch verschiedene Perspektiven eingenommen werden, was eine Flexibilisierung des Erkennens alternativer didaktischer Handlungsmöglichkeiten befördert und zu einer möglichen Veränderung subjektiver Theorien beiträgt.
Hiermit wird aber auch deutlich, dass die Betrachtung von Unterrichtsvideos nicht per se einen Lern- oder Erkenntnisgewinn verspricht, sondern strukturierende und didaktische Kontextualisierungen verlangt (ebd., S. 38). Geschieht dies nicht, muss mit unerwünschten Nebenwirkungen gerechnet werden, derart, dass intendierte Absichten der Objektivierung ins Gegenteil umschlagen und stattdessen der Bildung von Vorurteilen über die unterrichtende Person bzw. ihren Unterricht Vorschub geleistet wird. Um dem entgegenzusteuern stellen ein konstruktiver Umgang mit Unterrichtsvideos, die Trennung von Beobachtung und Interpretation, das Anleiten und Begleiten differenzierten Explizierens von Stärken und Schwächen des Unterrichtshandelns und übergreifend die respektvolle Wertschätzung der für Aufnahmen sich bereit erklärten Personen Vorbedingungen für den erfolgreichen Einsatz von Videos als didaktisches Medium dar (Krammer & Hugener 2005, S. 59). Im didaktischen Setting ist dabei die Trennung von Beobachtungs- und Interpretationsphasen nach eigener Anschauung das wichtigste Moment, da sich andernfalls eine Tendenz zum vorschnellen Bewerten ergibt.
In Abhängigkeit davon, ob die gefilmte Person selbst am Diskurs teilnimmt oder nicht, kann eine eher positive (bei Anwesenheit) oder negative Tendenz (bei Abwesenheit) ausgemacht werden. Auch wenn von einem großen Lerngewinn bei Personen ausgegangen werden kann, die ihren eigenen Unterricht aus verschiedenen Perspektiven diskursiv aufgezeigt und widergespiegelt bekommen, scheint die Arbeit mit „neutralen“ Unterrichtsvideos, an denen keine an der Diskursgruppe beteiligten Personen involviert waren, deshalb didaktisch am sinnvollsten, da für einen diskursiven Prozess in Ausrichtung auf eine konkrete und anwesende Person ein Vertrauensverhältnis vorauszusetzen ist, das in hochschulischen Veranstaltungen selten gewährleistet werden kann. Außerhalb eines entsprechenden Diskursklimas besteht die Gefahr von Videobesprechungen, die sich auf einer platitüdenhaften Ebene bewegen („Deine Unterrichtsprache war frei von Beleidigungen“).
Möchte man Videos nutzen, um eine individuelle Beratung mit Bezug auf den aufgenommenen Unterricht zu leisten, so haben sich multimediale netzbasierte Lösungen bewährt. Sie stellen Formen der synchronen wie auch asynchronen Kommunikation bereit und bieten über digitale Dateien handhabbare Lösungen der Distribuierung (Petko et al. 2003; Krammer & Hugener 2005). Auch individuelle Beratungen setzen eine sensible Vorbereitung von betroffenen Personen mit Blick auf die konstruktive Auseinandersetzung mit dem eigenen Unterrichten anhand von Videobeispielen voraus. Bei Unterlassung bestehen Tendenzen einer einseitig negativen Einschätzung des eigenen Unterrichtens in Richtung auf auszumachende Überforderungen, was einem konstruktiven Umgang mit individuellen Erfahrungen hinsichtlich Veränderungs- und Verbesserungspotentialen nicht zuträglich ist (Stadler 2003).

10.4 Entwicklung niedrig-inferenter Kategorien- und Zeichensysteme

Aufzeichnungen von Unterricht enthalten eine Fülle von Informationen. Kodiersysteme dienen als Methode der systematischen Beobachtung und Strukturierung des Materials (Barthold et al. 2006, S. 386). Niedrig- bzw. mittel-inferente Kodierverfahren verwenden Kategorien- bzw. Zeichensysteme (Seidel 2003b, S. 100), die sich auf beobachtbare Unterrichtgeschehnisse oder Handlungen beziehen. Der Grad der Inferenz lässt sich hinsichtlich der Komplexität möglicher Schlussfolgerungen für das jeweils spezifische Beobachtungsverfahren unterscheiden.
Niedrig- bis mittel-inferente Kategorien erlauben Aussagen über den zeitlichen Verlauf von Unterrichtsgestaltung oder die Auftretenshäufigkeit von Verhaltensweisen. Unmittelbare Ableitungen mit Bezug auf Lehrerkompetenzen oder Faktoren der Unterrichtsqualität können daraus noch nicht gezogen werden. Vielmehr liefern die Daten eine deskriptive, basale Beschreibung über die jeweils vorgefundene Unterrichtsrealität. Die Aussagenreichweite der Daten kann aber erhöht werden, indem diese mit den Ergebnissen anderer Instrumente verglichen, in einen größeren Theoriezusammenhang gestellt werden oder es auch zur Triangulation methodischer Zugänge kommt. So kann bspw. der oftmals berichtete hohe Gesprächsanteil von Lehrern (Kobarg & Seidel 2007; Pauli & Lipowsky 2007) an der Unterrichtskommunikation isoliert betrachtet nicht eindeutig als positiv oder negativ eingeschätzt werden. Gleichwohl besteht ein erziehungswissenschaftlich breiter Konsens darüber, dass partizipative Unterrichtsformen unter aktiver Schülerbeteiligung und -mitgestaltung günstige Voraussetzungen für Lehr- und Lernprozesse abgeben, besonders mit Blick auf die Förderung problemorientierten Lernens (Reusser 2005a). Ein lehrergesteuerter Frontalunterricht verlangt von Schülern die Anpassung an individuelle Sprachgewohnheiten des Lehrers bzw. die Rekonstruktion dahinter liegender Denkprozesse.
Untersuchungen zur Effektivität sog. „direkter Instruktionen“ im Vergleich zu eher offenen Unterrichtsformen, die basierend auf dem Prinzip der Binnendifferenzierung eine Öffnung des Unterrichts anstreben, um Schülern in höherem Maße eine Selbststeuerung ihrer Lernprozesse zu ermöglichen, kommen jedoch zu divergierenden Ergebnissen (Weinert 1996, 1998). Diese Divergenzen lassen sich auch mit methodischen Schwierigkeiten erklären, kaum alle intervenierenden Variablen erfassen zu können (Ulich 1983; zit. n. Hascher & Wehr 2005). Die Metaanalyse von Gage & Berliner (1996) präsentierte etwa hinsichtlich der Entwicklung von Schülerleistungen übergreifend bessere Ergebnisse für Formen der direkten Instruktion. Besonders die Leistungsmotivation sowie sprachliche Leistungen waren in schülerzentrierten Unterrichtsformen weniger ausgeprägt. Dagegen konnten bei den schulischen Leistungen im Rechnen und Lesen keine Unterschiede festgestellt werden. Trotz der wenig konsistenten Ergebnisse hat sich im Anschluss an Weinert (1996) die Meinung durchgesetzt, dass leistungsstarke Schüler von offenen Unterrichtsformen profitieren, während leistungsschwächere Schüler effektiver in angeleiteten Strukturen lernen, weil sie nicht in der Lage sind, eine Selektion angemessener Aufgabenformen vorzunehmen und Aufgaben kontinuierlich zielführend alleine zu bearbeiten. Insgesamt verweist die Diskussion auf die Frage, wie es gelingen kann, offene Unterrichtsformen und instruierende

Lehrerhandlungen als beiderseits notwendige Strukturierungs- und Beratungsleistungen miteinander zu verzahnen (Brunner & Schratz 2002; Forneck 2002; Gudjons 2003). Dieser Exkurs über zu interpretierende Gesprächsanteile von Lehrpersonen an der Unterrichtskommunikation als Indikator für ein instruktionistisches Unterrichtsverständnis verdeutlicht die Problematik der Aussagenreichweite der Ergebnisse niedrig-inferenter Kodierverfahren.

Die Entwicklung von Kategorien- und Zeichensystemen ist vor die Herausforderung gestellt, Indikatoren beschreibend zu fassen, die ein zu beobachtendes Merkmal auf der Handlungs- bzw. Sichtebene trennscharf operational definieren. Die eigene Videostudie beruft sich auf ein sog. „disjunktes“ Kodiersystem, d.h. jeweils nur eine beobachtete Merkmalsausprägung wird nach einem vorher definierten Stichprobenplan einem zeitlichen Intervall als Indikator zugeordnet. Als allgemeine Grundregel gilt die Kategorienzuordnung (Kodierung) zu einem festgelegten Intervall nach der Logik des jeweils beobachteten größten zeitlichen Umfangs einer Kategorie.
Die Indikatoren müssen einerseits trennscharf sein, um disjunkt zugeordnet werden zu können. Andererseits sind sie umfassend und auf einer einheitlichen Abstraktionsebene zu formulieren, um jeweilige Merkmalsausprägungen in Gänze beschreiben und erfassen zu können. Sowohl die Trennschärfe als auch die umfassende Beschreibung auf einer bestimmten Abstraktionsebene stellen nicht unerhebliche Herausforderungen dar. Dabei wird deutlich, dass nicht allein diskursive bzw. theoretisch geleitete Entscheidungen darüber bestimmen, welche Merkmalsbereiche anhand welcher Indikatoren beobachtbar sind. Die Entwicklung von Beobachtungskategorien erfolgt zwingend sowohl theorie- als auch datengestützt und rekurriert dabei auf Standards inhaltsanalytischer Beobachtungsverfahren (Bos & Tarnai 1999; Hugener et al. 2006).
Die Berücksichtigung vorliegenden Videomaterials als explizite Einflussgröße für induktiv entwickelte Kategoriensysteme bzw. die anschließende Überprüfung in Kodierungsverfahren eröffnen einen Zirkel angemessener Analyseschritte im Umgang mit Videodaten (Jacobs et al. 1999). Verschiedene Kategorien stehen innerhalb eines Kategoriensystems in einem Abhängigkeitsverhältnis zueinander, indem sie entweder bereits kodierte Sequenzen detaillierter untersuchen oder eine gänzlich neue Perspektive einnehmen. Seidel (2003b, S. 104) legt einen prototypischen Vorschlag über eine Kategorienentwicklung vor.

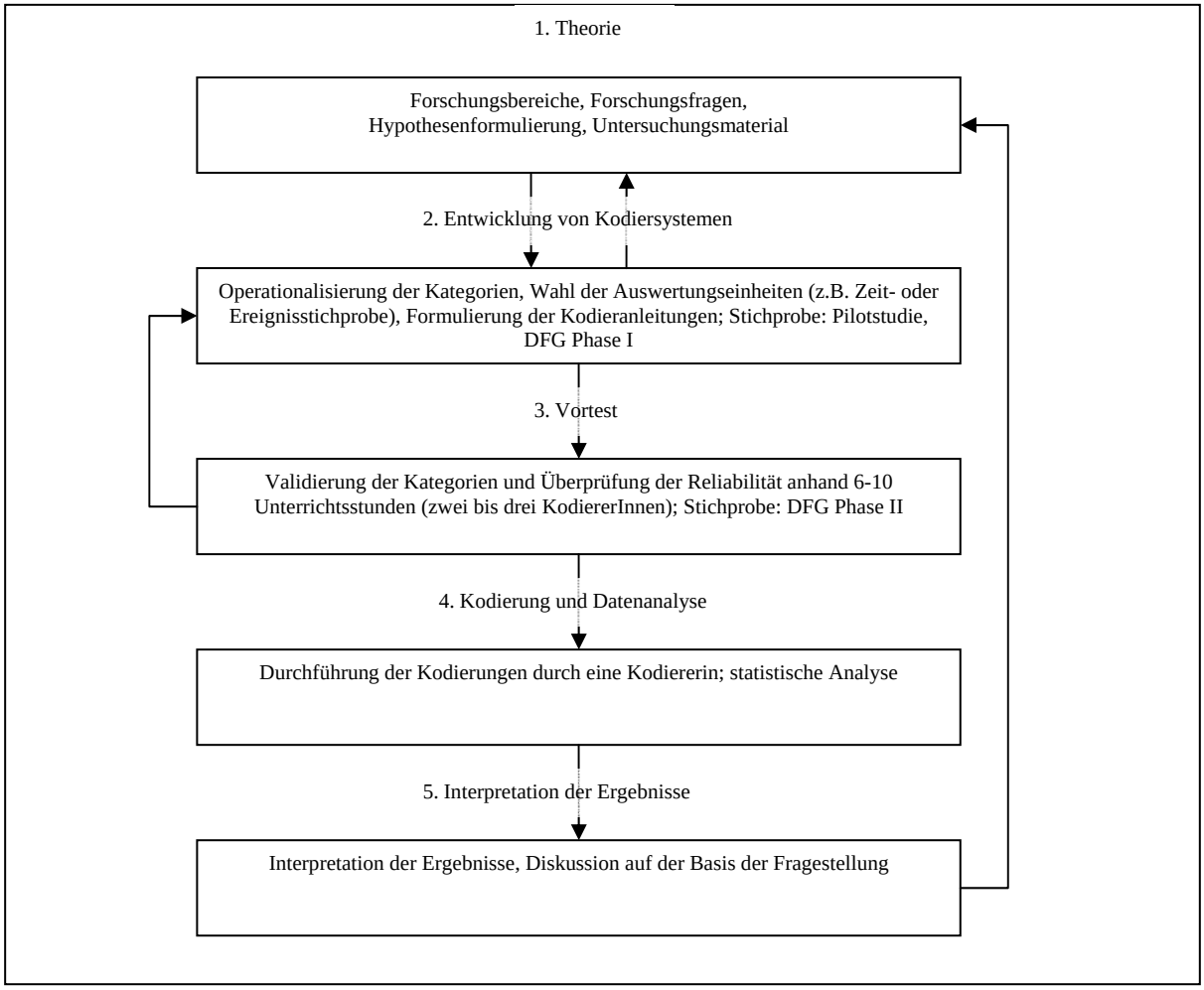

Abbildung 10-1: Vorgehen bei der Entwicklung von Kategoriensystemen (Seidel 2003, S. 104)

Das Schema verdeutlicht die mit Videostudien verbundenen Herausforderungen und verweist auch auf diesbezügliche Einschränkungen. Nicht alle wünschenswerter Weise zu untersuchenden Unterrichtsbereiche lassen sich verhaltensnah über ausgewählte Indikatoren niedrig-inferent erfassen. In besonderer Weise gilt dies für Lern- und Denkprozesse, in die Videostudien über die Sicht- oder Handlungsebene nur eingeschränkte und mittelbare Einblicke gewähren. Ein unmittelbarer Zugriff ist damit nicht gegeben; videobasierte Untersuchungen liefern lediglich Daten über die aufgezeichnete Unterrichtsrealität, womit mögliche Rückschlüsse auf vorangegangene Planungsprozesse oder auf Planungs- und Durchführungsdifferenzen ausgeschlossen bleiben. Es stellt sich die Frage, was überhaupt sinnvoll, d.h. intersubjektiv überprüfbar, beobachtbar ist.
Die eigene Studie lehnt sich an das von Seidel (2003b) vorgelegte Entwicklungsschema an. Dabei konnten zahlreiche Kodierungsversuche der schemagestützten Überprüfung nicht standhalten, wurden verworfen oder erfuhren eine Überarbeitung. Es kam zu Ausschlussverfahren, wenn es nicht gelang, trennscharfe und umfassende Modellierungen vorzunehmen, wenn Aspekte nicht

beobachtbar erschienen oder Interraterreliabilitäten zwischen mehreren Kodierern zu gering ausfielen.
Anfänglich war beabsichtigt das Standardkonzept von Oser (vgl. Kap. 5.2) in Teilbereichen in ein Kategoriensystem zu überführen. Dies gelang allerdings nur unzureichend, weil die Standards auf einem pädagogisch-psychologischen Abstraktionsniveau formuliert sind, das eine Erfassung über niedrig-inferente Kategorien nicht oder nur sehr verkürzend erlaubt. Beispiele für die nicht praktikable Überführung der Oserschen Standards in eine Beobachtungsmethodik liefern die Standardgruppen 1 „Lehrer-Schüler-Beziehungen und fördernde Rückmeldung“ bzw. 2 „Diagnose und Schüler unterstützendes Handeln“ (Oser 2001, S. 230ff.). Nicht ohne Zufall stehen diese Standardgruppen am Anfang seines Konzeptes. Immerhin formulieren sie zentrale Anforderungen an das Lehrerhandeln in der Interaktion mit Schülern. Die von Oser entwickelten Items z.B. „zu verhindern, dass Schülerinnen und Schüler wiederholte Erfahrungen machen, die zu gelernter Hilflosigkeit führen“ (Standardgruppe 1) oder „zu diagnostizieren, welche Ursachen Misserfolg, Aggression, Ängste, Blockierungen etc. haben und darauf angemessen zu reagieren“ (Standardgruppe 2) beschreiben jedoch Kompetenzprofile, die einer Beobachtung nur schwer zugänglich sind und kaum durch niedrig-inferente Kategorien erfasst werden können. Zudem wäre das Filmen von Probanden über einen längeren und kontinuierlichen Zeitraum Voraussetzung, um entsprechend anspruchsvolle Kompetenzprofile dokumentieren zu können. Aus Ressourcengründen kam für die vorliegende Studie ein solcher Zugang nicht in Frage. Versuche diagnostische Kompetenzen der Probanden niedrig-inferent zu untersuchen, wurden insgesamt verworfen. Ansätze von Oser finden hingegen in den hoch-inferenten Beobachtungsverfahren Berücksichtigung.
Auch andere Beobachtungsdimensionen konnten nicht realisiert werden. Das betrifft etwa den für sich genommen höchst interessanten Frageaspekt, inwieweit Studierende des Lehramtes Störungen von Schülern überhaupt wahrnehmen und wenn ja, mit welchen disziplinarischen Mitteln sie darauf reagieren. Es ließ sich kein geteiltes Verständnis über Störungen bzw. darüber schaffen, wie zweifelsfrei jeweils festzustellen ist, ob eine Störung nicht wahrgenommen, bewusst ignoriert wird oder in den Augen des Probanden gar keine Störung darstellt. Erschwerend kommt hinzu, dass die Stichprobe sich aus Videoaufnahmen in verschiedenen Schulformen, Grund- und Hauptschulen, zusammensetzt. Auch hier konnte nicht von einem geteilten Verständnis darüber ausgegangen werden, was einerseits als Störung gilt und was andererseits ein geduldetes Verhalten darstellt, das dem psychologischen Entwicklungsstand von Kindern geschuldet ist. Am Beispiel der Unterrichtsstörungen zeigt sich, dass niedrig-inferente Beobachtungsverfahren dann an Grenzen stoßen, wenn eine Beschreibung der Unterrichtsgestaltung über die Erfassung konkreter Ereignisse oder Handlungs- und Organisationsformen nicht mehr präzise möglich ist, sondern in eine interpretative Rekonstruktion der Wahrnehmung von Protagonisten mündet. Im letzten Fall ist der niedrig-inferente Beobachtungsrahmen, der sich auf einen geringen Grad abverlangter Schlussfolgerungen beruft, gesprengt. Hoch-inferente Beobachtungen rücken in das Kalkül, die neben niedrig-inferenten Verfahren Berücksichtigung in der eigenen Untersuchungsanlage fanden (vgl. Kap. 10.5).

Insgesamt kann als Ergebnis der Studie festgehalten werden, dass niedrig-inferente Beobachtungsverfahren, festgemacht an der Formulierung geeigneter Indikatoren oder der Berechnung von Übereinstimmungsmaßen, in ihrer Funktion überstrapaziert werden, wenn interaktives Lehrer- und Schülerhandeln im Beobachtungsfokus steht. Grundsätzlich erscheint es dem niedrig-inferenten Beobachtungsmodus eher zu entsprechen, eine Konzentration auf Handlungen einzelner Personengruppen vorzunehmen, Lehrer- und Schülerverhalten also getrennt voneinander zu beobachten, um die Daten in einem zweiten Schritt aufeinander zu beziehen. Die von Anfang an vollzogene Integration von Handlungselementen verschiedener Personengruppen innerhalb niedrig-inferenter Beobachtungsverfahren erschwert hingegen eine präzise Zuordnung zu Kategorien, die dann nicht mehr unbedingt niedrig-inferenten Charakter aufweisen. Oftmals lässt es sich jedoch nicht völlig vermeiden, zugleich die Lehrperson wie auch die Schüler in den Blick zu nehmen. Dies betrifft etwa die Beobachtung von Unterrichtskommunikation oder Arbeitsformen.
In der eigenen Studie wurden zwei Hauptdimensionen als niedrig-inferente Zugänge entwickelt:[86]

- Beobachtung von Sichtstrukturen
- Lernorganisation während des Klassenunterrichts

Die Sichtstruktur umfasst die sichtbaren und den Unterricht konstituierenden Lehr- und Lernhandlungen in Form von Sozialformen, Aktivitäten, Methoden und Medien und Kommunikation, die von Oser & Patry (1990) bzw. Oser & Baeriswyl (2001) treffend mit dem Begriff der „Choreographien“ belegt worden sind. Solche Choreographien stellen, wie oben angedeutet, basale Unterrichtsstrukturen zur Ermöglichung von Lernprozessen dar. Damit wird im Wesentlichen nur der lehrerrelevante Anteil der Basismodell-Theorie von Oser & Patry (1994) in einem niedrig-inferenten methodischen Zugang übernommen. Die Theorie lässt Hypothesen darüber zu, wie sich prototypisch Lehr-Lernprozesse aus Schülersicht darstellen bzw. wie professionelle Handlungsschemata zu entwickeln sind, die Lernprozesse auf der Schülerseite durch Lehrerhandeln optimal unterstützen (Oser & Sarasin 1995).
Auch wenn durch den Unterricht angeregte Lernprozesse auf der Schülerseite eine äußerst interessante Thematik abgeben, fokussiert die eigene Studie jedoch die Entwicklung von unterrichtlichen Kompetenzen von Studierenden des Lehramtes in verlängerten Praxisphasen. Damit stehen die von lehrender Seite intendierten Handlungssequenzen zur Anregung von Lernprozessen im Mittelpunkt, weniger die Schüleraktivitäten (Gerber, Knierim & Labudde 2006). Neben dieser sich aus der Fragerichtung ergebenden Einschränkung lassen sich Lernprozesse von Schülern m.E. zudem nur bedingt über niedrig-inferente Kategorien beobachten, weil letztlich nur auf der Handlungsebene indikatorisch mittelbar auf evtl. dahinter liegende

[86] Die hier vorgestellten niedrig- und hoch-inferenten Beobachtungsverfahren wurden im von der Internationalen Bodensee-Hochschule geförderten Projekt „Standarderreichung in der Lehrerinnen- und Lehrerbildung“ entwickelt, wobei innerhalb des Projektes die Federführung für die Entwicklung der Beobachtungsverfahren bei der Pädagogischen Hochschule Weingarten bzw. personell bei Prof. Dr. Günter Dörr und mir lag (vgl. Baer et al. 2007; Kocher & Wyss 2006).

Lernprozesse geschlossen werden kann. U.U. gewinnt man auf diese Weise weniger Aufschluss über die Qualität der eigentlichen Lernprozesse oder den Grad der kognitiven Aktivierung der Schüler als dass man emotional-affektive Gemütslagen wie Motivation und Konzentration oder soziale Austauschprozesse als wichtige Lernausgangslagen erfasst. Ein weiterer Grund für den Verzicht auf die Untersuchung von schülerbezogenen Lernprozessen über niedrig-inferente methodische Zugänge liegt in dem gewählten Stichprobenplan. Grundsätzlich gibt es zwei verschiedene Stichprobenpläne:

- Zeitstichprobe („timesampling“)
- Ereignisstichprobe („eventsampling“)

Zeitstichproben verlangen, die gesamte Unterrichtsstunde in einem vorher festgelegten Intervallturnus (z.B. zehn Sekunden oder eine Minute) zu kodieren und so jedem Intervall eine Kodierung zuzuweisen. Im Falle eines Zehn-Sekunden-Turnus erhält man für 45 Minuten 270 Kodierungen. Wählt man dagegen als Stichprobenplan die sog. Ereignisstichprobe, ist die Aufmerksamkeit lediglich auf vorher festgelegte, spezifische und zu kodierende Unterrichtsereignisse gerichtet. Voraussetzend ist es dafür erforderlich die gesamte Unterrichtsstunde in einer binomischen Kodierung als Zeitstichprobe zu behandeln, um festzulegen, wann das interessierende Ereignis (z.B. Durchführung von Experimenten oder Gruppenarbeitsphase) stattfindet oder nicht. Der zweite Kodierungsdurchgang widmet sich, wiederum unter Bestimmung festgelegter Zeitintervalle, den ausgewählten Unterrichtsereignissen.
Im eigenen Vorgehen wurde das Videomaterial zunächst danach untersucht, ob Unterricht stattfindet oder nicht. Sequenzen vor dem eigentlichen Unterrichtsbeginn oder nach Unterrichtsende fanden keine Berücksichtigung. Die damit in die Ereignisstichprobe eingehenden Unterrichtssequenzen wurden weiterhin in einem Timesampling-Verfahren kodiert, wobei für alle Kategorien ein Zehn-Sekunden-Intervall festgelegt wurde. Da die Untersuchung explorativ angelegt ist, konnten im Vorhinein keine Hypothesen über besonders interessierende Unterrichtsphasen oder Ereignisse aufgestellt werden, die neben der grundsätzlichen Einteilung, inwieweit überhaupt Unterricht stattfindet (eine notwendige Unterscheidung, die sich aus den Aufnahmebedingungen ergab), weitere Ereignisstichproben gerechtfertigt hätten.
In der Hauptdimension unterrichtlicher Sichtstrukturen fanden über die Klärung des eigentlichen Unterrichtsstatus hinaus folgende Teildimensionen Berücksichtigung:

- Unterrichtliche Arbeitsformen
- Allgemeindidaktische Unterrichtsphasen
- Aktivitäten im Klassenunterricht
- Kommunikation im Unterricht

Die Dimensionen sind eng an die Entwicklungen von Seidel (2003b; vgl. auch Kocher & Wyss 2006) angelehnt, wurden aber sowohl auf der Ebene der Kategorien als auch auf der Ebene der Indikatoren unterschiedlich stark modifiziert (vgl. Anhang). „Unterrichtliche Arbeitsformen“ zielen auf die Differenzierung von Sozialformen (Plenum, Stillarbeit, Partnerarbeit etc.). Als wesentliches Kompetenzmerkmal sollte

aber in dieser Kategorisierung beurteilt werden, inwieweit eingesetzte Sozialformen von der Lehrperson in ein didaktisches Gesamtkonzept der Unterrichtsstunde integriert sind oder ob Störungen von Schülerarbeitsphasen durch die Lehrperson darauf schließen lassen, dass Planungen offenbar nicht in Gänze eine Umsetzung erfahren bzw. die Lehrperson zumindest diesbezügliche Einschätzungen durch störende Unterbrechungen signalisiert. Die Fähigkeit zu erkennen, wann Schüler unabhängig von Planungselementen und Beratungsleistungen durch die Lehrperson für sich ein Arbeitssetting gefunden haben, in dem die Lehrperson eigentlich nur mehr als Störungsparameter auftauchen kann, nimmt hier hohen Stellenwert ein. Diese Gewichtung entspricht einem pragmatisch-konstruktivistischen Lernverständnis als pädagogischem Leitbild gemäß einer Ermöglichungsdidaktik (Wellenreuther 2004, S. 65ff.; Arnold 2007; Reusser 2006). Bei entsprechend störendem Verhalten der Lehrperson liegt ein Indikator für ein mangelndes Augenmerk auf die Schülerinteressen vor. Unabhängig von Arbeitsformen der Schüler betonen Lehrpersonen in diesem Fall ihre Lehrerrolle über und setzen ihr Planungskonzept ggf. auch gegen Schülerinitiativen und -interessen durch. Die Wahrnehmung dessen, wann man als Lehrperson Rückzugsmöglichkeiten hat und es tendenziell gelungen ist, überflüssig zu werden und man also durch zusätzliche Interventionen schülerbezogene Lernprozesse eher stört als fördert, stellt ein wichtiges Zielelement innerhalb einer konstruktivistisch orientierten Lehrerbildung dar. Sie kann vermutlich nur in zeitlich ausgedehnten Praxislernsituationen erfahrbar gemacht werden.

Das Konstrukt „Allgemeindidaktische Unterrichtsphasen“ gilt der Identifizierung von methodischen Verlaufsformen (Meyer 2001, S. 47) auf einer Meso-Ebene, die sich unabhängig von der Klassenstufe, dem Fach oder der Schulform ausfindig machen lassen. Bei gelingendem Nachweis ist das Plausibilitätskriterium erbracht, dass Unterrichtsaufzeichnungen über die zu unterscheidenden Rahmenbedingungen hinweg als schlüssige und gemeinsame Stichprobe angesehen werden können. Auch Abweichungen sprechen für die Möglichkeit einer Vergleichbarkeit, weil sich auf einem relativ hohen didaktischen Abstraktionsniveau Gemeinsamkeiten und Unterschiede herausstellen lassen. Die Kategorie impliziert zudem die Hintergrundannahme, dass sich in der Planung und Durchführung unterrichtlicher methodischer Verlaufsformen allgemeinpädagogische Kompetenzen ausdrücken, die jenseits individueller fachwissenschaftlicher und fachdidaktischer Expertise angesiedelt sind und damit als Beleg dafür gewertet werden könnten, dass gerade derart erworbene pädagogische Kompetenzen hohe Bedeutung für die Choreographie des Unterrichts einnehmen (Oser & Patry 1990; Oser & Baeriswyl 2001). Videostudien bieten in der Untersuchung von Sichtstrukturen des Unterrichts einen Analysezugang, der durch andere methodische Verfahren nicht geleistet werden kann. Sie bilden, mit Einschränkungen, die für Schüler vorfindbaren didaktischen Settings von Unterricht ab. Vor einem ermöglichungsdidaktischen Bezugshorizont sind nun Interpretationsräume hinsichtlich damit angelegter Lehr-Lernpotentiale eröffnet.

„Aktivitäten im Klassenunterricht“ beschreiben die Arbeit mit fachlichen Inhalten, die in der Herstellung der Klassenöffentlichkeit thematisiert werden. Die von Lehrpersonen angestoßenen Aktivitäten spiegeln das unterrichtliche Handlungsrepertoire wider und können vorsichtig in Richtung einer mehr oder

weniger beherrschten Methodenvielfalt gedeutet werden. Dabei stellt Methodenvielfalt keinen Selbstzweck als allgemeine Motivierungsstrategie dar, sondern ist im Sinne professioneller pädagogischer Handlungskompetenz auf didaktische Kontextuierung angewiesen. Der Einsatz verschiedener Methoden und Medien hängt davon ab, wann sie Unterrichtszielen sinnvoll zugute kommen und welche Schülergruppen jeweils davon profitieren (Helmke 2003, S. 65; Brophy 2000; Wiechmann 2000). Die Frage des angemessenen Methoden- und Medieneinsatzes kann allerdings durch rein niedrig-inferente Zugänge nicht beantwortet werden, sondern findet in hoch-inferenten Ratingverfahren Berücksichtigung. Dabei wird davon ausgegangen, dass Variationen im Methoden- und Medieneinsatz prinzipiell einer Adaption an unterschiedliche Lernstile und Rezeptionsfähigkeiten der Schüler entsprechen, allerdings eine hohe Variation nicht automatisch positiv zu bewerten ist („Methodenflucht“). Eine Festlegung im eigenen Manual lautete, dass Aktivitäten, die der Klassenöffentlichkeit bedürfen, nicht kodiert werden, wenn Einzel-, Partner- oder Gruppenarbeitsphasen stattfinden. Im Kern liefert die Dimension der Aktivitäten im Klassenunterricht Aussagen auf einer Mesoebene über die Bedeutsamkeit von Aktivitätsmerkmalen für den Unterricht auf Klassenebene.
Auf der Sichtebene galt es schließlich zu untersuchen, wie Gesprächsanteile in der „Unterrichtskommunikation“ zwischen Lehrpersonen und Schülern verteilt sind. Auf die problematisch vorzunehmende Interpretation der Ergebnisse ist oben bereits hingewiesen worden. In diesem Zusammenhang besteht ein verobjektivierender methodischer Versuch in der im Manual festgelegten Regel, dass bei Schülerarbeitsphasen als Kommunikationsform in jedem Fall die Kategorie „Mischformen“ kodiert wird, weil angesichts der nicht an die Klassenöffentlichkeit gerichteten Kommunikationsformen häufig nicht zweifelsfrei entschieden werden kann, wer den größeren Gesprächsanteil inne hat.

Die zweite zentrale Kategoriendimension dient der niedrig-inferenten Untersuchung der Lernorganisation im Klassenunterricht (vgl. Seidel 2003b, S. 101). Im Unterschied zu den Sichtstrukturen gerät das unmittelbar auf Schülerverhalten bezogene Lehrerhandeln in den Blick, das entsprechende positive oder negative Auswirkungen auf Lernaktivitäten der Schüler entfaltet. Konkret interessieren im Rahmen der Dimension Kategorien zu folgenden Bereichen:

- Strukturierung
- Differenzierung

Um Lernprozesse zu optimieren ist eine gute Strukturierung des Unterrichts und des Lehrervortrags notwendig (Helmke 2003, S. 61). Wellenreuther (2004, S. 339f.) hebt die Bedeutung von Strukturierungsleistungen für den Aufbau kognitiver Schemata im Rahmen der direkten Instruktion hervor. Strukturierungsleistungen durch die Lehrperson stellen einen wichtigen Bereich von Vermittlungsstrategien dar, die es Schülern erleichtern, einen Überblick über avisierte Unterrichtsinformationen zu erhalten und gleichzeitig über metakognitive Adressierung Lernprozesse zu effektivieren (Kiel 2005, S. 331; Bauer, Kopka & Brindt 1999). Dazu gehören die Aktivierung des Vorwissens, die darauf bezogene übersichtliche Gliederung des Unterrichtsangebotes, die Einbettung der Stunde in den Ablauf weiterer Stunden,

präzise und verständliche Arbeitsaufträge, Zusammenfassungen und Abgrenzungen der Unterrichtsinhalte sowie Rückmeldungen über Qualität und Quantität geleisteter Lernprozesse. In dieser Lesart verstehen sich Strukturierungselemente als metakognitive Orientierungen mit förderlicher Auswirkung auf schülerbezogene Lernprozesse. Voraussetzung sind profunde Kenntnisse der Lehrpersonen über die menschliche Informationsverarbeitung und Gedächtnisabspeicherung. Zusätzlich bzw. zugrunde liegend ist eine klare Wissensstrukturierung zwecks Bewahrung des thematischen Überblicks erforderlich. Hiermit werden hohe Ansprüche an Lehrpersonen formuliert bzw. Wissensbasen vorausgesetzt, bei denen man nicht unbedingt voraussetzen kann, dass die Studierenden im Praxisjahr diesen gerecht werden. Die Entwicklung der Kategorie basiert jedoch auf der Hintergrundannahme, dass sich evtl. gerade für diesen Zeitraum Veränderungen in unterrichtlichen Kompetenzen längsschnittlich aufzeigen lassen, die in der hier vorgestellten Modellierung eine äußerst bedeutsame Komponente der Klassenführung darstellen (Kiel 2005, S. 331; Wahl, Weinert & Huber 2001, S. 271ff.).

Ähnlich verhält es sich mit den Kategorien zu der Dimension „Differenzierung“. Der Umgang mit heterogenen Lerngruppen stellt eine zentrale Herausforderung für das unterrichtliche Handeln von Lehrpersonen dar. Der Grad der Heterogenität in Klassenzusammensetzungen ist in den letzten Jahrzehnten – bezogen auf Vorwissensbereiche, Motivation, Interessen und soziale Herkunft – gestiegen. Diese Entwicklung ist in großen Anteilen gesellschaftlichen Dynamisierungsprozessen mit Auswirkungen auf die schulische Realität geschuldet. Nutzungsmöglichkeiten der neuen Medien (Computer, Internet, Mobiltelefon) tragen ebenfalls zur Heterogenität in den Klassenzimmern bei.

Mögliche Antworten auf diese Problemlage liegen darin, Unterricht über Differenzierungsmaßnahmen zu individualisieren und an die Bedürfnisse und Interessen von Schülern zu adaptieren (Helmke 2003, S. 72; Wellenreuther 2004, S. 161ff.; Bischoff, Brühwiler & Baer 2005). Mittlerweile besteht ein erziehungswissenschaftlich breiter Konsens über notwendige Differenzierungsmaßnahmen als Bedingung für gelingende Lehr-Lernprozesse, die allerdings nicht unbedingt in jeder Stunde realisiert werden können. Der in der Vergangenheit und teilweise noch heute dominierende Frontalunterricht als Hauptunterrichtsform gilt als überholtes Konzept (Helmke 2003, S. 73). In diesem Zusammenhang konstatiert Gudjons (2006), dass vom Frontalunterricht abweichende Unterrichtsformen (z.B. Gruppenunterricht) nach wie vor ein Schattendasein führen, obwohl theoretische Vorteile selbstgesteuerter Unterrichtsphasen überzeugen. Gründe für die Favorisierung lehrergesteuerter Unterrichtsformen mögen in der Anschauung liegen, den Unterricht „im Griff zu haben“ und damit die Rolle als Lehrperson zu legitimieren. Gruppenarbeitsformen sind bei durchdachter Planung in der Unterrichtsvorbereitung oftmals aufwendiger und bergen Potentiale der Verunsicherung in sich, sowohl im Hinblick auf damit verbundene Unruhen als auch auf den unterrichtlichen Zeitaufwand. Besonders bei fehlender Übung der Schüler im Umgang mit diesen Arbeitsformen kann der Eindruck von Zeitverschwendung befördert werden.

Die in den Kategorien zur Differenzierung modellierten Konstrukte sollen untersuchen, wie Lehrpersonen Arbeitsphasen anlegen und dabei Momente der Differenzierung berücksichtigen. Ausgangsannahme war die These, dass eine

Zunahme von Differenzierungselementen für eine positive Entwicklung von unterrichtlichen Kompetenzen steht.

Zusammenfassend versprechen die im Rahmen der Lernorganisations-Dimension konstruierten Kategorien zur Differenzierung und Strukturierung Aufschluss darüber, wie Lehrpersonen in Phasen direkter Instruktion – und als Indikator für Schülerorientierung auch in lehrergestützten Unterrichtsphasen – metakognitive Adressierungen berücksichtigen bzw. inwieweit sie in schülerorientierten Arbeitsphasen Differenzierungselemente einbringen (und damit nicht auch in Schülerarbeitsphasen dem Hauptvorwurf des Frontalunterrichts entsprechen, gleiche Aufgabenformen für alle Schüler zu favorisieren).

10.5 Entwicklung hoch-inferenter Beobachtungsverfahren

Während sich niedrig-inferente Kategoriensysteme auf die strukturelle Beschreibung von Unterrichtaktivitäten beschränken und damit kurze, vergleichsweise häufig auftretende Verhaltensäußerungen fokussieren, die über den gesamten Stundenverlauf aggregiert werden, können hoch-inferente Schätzverfahren eingesetzt werden, um Qualitätseinschätzungen von Unterrichtsmerkmalen vorzunehmen (Clausen, Reusser & Klieme 2003, S. 124). Auch niedrig-inferente Kodierungen enthalten zum Teil (je nach Anlage) Hinweise auf Dimensionen von Unterrichtsqualität. Diese Hinweise sind aber über die gesamte Unterrichtsstunde verstreut und werden mit dem methodischen Zugang, wegen der Beschränkung auf eindeutig beobachtbare operationalisierte Konstrukte, nur unzureichend erfasst.
Die oben dargestellte Theoriebedürftigkeit für die Interpretation der Daten unterliegt bei hoch-inferenten Zugängen gewissermaßen einer Umkehrung. Hoch-inferente Verfahren setzen voraus, dass der zugrunde gelegte Theoriebezug in den „Köpfen“ der Beobachter verankert ist. Damit fällt der Expertise der Beobachter, die über profunde Kenntnisse bezüglich der Gestaltung von Lehr-Lernprozessen verfügen müssen (Petko et al. 2003, S. 276), eine zentrale Rolle zu. Auf der Basis dieser Expertise stellen die Beobachter Interpretationen an, die über die konkret beobachtbaren Unterrichtsereignisse hinausgehen. Damit werden die Einschätzungen potentiell anfälliger für systematische und unsystematische Beurteilungsfehler (Clausen, Reusser & Klieme, S. 127). Unsystematische Fehler liegen dann vor, wenn die Beobachter mit der Komplexität der zu beobachtenden Ereignisse überfordert sind und spezifische Unterrichtsmerkmale nicht wahrnehmen bzw. zum Zeitpunkt der Einschätzung am Ende der Stunde nicht mehr alle entscheidenden Situationen vor Augen haben. Zudem weisen Lehrpersonen nicht immer konsistente Verhaltenweisen im Hinblick auf die zu untersuchenden Unterrichtsmerkmale auf, was zu der Schwierigkeit führt, inkonsistente Verhaltensweisen in ihrer einzuschätzenden Bedeutung gegeneinander abzuwägen.
Systematische Beobachtungsfehler liegen dann vor, wenn sich die Unterschiede zwischen einzelnen Beobachtern auf bestimmte unterrichtliche Topoi beziehen oder auch biographische Einflüsse in der Einschätzung sichtbar werden (vgl. Kap. 10.2).
Demgemäß sind hoch-inferente Einschätzungen in einem größeren Maße fehlersensitiv als niedrig-inferente Kodierungen. Vor diesem Hintergrund (und auch

vor dem Hintergrund eigener Erfahrungen) überrascht es, dass Petko et al. (2003, S. 276) zu der Einschätzung gelangen, die Schulung von Beobachtern, die einen Expertenstatus für sich beanspruchen können, würde weniger aufwendig als die Schulung von Beobachtern für niedrig-inferente Verfahren ausfallen. Weiterhin führen die Autoren aus, dass der Rekrutierung entsprechender Beobachter eine wichtige Rolle zufällt, wobei die hoch-inferenten Dimensionen intuitiv erschlossen werden können (ebd.). Genau diese intuitiv erschließbare Einschätzung von Qualitätsmerkmalen des Unterrichts wird hier aber angezweifelt. Auf einer individuellen Ebene mag sie weiterführend sein. Spätestens aber, wenn es in diskursiven Prozessen zum Abgleich individueller Einschätzungen mit den Urteilen anderer Beobachter kommt und dabei gewichtige Unterschiede festgestellt werden, gelangen intuitive Interpretationsprozesse an ihre Grenzen und müssen durch kriteriumsorientierte Verfahren ersetzt werden, die einen interindividuellen Diskurs erst erlauben.[87]
Wie bereits oben dargestellt ist eine operationalisierte Definition der Kriterien, die hoch-inferente Verfahren fundieren, nur in Ansätzen möglich, weil sie ansonsten ihren eigentlichen Anspruch und damit ihre Bedeutung einbüßen. Auf einer methodischen Ebene geht es also darum, Manuale zu entwickeln, die einerseits beschreibend darstellen, welches Verständnis von Unterrichtsqualität für die Einschätzung zugrunde gelegt wird, und andererseits noch genügend Freiheitsgrade offen lassen, um Beobachtern für ihre Analyse von Lehr-Lernprozessen in unterrichtlichen Handlungsfeldern ausreichend Gelegenheit zu geben von ihrer Expertise auch Gebrauch zu machen. Dabei muss zwischen der Trainingsphase, in der die Kriterien explizit vor Augen geführt werden, um individuelle systematische Beobachtungsfehler zu verringern und Übereinstimmungen zu erhöhen, und der eigentlichen Ratingphase unterschieden werden, in der das nicht mehr notwendig ist (so denn eine hinreichende Übereinstimmung festgestellt werden konnte). Wichtig für die Anlage der Trainingsphase ist zu definieren, welches Übereinstimmungs- oder Reliabilitätsmaß für die Überprüfung zugrunde gelegt wird bzw. welche Werte zur Weiterverarbeitung der Daten angesetzt werden.

Hoch-inferente Einschätzungen beziehen sich nicht zwingend auf die gesamte Unterrichtsstunde, sondern es können auch kleinere Analyseeinheiten bestimmt werden (Waldis et al. 2006, S. 177). Damit stellt sich die Frage, an Hand welchen Bestimmungsverfahrens sich die kleineren Analyseeinheiten ergeben. Eine nahe liegende Möglichkeit besteht in der Verwendung der Daten von niedrig-inferenten Kodierungen, die bereits im Vorfeld klar abgrenzbare Unterrichtsereignisse (z.B. methodische Verlaufsformen) kenntlich gemacht haben. Diese Vorgehensweise eignet sich besonders dann, wenn eine eingegrenzte Fragestellung, z.B. im fachdidaktischen Bereich, untersucht werden soll. Damit einhergehend sind grundsätzliche Fragen der Standardisierung zu klären, die vor allem das Fach oder das Unterrichtsthema betreffen, um eine höhere Vergleichbarkeit der

[87] Clausen, Reusser & Klieme (2003, S. 130f.) weisen in ihrer Überprüfung der Qualität von Beurteilerergebnissen z.T. niedrige Werte in Form von Generalisierbarkeitskoeffizienten aus (z.B. für die Dimensionen Klarheit oder diagnostische Kompetenz im Sozial- bzw. Leistungsbereich). Daraus kann man schlussfolgern, dass es auch in dieser einflussreichen Studie für Teildimensionen nicht gelang, intersubjektive Abweichungen auszuschließen, obwohl ein hoher methodischer Aufwand betrieben wurde.

aufgezeichneten Stunden zu erreichen (Seidel 2003b). Die so intendierte Integration von niedrig- und hoch-inferenten Fragestellungen erscheint schlüssig, konnte aber in der eigenen Studie nicht angewandt werden, weil es um eine übergreifende und längsschnittliche Untersuchung der Entwicklung von unterrichtlichen Kompetenzen von Studierenden in verlängerten Praxisphasen geht. Aus organisatorischen Gründen war eine Standardisierung der Schulformen, Klassenstufen, Fächer oder Unterrichtsthemen nicht möglich. Eine Studie, die sich nur auf bestimmte Unterrichtsereignisse bezieht, wurde ausgeschlossen, um mögliche Verzerrungseffekte zu vermeiden, die sich aus der fehlenden Standardisierung ergeben. Die hoch-inferenten Einschätzungen bezogen sich deshalb auf die gesamte Unterrichtsstunde.
Insgesamt lässt sich nicht abschließend festhalten, welche Analyseeinheit die beste Grundlage für hoch-inferente Einschätzungen bietet, sondern der Bezugsrahmen ergibt sich jeweils individuell in Abhängigkeit der Fragestellung. Aus meiner Sicht werden hoch-inferente Einschätzungen durch die Wahl kleiner zeitlicher Analyseeinheiten aber eher erschwert, weil das interessierende Verhalten in einem einzelnen Intervall unter Umständen nicht beobachtet werden kann. Trendel, Wackermann & Fischer (2007) legen allerdings eine interessante Studie vor, in der sie die Basismodelltheorie (Oser, Dick & Patry 1992; Oser & Baeriswyl 2001) als Folie für die Unterrichtsanalyse anwenden und damit versuchen die Tiefenstrukturen des Unterrichts im Minutentakt zu klären.

Die Ergebnisse hoch-inferenter Unterrichtsbeurteilungen liegen zumeist in Form von Skalenwerten für die untersuchten Konstrukte vor, womit Mittelwerte für die Items berechnet werden können. Diese Mittelwerte lassen sich auf Skalenniveau aggregieren und können auch über die Probanden zusammengefasst werden, um längsschnittliche Betrachtungen zu ermöglichen. Ratingskalen werden als ein Messinstrument zur Transformation von qualitativen Sachverhalten in quantifizierbare Größen verstanden.[88] In Anlehnung an Clausen, Reusser & Klieme (2003) wurde in einer Vorstudie (Kocher & Wyss 2006; Baer et al. 2007) das vierstufige Antwortformat übernommen. Die Skalenbezeichnungen reichten in dieser Studie von 1 („trifft überhaupt nicht zu“) bis 4 („trifft voll und ganz zu“) und sind bipolar angelegt, d.h. pro Pol stehen zwei Items zur Verfügung, um eine Zustimmung oder Ablehnung in der Beobachtung zu bezeichnen und eine Differenzierung

[88] Besonders bei Videostudien, die einen hoch-inferenten methodischen Zugang verwenden, wird die vieldiskutierte Dichotomie zwischen quantitativen und qualitativen Methoden durchlässig. Nehmen die interpretativen Prozesse auf der Beobachterseite mit zunehmender Inferenz zu, so kann man die These aufstellen, dass solche Verfahren dem interpretativen Paradigma (Bohnsack 2000) zuzuordnen sind und damit als qualitative Forschungsmethodik angesehen werden müssen. Diese Entscheidung hängt davon ab, ob man den Prozess der Datengewinnung oder den Prozess der Datenauswertung höher gewichtet und dabei die Rolle des Raters mit einbezieht. Bei Videostudien fallen Datengewinnung und Datenauswertung häufig in einer Person zusammen, was auf die prinzipielle Notwendigkeit verweist, für die Datengewinnung mindestens eine weitere Person mit einzubeziehen und die Ergebnisse hinsichtlich von Reliabilitätsmaßen zu vergleichen. Die Interpretation der Daten sollte ebenfalls in diskursiven Prozessen innerhalb einer Forschergruppe validiert werden. Unter diesen Voraussetzungen vertrete ich die Auffassung, dass auch hoch-inferente Verfahren einer quantitativen Methodik entsprechen. Die Frage einer Repräsentativität der Daten bleibt bei Videostudien, im Hinblick auf die oftmals geringen Fallzahlen, allerdings ungeklärt.

auszudrücken. Die gradzahlige Skala verlangt vom Beobachter sich für einen Pol zu entscheiden.[89]
Als Ergebnis der Vorstudie wird in der hier vorliegenden Untersuchung vom vierstufigen Antwortformat abgewichen. Vergegenwärtig man die kognitiven bzw. interpretativen Prozesse des Raters in seiner Einschätzung, so lässt sich ein Kontinuum absteigender Abstraktion nachzeichnen. Am Anfang steht die Frage, auf welchem Pol die Einschätzung zu verorten ist, um in einem zweiten Schritt eine Differenzierung auf dem jeweiligen Pol vorzunehmen. Während die Zuordnung zu einem Pol aus eigener Erfahrung noch vergleichsweise einfach vorzunehmen ist, fällt eine Differenzierung bei nur zwei zur Verfügung stehenden Items mitunter recht schwer, womit Deckeneffekte nicht auszuschließen sind. Die Skala hält also zu wenig Differenzierungsmöglichkeiten bereit, was besonders für längsschnittliche Studien fatale Auswirkungen haben kann, weil sich Entwicklungen über die Zeit nicht darstellen lassen, wenn schon zu einem frühen Zeitpunkt ein zu hoher Skalenwert vergeben wurde.
Deckeneffekte, die sich aus einem zu hohen Kompetenzniveau bereits zum ersten Messpunkt ergeben, womit Steigerungen in der längsschnittlichen Entwicklung kaum noch realisierbar sind, wurden bei Studierenden, die sich noch am Anfang ihrer Ausbildung und ihrer berufspraktischen Entwicklung befinden, theoretisch ausgeschlossen. Diese Einschätzung hat einen Einfluss auf die Ergebnisinterpretation. Auf Grund der Vorerfahrungen wurde die Entscheidung getroffen, anstatt einer vierstufigen eine sechsstufige Skala zu verwenden. Der theoretische Nachteil einer Erhöhung der Varianz, die in der Überprüfung von Interraterreliabilitäten potentiell zu schlechteren Werten führt, soll durch den Vorteil einer differenzierteren und damit präziseren Einschätzungsmöglichkeit ausgeglichen werden.
Für die Analyse der Daten und ihre Aussagenreichweite ist die Definition des Skalenniveaus ausschlaggebend. Das Niveau beschreibt die Messbarkeitseigenschaften einer Skala und inwieweit Differenzierungen zwischen den Items vorgenommen werden. Bei den vorgestellten Ratingskalen handelt es sich per se um ein Ordinalskalenniveau, d.h. jede Merkmalsausprägung wird genau einer Kategorie zugeordnet, wobei die Quantifizierungen einer Reihenfolge entsprechen. Zusätzlich wird bei Ratingskalen postuliert, dass die semantischen Abstände zwischen den Polen gleich sind (Kriterium der Äquidistanz). Auch die quantifizierbaren Abstände zwischen den Rangplätzen werden als konstant betrachtet, weil man annimmt, dass die Rater in der Lage sind, gleiche Intervalle in ihren Einschätzungen zu berücksichtigen. Als Ergebnis dieser theoretischen Setzung erfüllen Ratingskalen die mathematischen Voraussetzungen von Intervallskalen (Bortz & Döring 2006, S. 180) und ermöglichen damit die entsprechenden statistischen Rechenverfahren (z.B. Berechnung des arithmetischen Mittels).
Die Beobachter sollen in ihrer intervallskalierten Wahrnehmung der Ratingskala dadurch unterstützt werden, dass nur die beiden Pole („trifft voll und ganz zu“ bzw. „trifft überhaupt nicht zu“) benannt werden, während die dazwischen liegenden Items kein gesondertes Antwortformat erhalten. Werden nur die Endpunkte einer

[89] Dieser „Entscheidungszwang“ wurde mit Absicht intendiert, um der bekannten Tendenz von Beobachtern, bei Entscheidungsschwierigkeiten einen neutralen Mittelwert zu wählen, entgegenzuwirken.

mehrstufigen Ratingskala benannt, womit die Extreme eines Kontinuums angedeutet sind, kann diese als approximativ intervallskaliert angenommen werden (Wirtz & Caspar 2002; Bortz & Döring 2006; Übersax 2001). Damit wird die These zugrunde gelegt, dass Beobachter im Einschätzungsprozess den einzelnen Items gleiche Intervalle zuordnen, weil sie die Gegensätzlichkeit der Pole erkennen. Eine gesonderte Kennung der Items könnte diesen konstruktiven Prozess unter Umständen auf Grund semantischer Besonderheiten in der Wahrnehmung nur stören oder auch die Beobachter überfordern (Bortz & Döring 2006, S. 180).
Neben der methodischen Diskussion, welches Skalenniveau bei Ratingskalen vorliegt, stellt sich die Frage, welche Aussagen diese Skalen liefern bzw. in welcher Richtung und Aussagereichweite die Daten interpretiert werden können. Setzt man sie zur Einschätzung von Qualitätsmerkmalen ein, so wird suggeriert, dass hoch-inferente methodische Zugänge unmittelbar Qualitätsmerkmale abbilden. Gleichzeitig handelt es sich aber nur um hoch-inferente Beobachtungsskalen, die zwar angesichts des zugrunde gelegten Theoriebezugs bedeutsam und insofern – bezogen auf die Fragestellung – valide erscheinen mögen, für sich genommen aber nicht als „Noten“ verstanden werden dürfen (auch wenn die Skalierung von „1“ bis „6“ dies nahe legen mag). Sie sind das Ergebnis theoriegeladener Konstrukte, die letztlich bloß beurteilen, inwieweit ein vorher definiertes Verhalten in welchem Umfang beobachtet werden kann. Damit liefern sie zwar – im Vergleich zu niedrig-inferenten Kodierungen – unmittelbarer zu interpretierende Ergebnisse, die aber nur im Rekurs auf die vorher operational definierten Konstrukte einen Aussagegehalt beanspruchen können.
Im Weiteren sollen die Kompetenzdimensionen, die für diese Studie entwickelt wurden, vorgestellt werden. Forschungstheoretisch wurde das Angebot-Nutzungs-Modell unterrichtlicher Wirkungen zugrunde gelegt (Helmke 2003, S. 42). Für die Entwicklung der Kompetenzdimensionen unterrichtlichen Handelns bedeutete dies den Versuch, Kompetenzfacetten zu finden, die Lerngelegenheiten schaffen oder ermöglichen, ohne von einer direkten Beeinflussung auf die Lernprozesse bei Schülern auszugehen. Die konsequente Berücksichtigung des Angebot-Nutzungs-Modells verlangt bei der operationalen Definition von hoch-inferenten Kategorien ein Unterrichtsverständnis indikatorisch zu explifizieren, das übergreifend als „Ermöglichungsdidaktik“ (Tietgens 2007) verstanden werden kann. Lehrpersonen schaffen durch ihr Handeln Lernumgebungen, die Lernprozesse auf Schülerseite ermöglichen, aber nur mittelbar beeinflussen.
Im theoretischen Teil der Arbeit wurde versucht der Kompetenzthematik in ihrer Komplexität gerecht zu werden und auf Problembereiche hinzuweisen, die sich aus der Verwendung des Begriffes als Beschreibungsfolie für unterrichtliche Handlungsdimensionen ergeben. Im Terminus der „Ermöglichungsdidaktik“ sowie im Rekurs auf das Angebot-Nutzungs-Modell werden die durch die Systemtheorie und vor allem das Konstruktivismustheorem induzierten veränderten Unterrichtsauffassungen transparent, die auch für methodische Entwicklungsprozesse Relevanz entfalten. Baumert & Kunter (2006) postulieren für die Untersuchung komplexer Unterrichtsprozesse die methodische Berücksichtigung ko-konstruktiver Handlungsdimensionen, die sich aus der wechselseitigen Interaktion von Schülern und Lehrpersonen ergeben, woraus sie ein „Angebot-Nutzungs-Modell unterrichtlicher Wirkungen mit doppelter Kontingenz“ ableiten. Dieser These soll hier

ausdrücklich zugestimmt werden. Für die methodische Entwicklung hoch-inferenter Kategorien wurde angestrebt dieser Sichtweise gerecht zu werden, indem sich die Skalen nicht nur auf Handlungsweisen der Lehrpersonen beziehen, sondern auch Reaktionen der Schüler aufnehmen. Die Beobachtungsleistung, die damit für das hoch-inferente Rating zu erbringen ist, wird dadurch jedoch anspruchsvoller. Ein motivierender Unterricht kann allerdings nicht nur auf der Basis von Handlungselementen der Lehrperson eingeschätzt werden, hingegen sind auch Verhaltensweisen der Schüler als Reaktion auf das Lehrerhandeln in den Blick nehmen. Einschränkend ist allerdings festzuhalten, dass diese ko-konstruktive Perspektive selbst durch Videostudien nur in Ansätzen geleistet werden kann, weil die Verhaltensaspekte verschiedener Schüler, basierend auf dem jeweils unterschiedlichen Erleben des Unterrichtsangebots, voneinander abweichen und deshalb nicht immer eindeutige, für die ganze Klasse geltende Interpretationen erlauben (Kunter 2005, S. 190). Diese Fragestellung stand auch nicht im Zentrum der Studie und hätte einer aufwendigeren Aufnahmetechnik bedurft. In dieser Videostudie modellierte hoch-inferente Kategorien beziehen sich deshalb im Kern auf die Handlungselemente der Lehrpersonen. Damit folgt die Studie jedoch keinem Verständnis eines lehrerzentrierten Unterrichts, sondern hierin zeigt sich ein Kompromiss, der sich aus der Notwendigkeit einer validen und reliablen Beobachtung ergibt, wenn die generierten Daten den Gütekriterien entsprechen sollen.

Für eine erste grundlegende Modellierung des Raterinventars sollten die Dimensionen Klassenführung, Diagnosekompetenz, Unterrichtsmethodische Kompetenz und Sachkompetenz übernommen werden (Helmke & Weinert 1997, S. 134), weil dies den Vorteil geboten hätte, die Daten mit den Ergebnissen anderer Instrumente des Gesamtprojektes, von dem die vorliegende Studie nur einen Ausschnitt bildet, vergleichen zu können. Von dieser Vorgehensweise wurde jedoch Abstand genommen, weil sich herausstellte, dass die Dimensionen sich nicht als Grundlage einer Videostudie eignen bzw. besser mit anderen Instrumenten erhoben werden können (vgl. Dörr & Küster 2006). Die Fragestellung, inwieweit die Studierenden im Praxisjahr über diagnostische Kompetenzen verfügen oder diese sogar ausbauen, kann nicht allein auf einer Beobachtungsebene untersucht werden, weil damit vor allem auf der Handlungsebene nicht unbedingt sichtbare Kognitionen verbunden sind. Zudem sind die Studierenden im Praxisjahr nur ansatzweise in der Lage sich ein eigenes diagnostisches Urteil zu bilden, indem z.B. die Notengebung maßgeblich in der Hand der sie betreuenden Mentoren liegt. Auch bezüglich der Sachkompetenz stellen Videostudien kein probates Untersuchungsinstrument dar, weil hier nur Ausschnitte erhoben werden, von denen man nicht zwingend auf vorhandene oder fehlende Sachkompetenz schließen kann. Die Dimensionen Klassenführung und Methodenkompetenz fanden hingegen weiter Berücksichtigung, allerdings nur als Teildimensionen. Vielmehr wurden die von Clausen, Reusser & Klieme (2003) in ihrer Videostudie vorgestellten Dimensionen als Strukturierungshilfe übernommen. Diese lauten:

- Instruktionseffizienz
- Schülerorientierung
- Kognitive Orientierung

- Klarheit und Strukturiertheit

Schon an der Bezeichnung der Dimensionen verdeutlicht sich ein stark an den Aktivitäten der Lehrperson ausgerichtetes Verständnis von Unterrichtsqualität. Gemäß der Orientierung am Angebot-Nutzungs-Modell von Helmke werden hier Unterrichtsmerkmale modelliert, die Lerngelegenheiten auf der Schülerseite wahrscheinlicher machen sollen. „Kognitive Orientierung" zielt z.B. nicht auf den Grad der kognitiven Aktivierung von Schülern (was sich in Form von Videostudien indikatorisch nur schwerlich beobachten lässt), sondern auf Aktivitäten von Lehrpersonen, die auf eine metakognitive Adressierung der Schüler schließen lassen.
Instruktionseffizienz meint basale Formen der Unterrichtsgestaltung, die mit den Teildimensionen „Umgang mit der Unterrichtszeit" und „Qualität der Organisation" näher beschrieben werden.[90] In dieser Dimension drücken sich verstärkt Anteile der Unterrichtsvorbereitung und -planung aus, wie z.B. die Vorbereitung der Unterrichtsmaterialien, ein geeigneter Einstieg, der Medienseinsatz oder die Variation zwischen Methoden und Sozialformen. Darüber hinaus werden hier aber auch klassische Aspekte der Klassenführung angesprochen, wie die umfassende zeitliche Planung der Stunde („time on task"), ein „geschmeidiger" Wechsel zwischen den Arbeitsformen und ein kompetenter Umgang mit Störungen.
Der letztgenannte Punkt zielt auf die Reaktion der Lehrperson auf Disziplinschwierigkeiten, die Wellenreuther (2004, S. 257) als „traditionellen Ansatz zum Klassenmanagement" auffasst. Den Begriff „traditionell" verwendet er deshalb in einer negativen Konnotation, weil sich darin erst eine Reaktion auf Schwierigkeiten ausdrückt, wenn es eigentlich schon zu spät ist, weniger eine vorausplanende Gestaltung des Unterrichts zur präventiven Verhinderung von Störungen. Diese Auffassung wird hier prinzipiell geteilt, allerdings übersteigt sie meiner Meinung nach die Grenzen von Verfahren, die auf eine methodische Beobachtung von Unterrichtsstunden fokussieren. Eine Analyse der Handlungselemente von Stunden, die präventiv auf die Vermeidung von Unterrichtsstörungen ausgerichtet ist, verlangt die anschließende Konfrontation der jeweiligen Lehrperson mit dem Videomaterial, um in Form eines „stimulated recalls" Aufschluss über die intendierten Handlungselemente zu erhalten. Nicht zuletzt erlauben die anderen Teildimensionen (z.B. Motivierungsfähigkeit) Rückschlüsse darauf, inwiefern in der aufgezeichneten Unterrichtsstunde disziplinäre Probleme angelegt sind.[91]

Die Dimension Schülerorientierung wurde konzipiert, um in einer ersten Teildimension zunächst zu untersuchen, wie die Lehrperson mit fehlerhaften Aussagen der Schüler umgeht und damit dokumentiert, ob ihr eine positive Fehlerkultur im Unterricht wichtig erscheint. Ist dies der Fall, sorgt sie dafür, dass individuelle Fehler nicht zu einer generellen Attribuierung werden und Schüler der

[90] Für die genauere Beschreibung der Dimensionen und Items vgl. Anhang.
[91] Der interessanten Frage des Umgangs mit Störungen, vor allem in den Grund- und Hauptschulen ein bestimmendes Thema, konnte in der eigenen Studie nur sehr begrenzt nachgegangen werden, weil die Videoaufnahmen eine disziplinierende Auswirkung auf die Schüler hatten. Bezogen auf das Verhalten der Schüler scheint mir das allerdings der einzige Faktor zu sein, für den Videostudien Artefakte generieren. Die disziplinierende Funktion der Videoaufnahmen wurde von allen Studierenden und Mentoren bestätigt.

Klasse nicht solche Muster zum Anlass nehmen, um sich über andere Schüler lustig zu machen. Damit ist sowohl ein rezeptives Verhalten als Reaktion auf Schüleräußerungen gemeint als auch ein konstruktives Handeln zur Schaffung einer positiven Fehlerkultur, um Schüler zu ermutigen ihre eigenen – notwendigerweise auch an Irrwege gebundenen – Lernfortschritte in der Klassenöffentlichkeit zu verbalisieren.
Eine zweite Teildimension fasst die Motivierungsfähigkeit, weil die Lehrperson diesbezüglich eine wichtige Vorbildfunktion inne hat und man kaum von Schülern eine interessierte und engagierte Unterrichtsbeteiligung erwarten kann, wenn die Lehrperson diese nicht auch vorlebt. Damit sind z.T. mentalitätsabhängige Verhaltensanteile angesprochen. Ein tatkräftiger und engagierter Unterrichtsstil, die Ermunterung von einzelnen Schülern sowie ein insgesamt wohlwollendes Verhalten den Schülern gegenüber dürfen von angehenden Lehrern als wichtige Voraussetzungen für einen gelingenden Unterricht erwartet werden.

In der dritten Dimension geht es um Prozesse der kognitiven Aktivierung. Gemeint ist damit ein Lehrerhandeln, welches gezielt auf die Unterstützung von Lernprozessen im Sinne einer konstruktivistisch orientierten Instruktion gerichtet ist. Erscheint die Begriffsaddition „konstruktivistische Instruktion“ auf den ersten Blick als ein Widerspruch, so wird damit genau die Herausforderung beschrieben, die Lehrpersonen erfüllen müssen, um diesen Widerspruch zu schließen.
Bestimmend für diese Sichtweise ist die Teildimension „Pacing“, die damit beschrieben werden kann, den Schülern ausreichend Zeit für selbstgesteuerte Lernprozesse zur Verfügung zu stellen. Jedoch ist anzunehmen, dass die Bereitstellung von Lernzeiten nicht ausreicht, um Schülern ein Verständnis neuer Begrifflichkeiten und Inhalte zu ermöglichen. Nimmt man das Angebots-Nutzungs-Modell von Helmke für Handlungsmöglichkeiten der Praxis und nicht nur als theoretische Bezugsfolie ernst, so resultiert daraus keine Beliebigkeit, die Verantwortung für gelingende Lernprozesse auf die Schüler zu verlegen. Vielmehr ist es die Aufgabe der Lehrperson Gelegenheitsstrukturen zu schaffen, die eine bedeutungsvolle Auseinandersetzung mit neuen Lerninhalten ermöglichen. Ihr kommt entsprechend eine Rolle als Mediator im Sinne einer lernbegleitenden Funktion zu.
In diesem Zusammenhang wurde eine zweite Teildimension – „Lehrperson als Mediator“ – konzipiert, mit der verstärkt der Frage nachgegangen werden sollte, ob die Studierenden ein Bewusstsein über die Notwendigkeit entwickelt haben, sich in die Lernprozesse ihrer Schüler hineinzuversetzen und diese didaktisch zu unterstützen. Aus dem englischsprachigen Raum hat sich der Begriff des „scaffolding“ durchgesetzt, der darauf zielt, Schülern ein unterstützendes Lerngerüst zur Verfügung zu stellen und gleichzeitig ihre individuelle Lernautonomie zu berücksichtigen. Empirisch fundierte Analysen der Kommunikation zwischen Lehrern und Schülern in offenen Lernsettings kommen jedoch zu dem Ergebnis, dass für geöffnete Unterrichtsszenarien typische Gesprächsanlässe nicht der inhaltlichen Beratung dienen, sondern vielmehr versteckte Hinweise, Bewertungen und organisatorische Anweisungen enthalten (Bräu 2006, S. 20 ff.). Solche Gesprächsinteraktionen wurden nicht als intermediäre Unterstützungsleistungen bewertet.

Die in dieser Dimension angesprochenen Handlungsmöglichkeiten können sowohl im Klassenunterricht (z.B. durch Nachfragen oder Hilfestellung bei der Ausformulierung von Ideen) als auch in Gruppen-, Partner- oder Einzelarbeitsphasen durch ein Coachen als lernstandsabhängige individuelle Beratung gezeigt werden.
Die Dimension der Lehrperson als Mediator stellt innerhalb des Raterinventars den für die Studierenden anspruchsvollsten Untersuchungsbereich dar. Deshalb wird angenommen, dass positive Veränderungen der Dimension innerhalb der drei Messzeitpunkte als geeigneter Indikator für eine Entwicklung der unterrichtlichen Handlungskompetenz gelten.

In der vierten und letzten Dimension geht es um die Klarheit und Strukturiertheit der untersuchten Stunden. Eine Teildimension bezieht sich auf die Klarheit und Verständlichkeit der Sprache, wozu auch die Korrektur inhaltlicher und grammatikalischer Fehler gehört. Ein wichtiger Bestandteil liegt zudem in der Verwendung eines altersgemäßen Sprachniveaus. Damit ist allerdings z.B. für die Grundschule nicht eine infantilisierende Sprache gemeint, sondern eine Wortwahl und Grammatik, die entsprechend ihres Alters von den Schülern verstanden werden kann. Ein Beispiel für ein solches sprachliches Verhalten gibt die Erklärung neu einzuführender Fachbegriffe und Fremdwörter ab. Der Einsatz von ironisch gemeinten Redewendungen ist in der Grund- und Hauptschule fehl am Platze, da nicht davon ausgegangen werden kann, dass die Schüler bereits ein Sprachgefühl entwickelt haben, um entsprechende Formulierungen richtig verstehen und einordnen zu können. Ebenfalls wird eine dialektgefärbte Sprache aus Gründen der eingeschränkten Verständlichkeit negativ eingeordnet.
Eine zweite Teildimension gilt der Gesprächführung der Lehrperson. Zentral sind hier einerseits die Reaktionen auf Schüleraussagen und -fragen (Wiederholung der Schüleraussage, stereotype Antworten sowie insgesamt eine Variation in der Reaktion auf Schülerantworten), andererseits der Einsatz von didaktisch sinnvollen Fragestellungen bzw. von aktivierenden Gesprächsimpulsen.
Abschließend wird die Zielklarheit der Stunde geprüft. Bestimmend sind die übergreifende Zielklarheit der Stunde oder einzelner Sequenzen, die Klarheit der Arbeitsaufträge sowie die Abstimmung und das Ineinandergreifen der einzelnen Unterrichtsphasen. Mit dem letztgenannten Punkt soll erhellt werden, inwieweit die Studierenden ein Verständnis dafür entwickelt haben, eine Lektion zusammenhängend zu planen und vor dem Hintergrund differierender Anforderungssituationen zu realisieren.

11 Ergebnisse

Am Anfang des Kapitels werden die Voraussetzungen und Ergebnisse der niedrig-inferenten Kodierungen vorgestellt. Ausgehend von der Datenlage beschränkt sich die Ergebnisdarstellung auf die Überprüfung der Interkoderreliabilitäten sowie auf die Prozentverteilungen über die Messzeitpunkte.
Im Weiteren werden die deskriptiven Kennwerte, Reliabilitäten und Voraussetzungen der Ratingskalen dargelegt. Anschließend erfolgen Signifikanztests (t-Test), um die Entwicklung über die Messzeitpunkte einschätzen zu können.

11.1 Interkoderreliabilitäten der niedrig-inferenten Kodierungen

Die Kodierungen erfolgten mit Hilfe der am Institut für die Pädagogik der Naturwissenschaften entwickelten Software „Videograph“ (Rimmele 2004). Bei der Software handelt es sich um einen Multimediaplayer, der die disjunkte Zuweisung vorher entwickelter Kodierungen auf festgelegte Zeitintervalle erlaubt. Darüber hinaus ist auch die Anfertigung von Transkripten für die Zeitintervalle mit der Software möglich, von der in dieser Studie allerdings kein Gebrauch gemacht wurde. Die elf Videoaufzeichnungen wurden über die drei Messzeitpunkte (N=33) als Zeitstichprobe in Zehn-Sekunden-Intervallen kodiert.
Das Kodiersystem wurde in aufwendiger Weise unter meiner Mitwirkung entwickelt und auf Reliabilitäten überprüft (vgl. Baer et al. 2007). Im Rahmen der eigenen Studie wurden zwei studentische Hilfskräfte für den Umgang mit dem Kodiersystem geschult und die Reliabilitäten auf der Basis von drei Videos zu verschiedenen Zeitpunkten überprüft. In Übereinstimmung mit der gängigen Forschungspraxis (Seidel 2003) galt das Training als abgeschlossen, wenn das Übereinstimmungsmaß Cohens Kappa einen Wert über .7 erreichte. Dabei wurden Übereinstimmungsmaße (Cohens Kappa) von .73 bis .91 erzielt.
Die Software „Videograph“ erlaubt den unmittelbaren Export der Kodierungsdaten in die Statistiksoftware „SPSS“, mit der die Daten weiter verarbeitet wurden.

11.2 Ergebnisse der niedrig-inferenten Kodierungen über die drei Messzeitpunkte

Um eine Auswertung der Kodierungen zu ermöglichen, wurden die Ergebnisse aller Probanden jeweils pro Messzeitpunkt aggregiert und im Hinblick auf die prozentuellen Verteilungen gemessen an der Anzahl der Gesamtintervalle der jeweiligen Stunden ausgewertet.

11.2.1 Ergebnisse für die Dimension „Unterrichtsstatus“

Die Dimension „Unterrichtsstatus“ hat einen nur geringen Aussagewert für das Anliegen der Studie. Sie dient vor allem dazu die Videoaufzeichnungen als Ergebnisstichprobe für die anderen Dimensionen zu quantifizieren. Inhaltlich kann

aus den Ergebnissen darüber hinaus abgeleitet werden, ob gravierende interne oder externe Störungen vorgelegen haben, die als Abbruch des Unterrichts gewertet werden müssen. Tabelle 11-1 zeigt die prozentuellen Verteilungen über die Messzeitpunkte:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Kein Unterricht	1,35	1,29	1,22
Vor Unterricht	2,91	2,59	2,47
Unterricht	92,78	93,33	93,39
Nach Unterrichtsende	2,94	2,75	2,90

Tabelle 11-1: Unterrichtsstatus

Aus den Ergebnissen kann man ablesen, dass die internen oder externen Störungen (kodiert als „kein Unterricht“) einen nur geringen Anteil an den Aufzeichnungen haben. Die anderen Kategorien können inhaltlich nicht ausgewertet werden, da sie von der jeweiligen Aufzeichnungsdauer abhängig sind. Abbildung 11-1 veranschaulicht die Zusammenhänge über die drei Messpunkte:

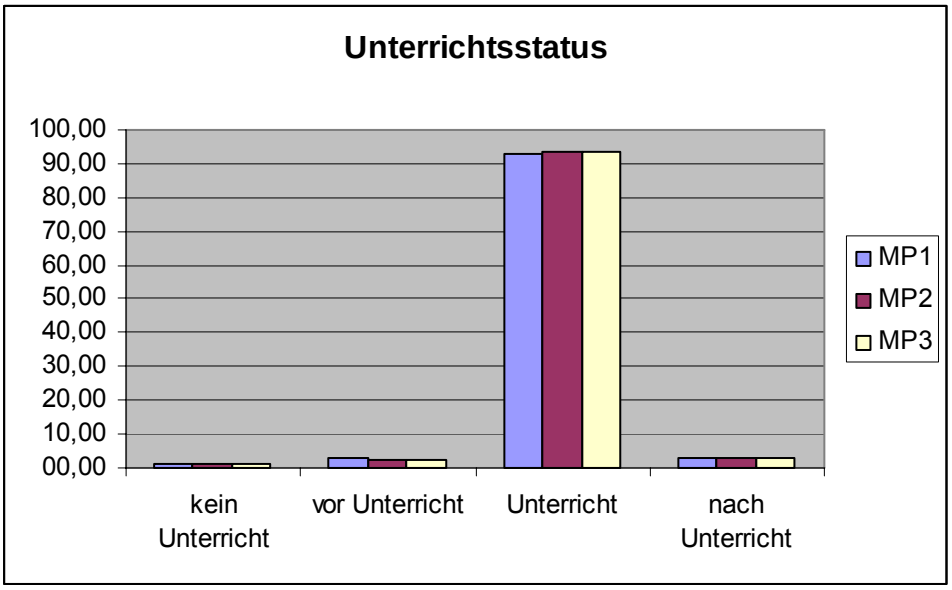

Abbildung 11-1: Unterrichtsstatus

11.2.2 Ergebnisse für die Dimension „Unterrichtliche Arbeitsformen“[92]

Mit der Dimension „Unterrichtliche Arbeitsformen“ wird untersucht, welche Sozialformen Probanden wählen. Hier interessiert vor allem das Verhältnis zwischen den Arbeitsformen im lehrerzentrierten Klassenunterricht und Schülerarbeitsphasen in Einzel-, Partner- oder Gruppenarbeit (bzw. in Kombinationen). Von Interesse ist

[92] In Anlehnung an die TIMMS-Videostudie (Seidel 2003) wird der evtl. irritierende Begriff „Arbeitsformen“ übernommen. Im Kern geht es hier zwar um Sozialformen, aber eben nicht ausschließlich.

zudem die Kategorie „Störungen durch die Lehrperson“, die dann vorliegt, wenn eine Schülerarbeitsphase durch eine Intervention der Lehrperson gestört wird. Inhaltlich lässt sich aus dieser Kategorie ableiten, inwiefern die Lehrperson mangelnde Einschätzungen dafür entwickelt hat, zu welchem Zeitpunkt die Schüler eigene Arbeitssettings gefunden haben oder die vorbereitende Arbeitsanweisung aus Sicht der Lehrperson nicht genügend klar entwickelt wurde, was eine nicht hinreichende Vorbereitung aufzeigt. Tabelle 11-2 dokumentiert die prozentuellen Kategorienwerte über die Messpunkte:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Keine	7,29	6,73	6,67
Klassenunterricht	45,97	47,48	46,76
Still-, Einzelarbeit	34,92	34,47	33,50
Partnerarbeit	0,07	0,07	0,08
Gruppenarbeit	5,62	5,79	7,82
Übergang	3,84	3,79	3,79
Mehrere Arbeitsformen	2,05	1,47	1,20
Störungen durch LP	0,26	0,24	0,21
Andere	0	0	0

Tabelle 11-2: Unterrichtliche Arbeitsformen

Betrachtet man die prozentualen Verteilungen, so fällt die starke Dominanz der Kategorien Klassenunterricht und Still- bzw. Einzelarbeit auf, welche über die drei Messpunkte nahezu konstant bleiben. Gruppenarbeitsformen kommen im Gegenzug deutlich weniger zum Einsatz, wenn sich auch gegen Ende des Praxisjahres eine Zunahme ausmachen lässt. Die Kategorie Übergänge, die die Zeit umfasst, die für den Wechsel von einer Arbeitsform in eine andere benötigt wird (z.B. Austeilen eines Arbeitsblattes), fällt mit durchschnittlich knapp vier Prozent vergleichsweise gering aus. Besonders in dieser Kategorie hätte man eine Veränderung erwarten können. Erklären lässt sich der Befund damit, dass keine aufwendigen Arbeitsformen (z.B. Werkstattunterricht) realisiert wurden. Insgesamt wird damit sehr stabil über die Probanden und Messzeitpunkte ein Unterricht beschrieben, der aus einer einleitenden Klassenunterrichtsphase und anschließenden Bearbeitung von Aufgabenstellungen in Still- oder Einzelarbeit besteht. Störungen dieser Arbeitsphase durch Herstellen der Klassenöffentlichkeit konnten kaum beobachtet werden. Hilfestellungen fanden individuell statt. Abbildung 11-2 veranschaulicht die Ergebnisse:

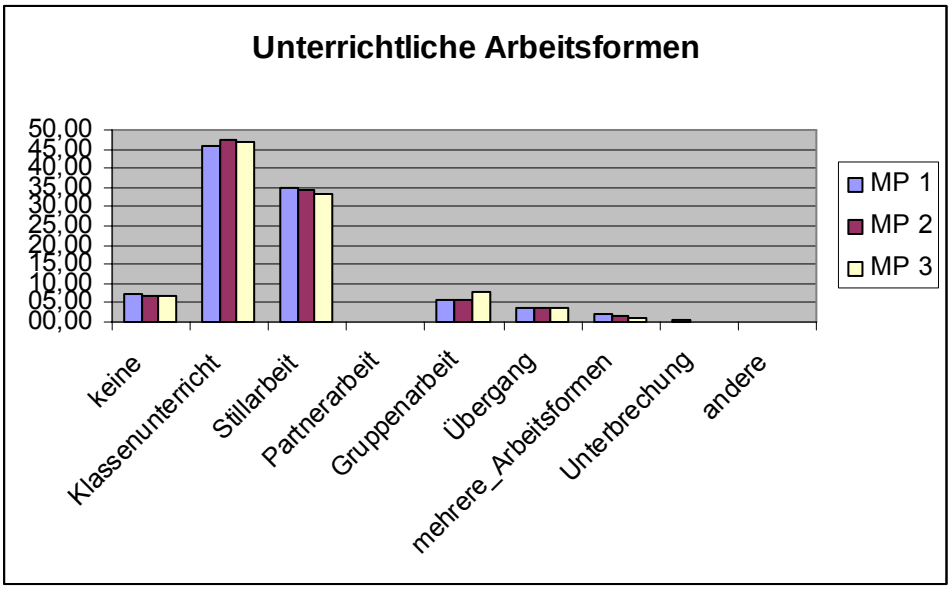

Abbildung 11-2: Unterrichtliche Arbeitsformen

11.2.3 Ergebnisse für die Dimension „Allgemeindidaktische Unterrichtsphasen“

Das Konstrukt „Allgemeindidaktische Unterrichtsphasen“ beinhaltet die makrodidaktischen Verlaufsformen des Unterrichts und damit verbundene Lerngelegenheiten für die Schüler. Aufschluss gewinnt man über die Unterrichtsplanung der Probanden und es können Rückschlüsse darüber gezogen werden, inwieweit günstige Zusammenhänge zwischen den Phasen hergestellt werden. Es besteht eine Verbindung zu der letztgenannten Dimension, weil die Kategorien „Einführung / Hinführung“ bzw. „Anwendungs- und Vertiefungsphase“ (Arbeitsphase) im Kodiermanual häufig mit den Kategorien „Klassenunterricht“ bzw. „Einzel-, Partner- und Gruppenarbeitsformen“ gekoppelt sind. Tabelle 11-3 zeigt die Ergebnisse:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Keine	7,33	6,76	6,70
Einführung	30,42	31,00	31,05
Arbeitsphase	47,30	47,08	47,61
Zusammenfassung	2,56	3,14	3,02
Prüfen	2,63	2,53	2,42
Andere	9,81	9,53	9,23

Tabelle 11-3: Allgemeindidaktische Unterrichtsphasen

Die höchsten Prozentwerte ergeben sich für die Kategorien „Einführung“ bzw. „Arbeits- und Vertiefungsphase“. Auch hier verdichtet sich das Bild eines Unterrichts, der stark durch Einführungen und das Bearbeiten von Aufgabenstellungen geprägt ist. Zusammenfassungen, die wichtig für die Schüler sein können, um den „roten Faden“ der Unterrichtsstunde nicht zu verlieren, oder abschließende metakognitive

Rückmeldungen über durchlaufene Lernprozesse erfolgen dagegen nur sehr selten. Auch Prüfungselemente konnten kaum beobachtet werden, was hier aber eher als ein Artefakt der Videoaufzeichnungen zu werten ist. Abbildung 11-3 veranschaulicht die Ausrichtung von Unterrichtsphasen:

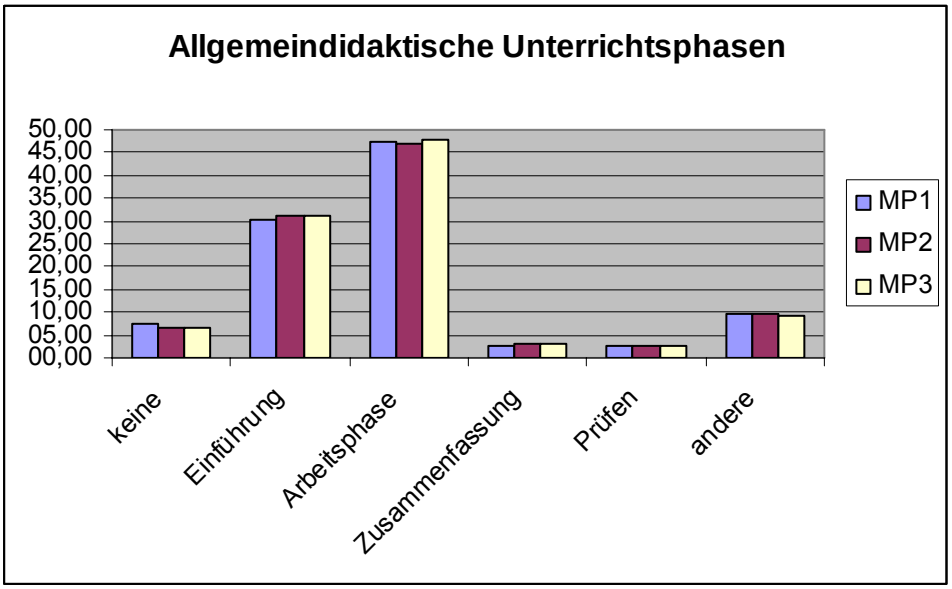

Abbildung 11-3: Allgemeindidaktische Unterrichtsphasen

11.2.4 Ergebnisse für die Dimension „Aktivitäten im Klassenunterricht“

Mit der Dimension „Aktivitäten im Klassenunterricht“ werden die Binnenstrukturen des Klassenunterrichts kodiert. Hier werden nur dann Kodes vergeben, wenn auch in der Dimension „Unterrichtliche Arbeitsformen“ die Kategorie „Klassenunterricht“ vergeben wurde. Im Untersuchungsfokus steht die Beschäftigung mit fachlichen Inhalten. Tabelle 11-4 zeigt die prozentualen Aktivitätsverteilungen:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Keine	54,16	52,64	53,35
LP-Vortrag	2,83	2,91	2,81
S-Vortrag	0	0	0
Unterrichtsgespräch	22,02	22,51	22,39
Diktat / Abschreiben	0,33	0,33	0,33
Diskussion	0	0	0
Film / Video	0	0	0
Lesen	2,52	2,63	2,67
Spielformen	4,97	5,84	5,90
Mehrere gleichzeitig	0	0	0,33
S leitet Unterricht	0	0	0
Abfrage von Resultaten	0,54	0,38	2,67
Andere	12,61	12,72	5,90

Tabelle 11-4: Aktivitäten im Klassenunterricht

Der hohe Anteil der Kategorie „keine“ ergibt sich aus dem Kodiermanual (vgl. Anhang), in dem ausschließlich die Beschäftigung mit Fachinhalten in der Klassenöffentlichkeit als „Aktivitäten im Klassenunterricht“ definiert wurde, und unterliegt deshalb hier keiner inhaltlichen Bewertung. In der Zahl spiegelt sich der hohe Anteil an Schülerarbeitsformen, der in der Höhe so nicht antizipiert worden war. Die Ergebnisse können deshalb irritieren.
Auffällig ist der geringe Prozentwert für die Kategorie „Vortrag der Lehrperson“. Der oftmals in der Literatur zitierte Frontalunterricht mit minutenlangen Vorträgen und Erklärungen bildet sich nicht ab. Umgekehrt kann man folgern, dass sich längere Erklärungsphasen nur selten finden, was für sich genommen nicht unbedingt positiv einzuschätzen ist. Es dominiert das für Deutschland typische diskursive Unterrichtsgespräch. Die Kategorie „Lesen“ konnte nur selten beobachtet werden, während die Kategorie „Spielformen“ in etwa doppelt so häufig kodiert wurde. Schülerorientierte Aktivitäten, z.B. Schülervorträge und die Übernahme der Unterrichtssteuerung durch Schüler, zeigten sich überhaupt nicht. Gleiches gilt für den Einsatz von Filmen als Unterrichtsmedium. Interpretativ ergibt sich damit ein etwas diffuses Unterrichtsbild. Weder konnten starke Schwerpunktsetzungen auf der Lehrpersonenseite durch längere Vorträge oder Erklärungen beobachtet werden, noch fanden explizite Elemente von Schülerorientierung und -aktivitäten in den aufgezeichneten Unterrichtsstunden Ausdruck. Vernachlässigt man die Schülerarbeitsformen, dominiert das fragend-entwickelnde Unterrichtsgespräch als typisches Unterrichtsmuster im Klassenunterricht. Angesichts der Diskussion um neue Unterrichtsformen, z.B. die Möglichkeiten offener Lehr-Lernformen, bildet sich ein eher konservativer Unterrichtsstil ab.
Eine mögliche Folgerung lautet, dass den Stunden eine klare Schwerpunktsetzung fehlt, was allerdings evtl. für ein Artefakt der Videoaufzeichnungen steht, wenn die Studierenden versuchen einen „goldenen Mittelweg“ zu finden. Die Verteilung von Aktivitäten visualisiert Abbildung 11-4:

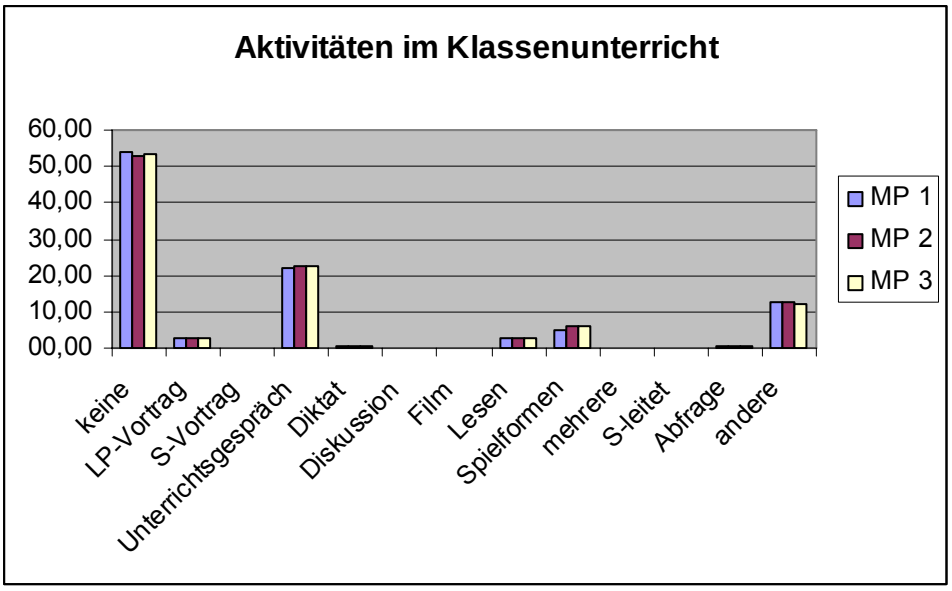

Abbildung 11-4: Aktivitäten im Klassenunterricht

11.2.5 Ergebnisse für die Dimension „Unterrichtskommunikation“

Die Dimension Unterrichtskommunikation untersucht die sprachlichen Anteile von Lehrern und Schülern an der Unterrichtskommunikation bzw. die Möglichkeit, dass keine verbale Kommunikation vorliegt. Die Dimension ist ausschließlich auf verbale Kommunikation ausgerichtet, nonverbale Anteile finden keine Berücksichtigung. Bei Schülerarbeitsformen, in denen nicht klar getrennt werden kann, ob Lehrer oder Schüler den jeweils größeren quantitativen Redeanteil haben, wird die Kategorie „Mischformen“ vergeben. Die Kategorie „niemand spricht“ wird gleichberechtigt zu den anderen Kategorien vergeben, d.h. auch wenn Redeanteile im Zeitintervall vorkommen, wird „niemand spricht“ kodiert, wenn damit der größte Zeitanteil innerhalb eines Intervalls markiert werden kann. Ein Vergleich mit der Realtime-Kommunikation ergibt, dass so die Kommunikationsanteile, gemittelt über die jeweils ganze Stunde, am besten zu erfassen sind. Tabelle 11-5 dokumentiert die prozentualen Verteilungen:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Keine	7,10	6,52	6,48
Niemand spricht	12,60	13,60	13,11
LP spricht	26,32	26,05	25,28
S spricht	5,60	6,09	6,50
Mischformen	48,27	47,60	48,50
Andere	0,12	0,15	0,15

Tabelle 11-5: Unterrichtskommunikation

Gemäß der Ergebnisse beträgt der Redeanteil der Lehrpersonen ca. 25% der Unterrichtszeit, während der Schüleranteil nur 5% beträgt. Diese Werte stimmen nicht mit denen anderer Studien überein, für die ein Lehrergesprächsanteil von ca. 70% berichtet wird (Klieme 2006a). Die große Differenz lässt sich mit dem hohen Anteil der Schülerarbeitsphasen erklären, in denen im Regelfall die Kategorie „Mischformen“ kodiert wurde, weil nicht ersichtlich war, wer den größeren Redeanteil innehatte. Die Dimension sollte vor allem die Redeanteile von Kommunikationsprozessen im Rahmen einer Klassenöffentlichkeit erfassen. Mit der aktiven Verwendung der Kategorie „niemand spricht“ wurde – im Unterschied zu anderen Studien – zudem der Redeanteil der Lehrpersonen methodisch eher reduziert, weil damit auch Satzunterbrechungen oder das Warten auf Antworten erfasst wurden. Es kann hier allerdings angesichts des fehlenden Vergleichsmaßstabes nicht entschieden werden, ob die quantifizierten durchschnittlichen 13% einen hohen oder niedrigen Wert darstellen. Eine Anhebung des Zeitintervalls, z.B. auf 30 Sekunden, könnte diesbezüglich veränderte Ergebnisse bringen. Abbildung 11-5 illustriert die errechneten Verteilungen:

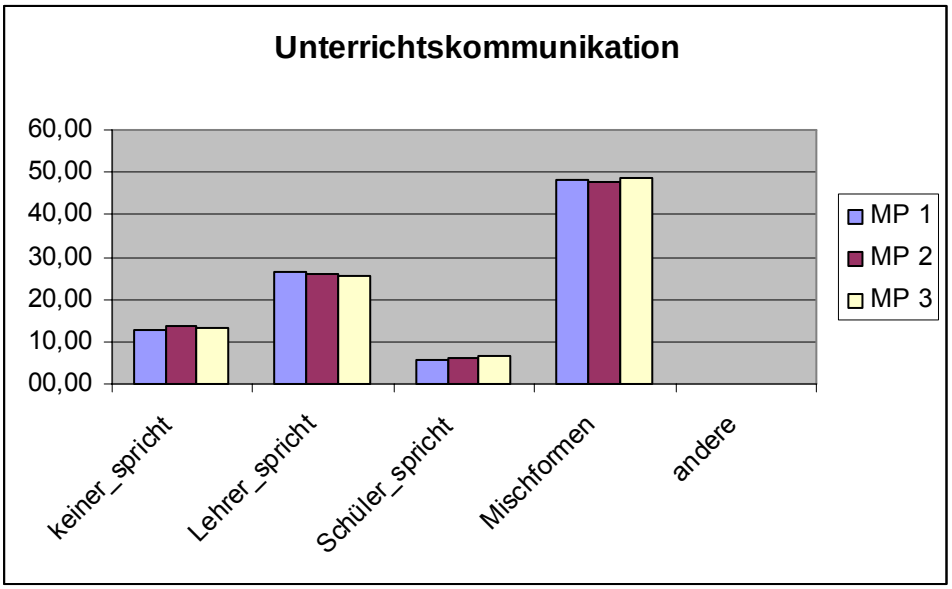

Abbildung 11-5: Unterrichtskommunikation

11.2.6 Ergebnisse für die Dimension „Strukturierung“

Mit der Dimension „Strukturierung“ sollen Unterrichtselemente erfasst werden, die im Sinne einer metakognitiven Adressierung verstanden werden können. Hauptaugenmerk liegt dabei auf der Art der Strukturierung und nicht auf den daraus resultierenden Aktivitäten. Derartige Strukturierungen schaffen Transparenz über den Unterrichtsverlauf und damit verbundene Lernziele und bieten den Schülern rückblickend ein Verständnis über ihren Lernverlauf. Tabelle 11-6 zeigt die Ergebnisse:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Keine	90,33	89,57	89,73
Ziel und Ablauf	0,08	0,08	0,07
Ziel weiterer Std.	0	0	0
Inhaltliche Aufträge	6,81	6,93	6,87
Zusammenfassen	2,56	3,14	3,02
Rückschau	0	0	0
Andere	0,23	0,29	0,31

Tabelle 11-6: Strukturierung

Zu ca. 90% können keine Strukturierungselemente ausgemacht werden. Dieser sehr hohe Prozentwert muss nicht überraschen, da nicht hinreichend davon ausgegangen werden kann, dass Formen der Strukturierung große Teile der Unterrichtsstunden einnehmen. So wurden nur die Phasen der mündlichen Instruierung durch die Lehrperson kodiert, nicht aber daraus entstehende etwaige Schüleraktivitäten. Es fällt auf, dass inhaltliche Arbeitsaufträge den größten Anteil der Strukturierung einnehmen, gefolgt von der Kategorie „Zusammenfassen“, die allerdings deutlich geringere Werte ausweist. Kategorien, die sich auf die Transparenz des

Stundenverlaufs oder Lernprozesse der Schüler beziehen, konnten praktisch nicht beobachtet werden. Daraus lässt sich folgern, dass die Studierenden nicht über ein Strukturierungsverständnis in Richtung metakognitiver Adressierungen verfügen. Allerdings kann ein Einfluss der besonderen Rolle der Studierenden als Praktikanten an den Schulen, die nur selten aufeinander folgende Stunden halten, nicht ausgeschlossen werden. Die Ergebnisse liefern Hinweise auf eine evtl. bessere Zugänglichkeit dieser Dimension über hoch-inferente Verfahren. Abbildung 11-6 veranschaulicht die Strukturierungs-Verteilungen:

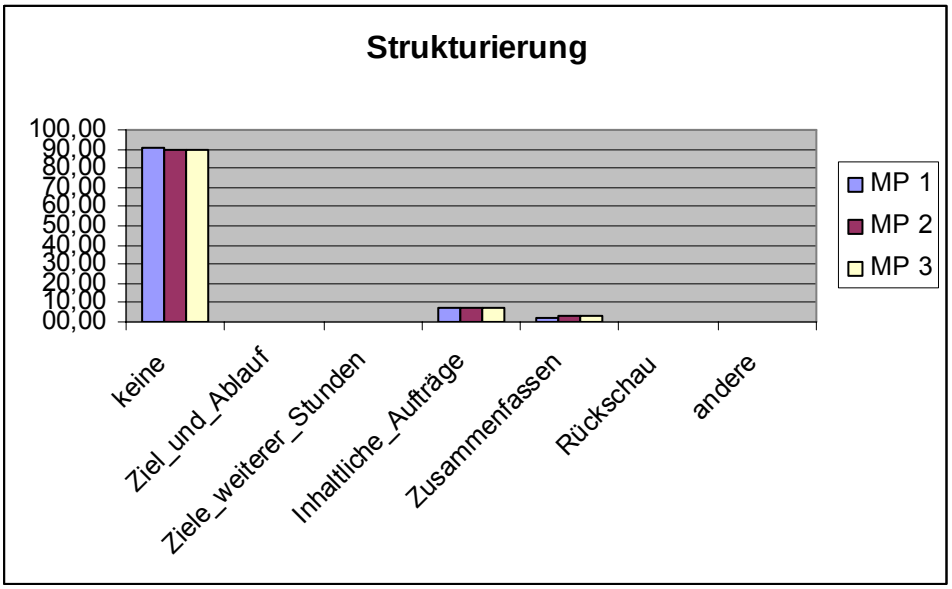

Abbildung 11-6: Strukturierung

11.2.7 Ergebnisse für die Dimension „Differenzierung“

Die Dimension soll untersuchen, ob von den Studierenden Differenzierungselemente eingesetzt werden, um eine Individualisierung des Unterrichts zu erreichen. Darüber hinaus geht es um die Frage, welche konkreten Differenzierungsmöglichkeiten die Studierenden nutzen. Tabelle 11-7 legt die Ergebnisse dar:

Kategorien	1. Messpunkt	2. Messpunkt	3. Messpunkt
Keine	94,94	96,20	96,69
Leistungsdifferenzierung-S	4,71	3,36	2,75
Leistungsdifferenzierung-LP	0,05	0,06	0,05
Neigungsdifferenzierung	0,30	0,38	0,51
Mischformen	0	0	0

Tabelle 11-7: Differenzierung

Ähnlich zu der Dimension „Strukturierung“ ist auffällig, dass Differenzierungselemente nur selten vorkommen. Mit einem Durchschnitt von ca. 96% findet ein undifferenzierter Unterricht statt. Bedenkt man den hohen Anteil von Schülerarbeitsphasen an der gesamten Unterrichtszeit, so überrascht dieses

Ergebnis. Als einziges Differenzierungselement kommt eine Leistungsdifferenzierung durch die Schüler selbst zum Tragen. In der Praxis geschieht dies durch eingesetzte Aufgabenblätter verschiedener Schwierigkeitsgrade. Andere Differenzierungen konnten nicht beobachtet werden. Auch wenn die qualitative Interpretation niedrig-inferenter Kodierungen nicht unproblematisch ist, kann konstatiert werden, dass dieses Ergebnis von allen Dimensionen am einseitigsten und damit im Sinne förderlichen Unterrichts negativsten ausfällt, zumal die Differenzierungen über die Messpunkte noch abnehmen (wenn auch in einem eher vernachlässigbaren Bereich). Wiederum kann nicht ausgeschlossen werden, dass die besondere Situation der Praktikanten eine Rolle für den geringen Differenzierungsgrad spielt. Abbildung 11-7 illustriert die Differenzierungsformen:

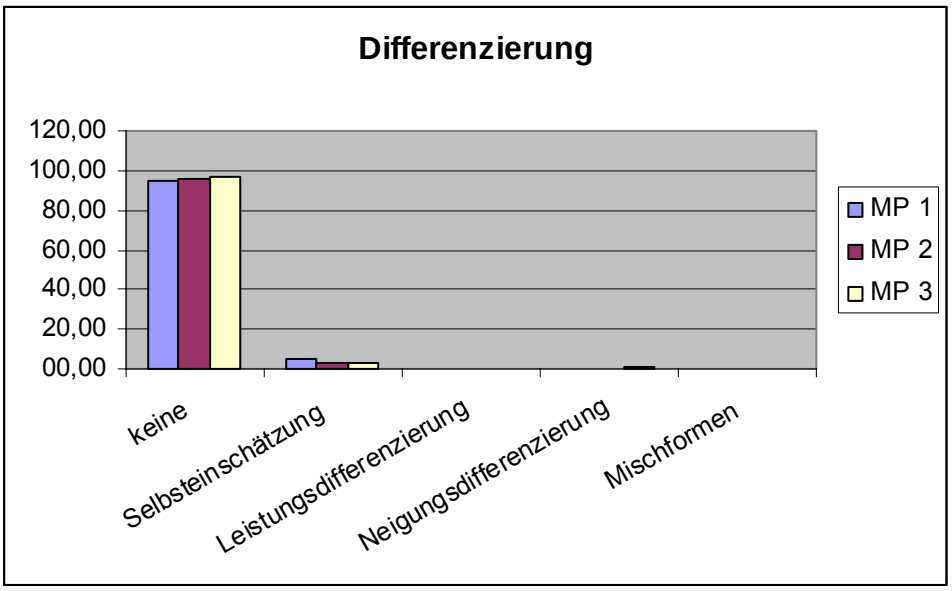

Abbildung 11-7: Differenzierung

11.2.8 Diskussion der Kodierungsergebnisse

Wenn auch keine Veränderungshypothesen für die drei Messpunkte aufgestellt worden sind, überraschen die Ergebnisse dennoch. Auf der methodischen Basis niedrig-inferenter Beobachtungen können kaum Veränderungen festgestellt werden. Über die Probanden und Messzeitpunkte lassen sich allerdings große Streuungen ausmachen, die unspezifisch ausfallen und bestenfalls auf einer Einzelfallebene untersucht werden können. Da einzelfallbezogene Personenprofile nicht im Interesse der Arbeit liegen, wird von einer solchen Perspektive jedoch abgesehen.
Ähnlich verhält es sich mit einem Vergleich aufgenommener Unterrichtstunden zwischen den am Modellversuch beteiligten Schulen (Grund- bzw. Hauptschulen). Er wäre prinzipiell interessant, ist aber von spezifischen Variablen beeinflusst (z.B. Schülerzusammensetzung, spezifische Herausforderungen an den Schulen), die in der Studie nicht kontrolliert wurden. Zudem würde ein solcher Vergleich verringerte Stichprobengrößen mit sich bringen, wodurch die Aussagen nicht an Validität gewinnen. Forschungsethisch wäre eine vergleichende Betrachtung schließlich problematisch, weil die Ergebnisse unmittelbar auf die jeweiligen Schulen

rückkoppelbar wären und eine Identifizierung der betreffenden Studierenden tendenziell einfacher möglich wäre. Aus diesen Gründen scheint es angemessener die Gesamtstichprobe als Untersuchungseinheit festzulegen.[93]
Aggregiert über die Probanden und Messzeitpunkte bleibt allerdings das unvermutete Ergebnis einer Stabilität der Ergebnisse. Somit können die Fragestellungen der Studie (vgl. Kap. 9) zu generellen oder spezifischen Entwicklungen im Rahmen der Teilnahme an verlängerten Praxislernangeboten auf dieser Datengrundlage nur dahingehend beantwortet werden, dass keine Entwicklungen aufzudecken sind (unter Vernachlässigung einer Einzelfallanalyse). Angesichts der kleinen Stichprobe hätte man anders vermuten können, dass unspezifische und inhaltlich schwer interpretierbare Veränderungen auftauchen.
In Folge der stabilen Datenlage erscheinen Signifikanzprüfungen über die Messpunkte obsolet, weil sie, basierend auf Prozentwerten, eingedenk der kleinen Stichprobe auf deutlich höhere Varianzen angewiesen wären. Es stellt sich nun die Frage der Ergebnisinterpretation. Grundsätzlich sind zwei Lesarten denkbar:

(1) Die Ergebnisse dokumentieren ein für das Praxisfeld Schule typisches, sehr stabiles Unterrichtsmuster, das in der Praxis schnell einsozialisiert wird.

(2) Die Ergebnisse müssen als Artefakte der Aufzeichnungssituation gewertet werden, wobei verschiedene Faktoren als Einflussgrößen differenziert werden können.

Für die erste Interpretation spricht die Tatsache, dass zwischen dem ersten allgemeinen Tagespraktikum und dem Praxisjahr konzeptionell keine Differenzen ausgemacht werden können. D.h. es fließen Unterrichtskonzeptionen in die Ergebnisse ein, die nicht nur an den am Praxisjahr beteiligten Schulen, sondern auch an anderen Schulen realisiert worden sind. Für das Tagespraktikum kann man zudem einen größeren Einfluss der Hochschule durch die betreuenden Mentoren vermuten, als das im Praxisjahr der Fall ist. Dies gilt sowohl für die Unterrichtsplanung als auch für Feedback im Anschluss an durchgeführte Unterrichtsstunden. Damit gehen mittelbar Einflüsse verschiedener Personengruppen in die aufgezeichneten Stunden ein, wobei hier allerdings nicht entschieden werden kann, was das für die Realisierung durch die Studierenden heißt.
Für die Aufnahmen innerhalb des Praxisjahres ist zudem anzunehmen, dass die Unterrichtsplanung im hohen Maße mit den betreuenden Mentoren der Schulen besprochen wurde und damit ein geteiltes Unterrichtsverständnis auch der Mentoren wenigstens in die Planung eingeht. Darüber hinaus waren die Mentoren bei allen Aufzeichnungen anwesend und es fanden sich kaum Anzeichen dafür, dass die Unterrichtsdurchführung ihrem Verständnis zuwiderlief. Diese Fragestellung wurde jedoch nicht systematisch untersucht.
Nimmt man an, dass die Ergebnisse nicht nur Ausdruck individueller Entwicklungsverläufe der Studierenden sind, sondern man daraus auch Rückschlüsse auf das Ausbildungsfeld Schule ziehen kann (angesichts der kleinen

[93] Entsprechende Argumente gelten für die Auswertungen der hoch-inferenten Einschätzungen (vgl. Kap. 11.5).

Stichprobe kann dies nur mit Einschränkungen behauptet werden), so spiegeln sie möglicherweise ein Unterrichtsmuster wider, welches nicht unbedingt spezifisch für das Praxisfeld Schule insgesamt, aber typisch für Beobachtungs- und gewissermaßen Testsituationen ist, denen angehende Lehrer im Laufe ihrer Ausbildung häufig gegenüberstehen, während das für die Berufsausübung nur selten zutrifft. In diesem Fall würde es sich nicht um ein Artefakt handeln, sondern um ein Moment der Berufssozialisation, welches charakteristisch für Erfahrungswelten der Lehrerbildung, aber nicht für das Handeln von praktizierenden Lehrern ist. Die noch in der Ausbildung Stehenden hätten dann eigene Wege gefunden, um mit extern an sie herangetragenen, testähnlichen Situationen umzugehen, die berufssozialisatorisch ein Bild des „Best-Practice-Unterrichts“ vermitteln, welches ökologisch für das Feld Schule nicht unbedingt valide ist. Für die eigene Fragestellung bedeutet diese Unterscheidung allerdings keine eingeschränkte Validität der Forschungsergebnisse, weil es um die Entwicklung von Studierenden in verlängerten Praxisphasen und nicht um die Berufsrealität von fertigen Lehrern geht. In diesem Zusammenhang wäre eine Untersuchung von Referendaren im Vorbereitungsdienst oder der Eingangsphase von Lehrern zu Beginn ihrer beruflichen Laufbahn interessant, um Vergleichsmöglichkeiten der diesbezüglichen Unterrichtsmuster vornehmen zu können. Speziell für die Berufseingangsphase in Deutschland liegen kaum empirische Daten vor (vgl. Larcher-Klee 2005 für die Schweiz).

Für die zweite Lesart spricht, dass speziell bei kleinen Stichproben Artefakte nie ganz ausgeschlossen werden können. Damit ist die Frage angesprochen, ob die Videoaufnahmen tatsächlich Entwicklungsprozesse der Studierenden im Laufe ihrer Ausbildung abbilden oder ob andere Faktoren für das unterrichtliche Verhalten ursächlich sind. Die Anwesenheit von Kameras im Klassenzimmer (und der sie bedienenden Personen) kann potentiell als Invasivitätsquelle dienen und unterrichtliches Verhalten verändern. Andere Studien (Clausen 2002; Seidel 2003) berichten von keiner nennenswerten Beeinflussung. Für die eigene Studie wird diese Auffassung geteilt, da die Studierenden sehr frühzeitig über die Videoaufnahmen in Kenntnis gesetzt wurden und mit der Teilnahme am Praxisjahr ihr Einverständnis dafür gegeben haben. Einschränkungen für die Videoaufnahmen wurden nur dahingehend vorgenommen, dass weder Schüler noch Lehrpersonen während der Aufnahmen den Raum verlassen sollten bzw. dass die aufgenommene Unterrichtsstunde für Außenstehende möglichst nachvollziehbar sein und nicht im Kern auf vorangegangenen Stunden basieren sollte. Damit wurde eine Verringerung der Varianz didaktischer Settings im Sinne von Sonderfällen (z.B. Vorbereitungen auf das Weihnachtsfest) angestrebt, um die Durchführung der Aufnahmen zu erleichtern und deren Vergleichbarkeit zu erhöhen.

Während der Aufnahmen konnte weder bei den Studierenden noch bei den Schülern ein auffälliges der Kamera zugewandtes Verhalten beobachtet werden, von dem man auf eine generelle Störung der Unterrichtsstunden schließen konnte, wenn auch teilweise ein Interesse einzelner Schüler an der Kamera feststellbar war. Eine anschließende mündliche Befragung der Studierenden und Mentoren deckte keine Besonderheiten auf, weshalb die eigentliche Durchführung der Videoaufnahmen als Ursache für Artefakte weitgehend ausgeschlossen wird.

Anders verhält es sich mit der Frage, inwieweit die Studierenden ihre Unterrichtsplanung untereinander abgesprochen haben. Hier mag sehr wohl ein Einflussfaktor vorliegen, der zum Teil zu Artefakten führen kann. Aus meiner Sicht ist es aber unwahrscheinlich, dass entsprechende wechselseitige Beeinflussungen ausschließlich für die Videoaufnahmen gelten. Danach handelt es sich nicht um ein Artefakt, sondern um einen methodisch kaum kontrollierbaren Einflussfaktor der Teilnahme am Praxisjahr. Darüber hinaus würde durch eine Kontrolle lediglich eine Stabilität innerhalb der Messpunkte, nicht jedoch im längsschnittlichen Vergleich, plausibel.
Ein dritter Faktor zur Erzeugung von Artefakten könnte im Erfahrungsmoment der Videoaufnahmen liegen. Die Studierenden mögen den Eindruck gewonnen haben, mit ihrer ersten Unterrichtsdurchführung die Videoaufnahme erfolgreich bewältigt zu haben und replizieren entsprechend die Durchführung. Bedenkt man aber, wie viel unterrichtliche Erfahrung die Studierenden im Praxisjahr sammeln, erscheint diese These eher unwahrscheinlich. Man kann jedoch vermuten, dass die Studierenden für die Videoaufnahmen eine Unterrichtsplanung vorziehen, die wenig Risiko zum Scheitern in sich birgt und z.B. nicht unbedingt Unruhe im Klassenzimmer provoziert. Die Ergebnisse lassen sich teilweise in diese Richtung deuten und können ein Erklärungsfaktor für die Stabilität über die Messpunkte sein. Dagegen steht, dass die Studierenden die Unterrichtsplanung im Regelfall nicht alleine vollzogen, sondern mit ihren Mentoren oder Kommilitonen abgesprochen haben. Absprachen mit den Mentoren lassen vermuten, dass ein geteiltes Verständnis darüber vorliegt, wie man Unterrichtsstunden plant und durchführt, die von schulexternen Personen beobachtet oder aufgezeichnet werden.
Abschließend soll festgehalten werden, dass die Unterrichtsaufzeichnungen nicht frei von Einflussquellen sind. In welchem Umfang das der Fall ist, kann nur ansatzweise vermutet werden. Der Einfluss der Kameras im Klassenzimmer auf das Unterrichtsverhalten wird für vernachlässigbar eingeschätzt. Die Beobachtungssituation an sich kann allerdings einen Einfluss ausüben, insofern die Videoaufzeichnungen nicht prototypisch für den Unterrichtsalltag stehen, der in seiner Facettenvielfalt durch wenige Aufnahmen nicht vollständig abbildbar ist.
Aufgezeichnet wurden Unterrichtssituationen, die Tendenzen zu einem risikoarmen Unterrichtsstil aufweisen und vermutlich typisch für das Verhalten von Personen in Beobachtungssituationen sind, wie sie von angehenden Lehrern häufiger erlebt werden. Das offenbar geteilte Unterrichtsverständnis zwischen Studierenden und Mentoren lässt aber wiederum die These zu, dass die aufgezeichneten Stunden in keinem hohen Maße von der unterrichtlichen Normalsituation abweichen. Deshalb wird angenommen, dass die Unterrichtsaufzeichnungen grundsätzlich geeignet sind, um längsschnittliche Entwicklungsprozesse zu untersuchen.
Die genannten Einschränkungen bieten Erklärungspotential für die Einordnung der Ergebnisse. Es bleibt abzuwarten, ob die durchgeführten Interviews diesbezüglich weitere Erklärungen liefern (vgl. Müller i.V.).

11.3 Interraterreliabilitäten

Hoch-inferente Ratingverfahren stellen große Anforderungen an die Expertise der Rater. Für die Durchführung sind umfassende internalisierte Modelle zur Beurteilung komplexer Unterrichtsprozesse notwendig. Die interindividuelle Überprüfung der Ergebnisse wird damit umso notwendiger, weil es im Vergleich zur Kodierung niedrig-inferenter Beobachtungsverfahren problematischer ist, kriterielle Beobachtungsvorgaben zu formulieren. Wie bereits oben angedeutet wird hier die Auffassung vertreten, dass hoch-inferente Ratings nur dann als Expertenurteile gelten können, wenn sie tatsächlich auf der Expertise der Rater basieren. Zu scharf definierte kriterielle Vorgaben führen dazu, dass nur noch von einer Expertise für das jeweilige Raterinventar gesprochen werden kann, womit die Absicht verbunden ist, Beobachtungsvarianzen zwischen mehreren Ratern zu minimieren. Diese Absicht ist methodisch durchaus nachvollziehbar, weil große Unterschiede zwischen mehreren Ratern nur dergestalt interpretiert werden können, dass offenbar unterschiedliche Konstrukte beobachtet werden oder unterschiedliche Wertewichtigkeitshierarchien herrschen. Die Bedeutsamkeit eines intensiven diskursiven Austausches wird damit elementar.
Im Zusammenhang mit dieser Diskussion ist die von Wirtz & Caspar (2002) eingeführte Unterscheidung zwischen der Beurteilerübereinstimmung und der Beurteilerreliabilität wesentlich. Die Autoren plädieren dafür, dann von einer Beurteilerübereinstimmung zu sprechen, wenn die punktgenaue Übereinstimmung zwischen den Beurteilern wichtig ist, weil Varianzen zu komplett verschiedenen Aussagen führen. Bei niedrig-inferenten Kodierverfahren ist das z.B. der Fall, wenn innerhalb einer Dimension unterschiedliche Kategorien ausgewählt werden. Daraus resultiert auch die Verwendung entsprechender Übereinstimmungsmaße, wie z.B. Cohens Kappa, mit denen nicht nur die Übereinstimmung berechnet, sondern auch die zufällig erwartbare Übereinstimmung mitberücksichtigt wird. Im Unterschied dazu sprechen sich die Autoren dann für die Verwendung von Reliabilitätsmaßen aus, wenn keine exakte Gleichheit der eingeschätzten Merkmalsausprägungen verlangt ist. Hingegen muss die relative Lage der Werte zum Mittelwert der Rater für die jeweiligen eingeschätzten Personen ähnlich sein. Dieser potentielle Zugang gilt nur für intervallskalierte Ratings, weil nur in diesem Fall ein Vergleich und die Interpretation der Distanzen zwischen den Werten auf der Skala möglich sind. Wirtz & Caspar (2002, S. 157) stellen fest:

„Bei intervallskalierten Ratings ist ein Vergleich der Distanzen zwischen den Werten auf der verwendeten Ratingskala möglich: zwischen allen benachbarten Werten liegen gleiche Abstände, und gleichen Unterschieden in den Ausprägungen des Merkmals in der Empirie muss eine gleiche Zahlendifferenz entsprechen."

Die Autoren kommen deshalb zu dem Ergebnis, dass bei intervallskalierten Ratings die Intraklassenkorrellation (ICC) im Allgemeinen die angemessene Methode zur Reliabilitätsbestimmung darstellt. Diese Auffassung wird hier ausdrücklich geteilt, auch wenn sich in anderen Studien Cohens Kappa (Clausen 2002) oder der Generalisierbarkeitskoeffizient (Clausen, Reusser & Klieme 2003) durchgesetzt haben. Die Reliabilitäten wurden an drei Filmen der Stichprobe zu unterschiedlichen

Zeitpunkten und über zwei Rater (davon jeweils einer meine eigene Person) berechnet. Tabelle 11-8 zeigt die Koeffizienten für die vier Dimensionen über drei Messpunkte:

Dimensionen	1. MP	2. MP	3. MP	Gesamt
Instruktionseffizienz	.571	.533	.774	.626
Schülerorientierung	.533	.571	.774	.626
Kognitive Aktivierung	.697	.583	.900	.730
Klarheit / Strukturiertheit	.632	.615	.764	.670

Tabelle 11-8: Koeffizienten der Intraklassenkorrelation

Gemittelt über alle Dimensionen und drei Messpunkte ergibt sich ein Koeffizient der Intraklassenkorrelation von .66. Der Wert liegt damit knapp unterhalb des allseitig für eine gute Reliabilität anerkannten Wertes von .7 (Greve & Wentura 1997). Wirtz & Caspar (2002, S. 160) stellen in diesem Zusammenhang heraus, dass die Ausprägung der Koeffizienten immer in Abhängigkeit des zu messenden Merkmals und der untersuchten Stichprobe beurteilt werden muss. Insofern können nicht „gute“ Reliabilitäten einer Skala, sondern nur Werte der jeweiligen Anwendung ausgewiesen werden. Die Intraklassenkorrelation ist zudem ein Maß der Varianzaufklärung und entsprechend auf Varianzen angewiesen. Die Reliabilitätsmaße werden umso größer, je größer die Varianz in der Stichprobe ausfällt. Bei kleinen Stichproben, die zudem als natürliche Gruppen relativ homogen ausfallen (wovon bei der eigenen Stichprobe ausgegangen werden kann), können auch kleinere Koeffizienten akzeptiert werden (ebd., S.161). Deshalb wird angenommenen, dass ein für die eigene Studie ausreichendes Reliabilitätsmaß vorliegt, um eine den Gütekriterien empirischer Forschung genügende hinreichende Reliabilität und Validität der Daten beanspruchen zu können.
Aus Ressourcengründen konnten nicht alle Videos von zwei oder mehreren Ratern ausgewertet werden, weshalb die nachfolgenden Daten aus hoch-inferenten Einschätzungen stammen, die von mir selbst durchgeführt wurden.

11.4 Reliabilitäten der Skalen

Das eingesetzte Raterinventar besteht aus vier Großdimensionen und neun Subskalen mit insgesamt 38 Items. Bei der Entwicklung wurden Items und Skalen von Clausen, Reusser & Klieme (2003) übernommen, z.T. aber auch eigene Items entwickelt. Zudem wurden die Beschreibungen der Items für das Manual selbst im Forschungsprojekt „Standarderreichung in der Lehrerinnen- und Lehrerbildung: Analyse der berufsfeldorientierten Ausbildung“ (Baer et al. 2007) unter meiner Mitwirkung formuliert. Zur Untersuchung der internen Konsistenz wurden die neun Subskalen auf ihre Reliabilität überprüft (Cronbachs Alpha). Tabelle 11-9 zeigt die Werte:

Dimension	Anzahl der Items	Cronbachs Alpha
Umgang mit der Unterrichtszeit	4	.570
Qualität der Organisation	4	.581
Umgang mit Fehlern	3	.888
Motivierungsfähigkeit	5	.900
Pacing	4	.631
Lehrperson als Mediator	7	.768
Sprachliche Qualität	3	.349
Gesprächsführung	5	.837
Zielklarheit	3	.782

Tabelle 11-9: Reliabilitäten der Subskalen (Cronbachs Alpha)

Die Auswertung erbringt zufrieden stellende Daten für fünf Skalen, mittlere Werte für drei Skalen und eher unbefriedigende Werte für eine Skala. Angesichts der kleinen Stichprobe stellen die Werte eine überzeugende Grundlage dar, um weitere varianzanalytische Untersuchungen auf dem Niveau der Skalen und nicht auf Einzelitemniveau zu berechnen.

11.5 Deskriptive Kennwerte

Tabelle 11-10 zeigt die Mittelwerte und Standardabweichungen für die neun Unterrichtskonstrukte über die drei Messpunkte:

	1. MP	SD	2. MP	SD	3. MP	SD	N
Umgang mit der Unterrichtszeit	3,05	0,38	4,34	0,59	4,7	0,56	11
Qualität der Organisation	3,77	0,52	4,57	0,43	4,75	0,53	11
Umgang mit Fehlern	3,06	0,49	3,73	0,85	4	1,1	11
Motivierungsfähigkeit	3,09	0,77	3,8	0,83	4,09	0,96	11
Pacing	3,32	0,54	3,93	0,39	4,39	0,61	11
Lehrperson als Mediator	2,6	0,51	3,05	0,54	3,47	0,77	11
Sprachliche Qualität	3,24	0,58	3,72	0,44	4,03	0,55	11
Gesprächsführung	3,2	0,46	3,71	0,61	4,09	0,53	11
Zielklarheit	3,32	0,74	3,88	0,64	4,36	0,66	11

Tabelle 11-10: Mittelwerte und Standardabweichungen der Stichprobe

Betrachtet man die Werte genauer, so fällt auf, dass alle Mittelwerte über die drei Messpunkte zunehmen. Vergleicht man die Minimum- / Maximum-Werte über die Messzeitpunkte, so fällt der Zuwachs auf der Skala mit dem geringsten Mittelwert zum ersten Messzeitpunkt („Lehrperson als Mediator“; M=2,6) am zweiten (M=3,05) und dritten Messzeitpunkt (M=3,47) vergleichsweise gering, aber stabil aus. Dagegen zeigt sich der Zuwachs auf der Skala, in der zum ersten Messzeitpunkt der höchste Mittelwert erreicht wurde („Qualität der Organisation; M=3,77), am zweiten Messzeitpunkt (M=4,57) deutlicher, wobei sich allerdings die Zunahme im Vergleich

des zweiten und dritten Messzeitpunktes (M=4,75) relativ gering ausnimmt. Bezüglich der genannten Skalen gibt es also unterschiedliche Entwicklungssprünge, wenn auch die positive Entwicklungstendenz erhalten bleibt. Dabei ist offensichtlich, dass die Skalen, für die zum ersten Messzeitpunkt die höchsten bzw. niedrigsten Werte festgestellt wurden, entsprechende Rangplätze auch zum dritten Messzeitpunkt innehaben. Inhaltlich kann man daraus unterschiedliche Schwierigkeitsgrade für jeweilige Verbesserungen im längsschnittlichen Vergleich folgern. Methodisch können z.T. Deckeneffekte nicht gänzlich ausgeschlossen werden.
Tendenziell fällt außerdem die Zunahme der Mittelwerte vom ersten zum zweiten Messpunkt größer aus als vom zweiten zum dritten Messpunkt. Die Abbildungen 11-8 – 11-16 illustrieren die Entwicklung:

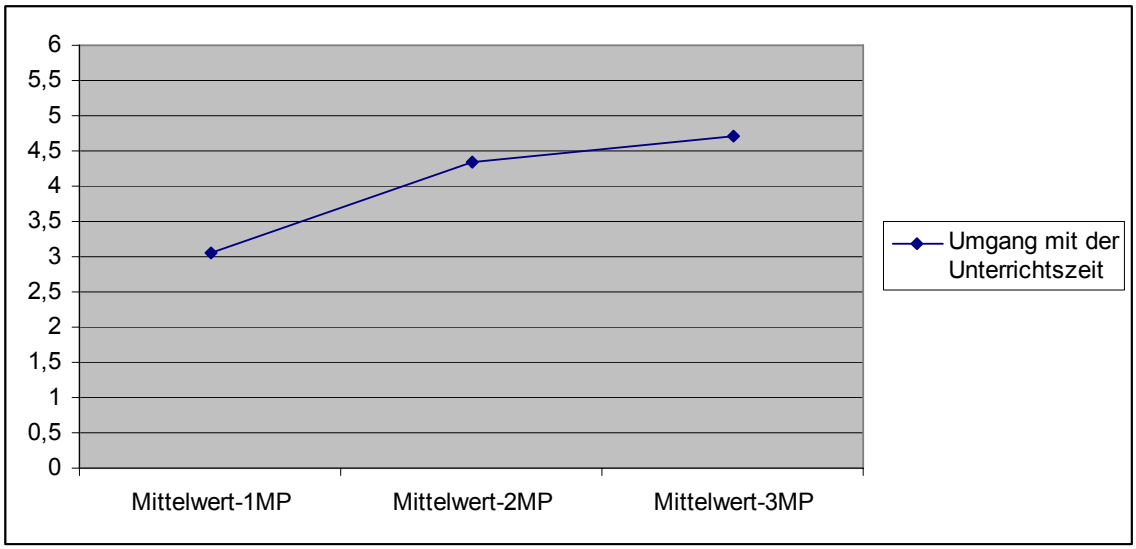

Abbildung 11-8: Umgang mit der Unterrichtzeit

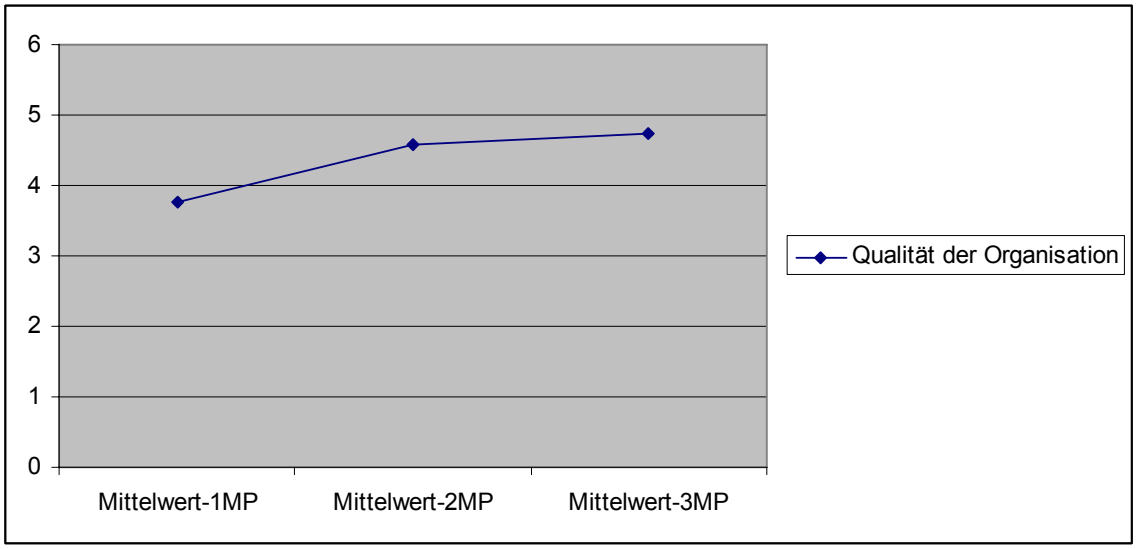

Abbildung 11-9: Qualität der Organisation

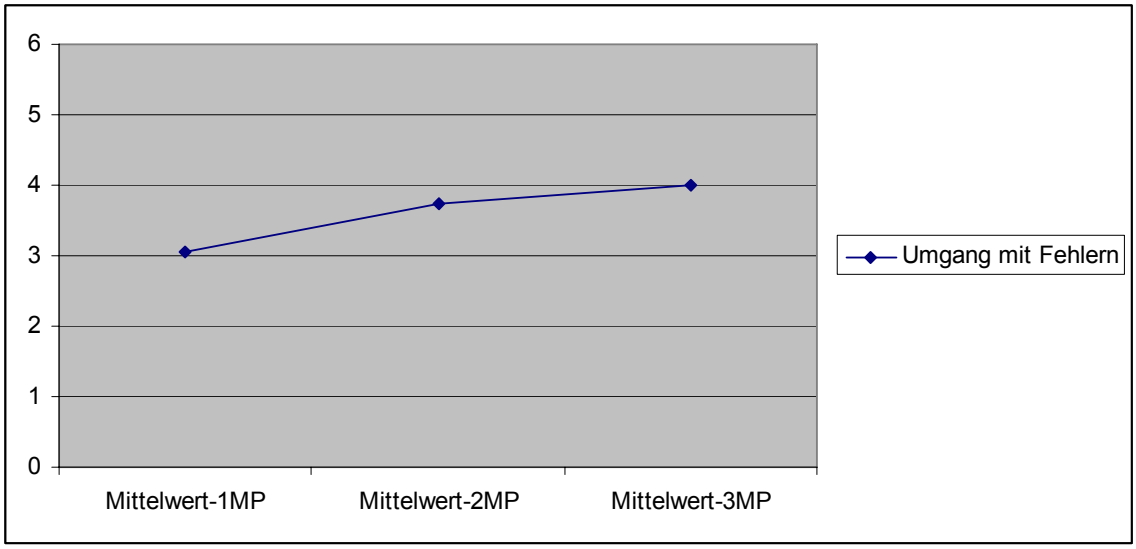

Abbildung 11-10: Umgang mit Fehlern

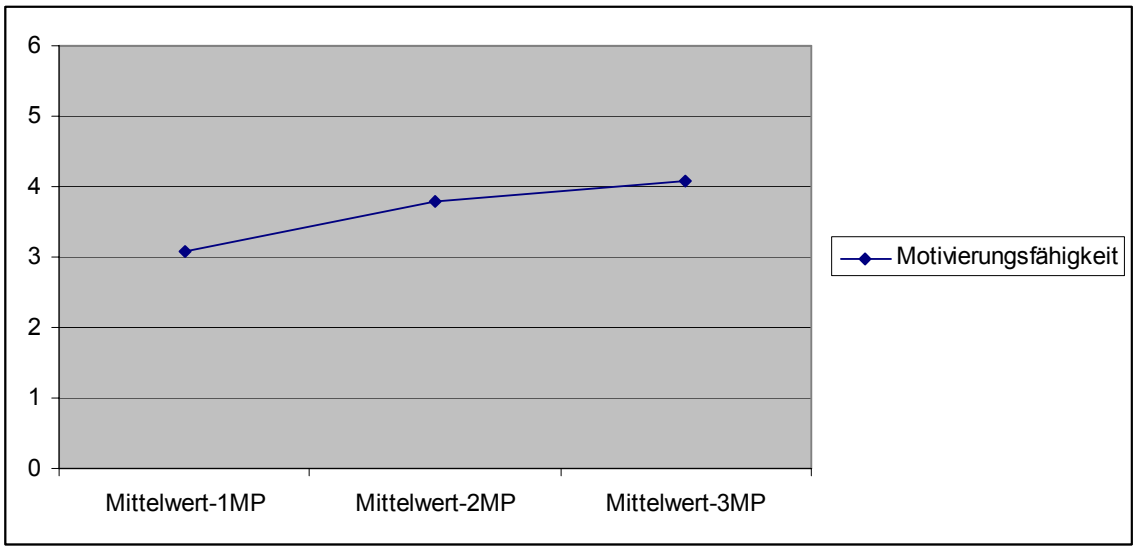

Abbildung 11-11: Motivierungsfähigkeit

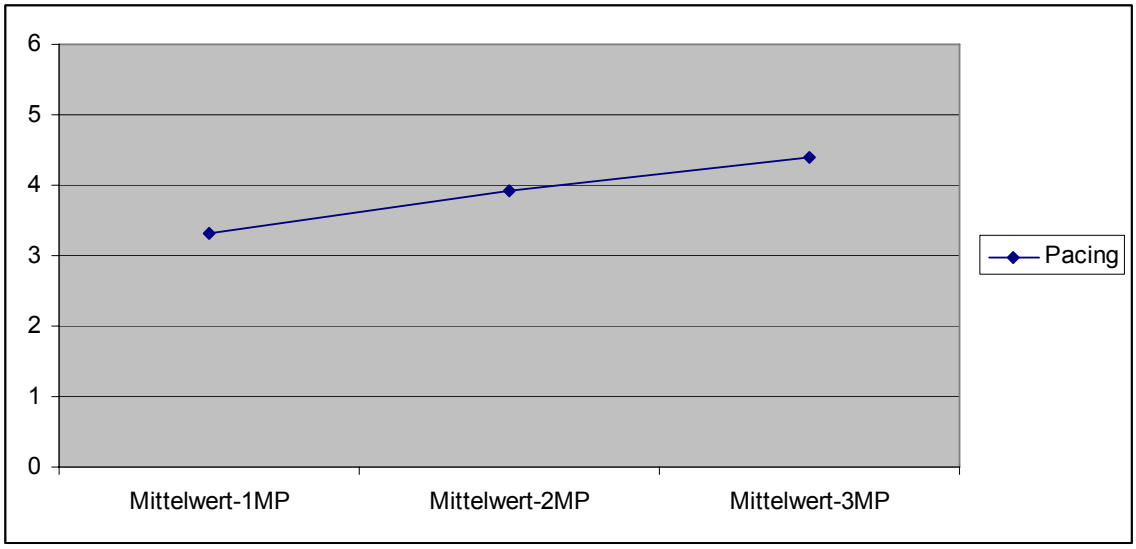

Abbildung 11-12: Pacing

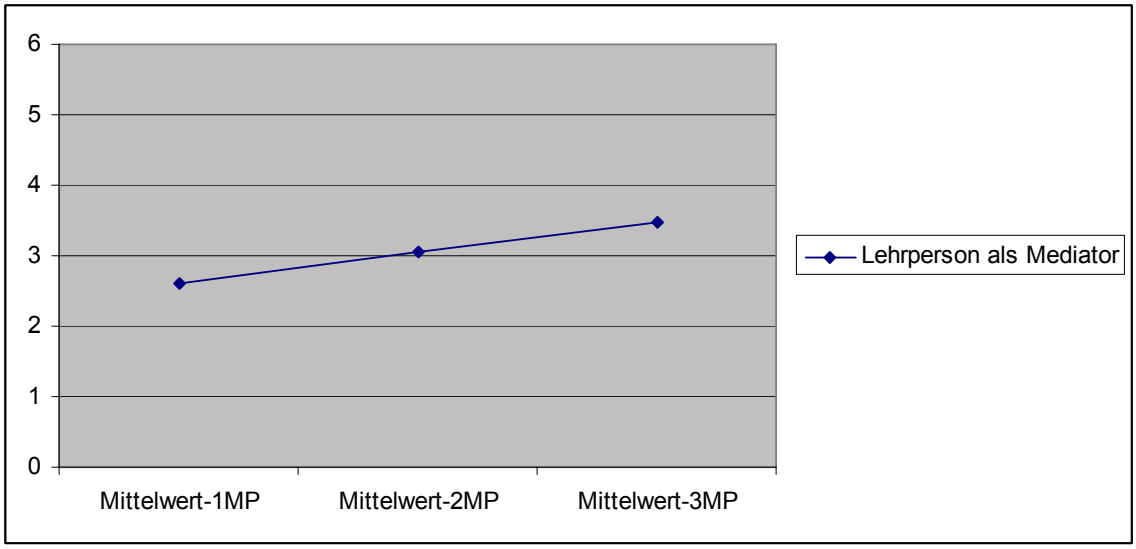

Abbildung 11-13: Lehrperson als Mediator

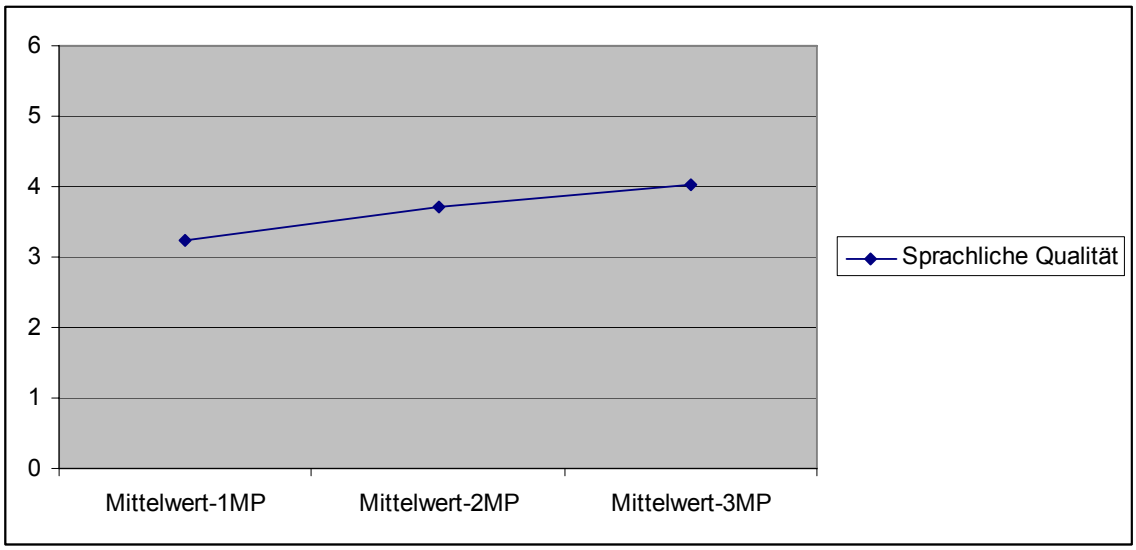

Abbildung 11-14: Sprachliche Qualität

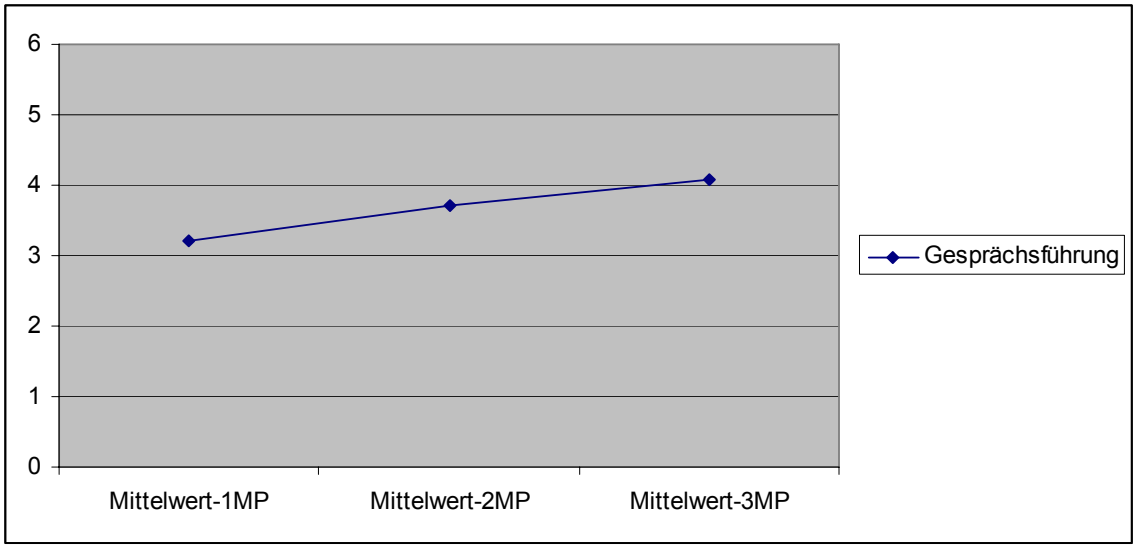

Abbildung 11-15: Gesprächsführung

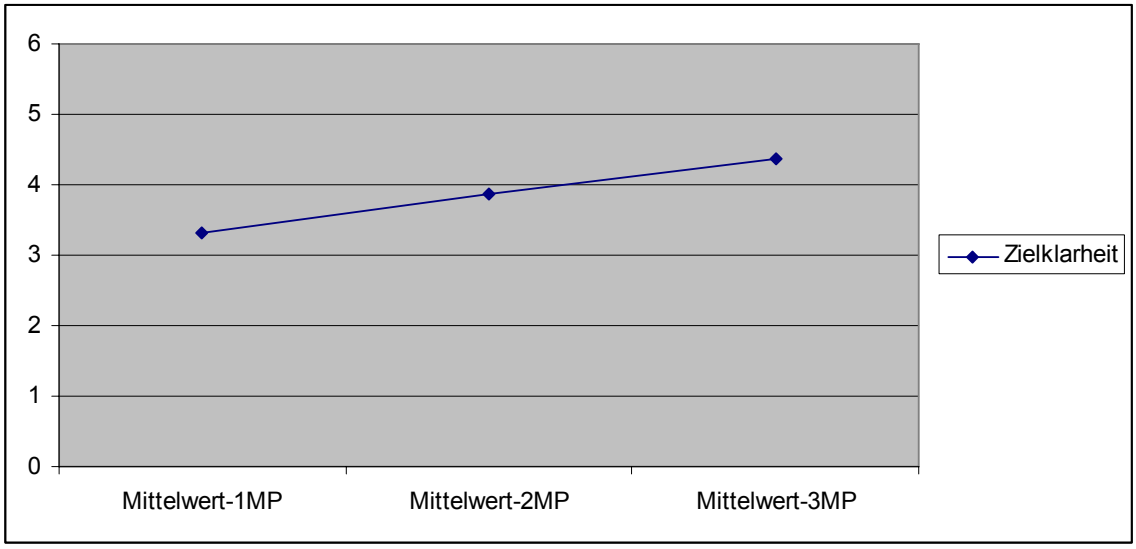

Abbildung 11-16: Zielklarheit

11.6 Überprüfung der Voraussetzung homogener Varianzen

Auch wenn Bortz (2005, S. 276) feststellt, dass die Varianzanalyse bei gleich großen Stichproben relativ robust auf Verletzungen ihrer Voraussetzungen reagiert, soll auf Grund der kleinen Stichprobe eine Überprüfung der Varianzhomogenität durchgeführt werden. Hierfür eignet sich der Levene-Test, der insgesamt den Vorteil aufweist, gegenüber Verletzungen der Normalverteilungsvoraussetzung unempfindlich zu sein. Bei der Durchführung des Levene-Test wird für jeden einzelnen Fall die absolute Abweichung vom Gruppenmittelwert gebildet (Janssen & Laatz 2003, S. 213). Damit wird überprüft, ob sich die Varianz innerhalb der Gruppen von der zwischen den Gruppen signifikant unterscheidet. Die diesbezügliche Überprüfung der Subskalen liefert Signifikanzwerte, die zwischen .3 und .9 liegen, womit die Annahme homogener Varianz weitgehend akzeptiert werden kann.

11.7 Überprüfung der Voraussetzung der Normalverteilung

Viele statistische Verfahren setzen eine Normalverteilung der Daten in der Grundgesamtheit voraus. Auch wenn viele Tests robust auf eine Verletzung reagieren (Janssen & Laatz 2003, S. 207), so ist die entsprechende Überprüfung für den Einsatz parametrischer Testverfahren notwendig. Diese liefern im Vergleich zu nonparametrischen Testverfahren aussagekräftigere Werte.
Abweichungen von der Normalverteilung können im Rahmen gewisser Grenzen hingenommen werden. Große Abweichungen, z.B. mehrgipfelige oder extrem schiefe Verteilungen, können dagegen nicht akzeptiert werden. Eine Überprüfung mit dem Kolmogorov-Smirnov und Shapiro-Wilk-Test (für kleinere Stichproben) kommt zu dem Ergebnis, dass keine Verletzung der Normalverteilung vorliegt. Leider bieten beide Testverfahren nur unscharfe Signifikanzgrenzen für die Überprüfung der Normalverteilung, weil keine eindeutigen Kriterien dafür existieren, wann die

Anwendungsvoraussetzung für einen Test nicht mehr gegeben ist, der eine Normalverteilung voraussetzt (ebd.). Die beiden genannten Testverfahren erkennen aber extrem schiefe Verteilungen und liefern die entsprechenden Signifikanzwerte (nämlich 0 Prozent Signifikanz als Wahrscheinlichkeit dafür, dass von einer Normalverteilung ausgegangen werden kann). Deshalb wird für die eigene Stichprobe davon ausgegangen, dass keine grobe Verletzung der Normalverteilung vorliegt. Hilfreich ist zudem die grafische Visualisierung in Normalverteilungsdiagrammen. Auch hier konnten hinsichtlich der eigenen Daten keine Auffälligkeiten entdeckt werden, die gegen die Annahme einer Normalverteilung sprechen.

11.8 Vergleich der Mittelwerte über die Messzeitpunkte (t-Test)

Die in Kap. 11.5 dargelegten deskriptiven Kennwerte der Stichprobe liefern Anlass zu der Hyothese, dass sich in den verschiedenen Subskalen Unterschiede in der Entwicklung über die Messzeitpunkte ausmachen lassen. Um Differenzen zwischen Dimensionen bzw. Messpunkten zu untersuchen, wurde der t-Test für abhängige Stichproben durchgeführt. Er basiert auf der mathematischen Grundlage, dass die Varianz des ersten Messpunktes die Varianz des x-ten Messpunktes berücksichtigt und umgekehrt (Bortz 2005, S. 141). Tabelle 11-11 legt die berechneten t-Werte für den Vergleich der drei Messpunkte dar:

	t (1.MP-2.MP)	t (2.MP-3.MP)	df
Umgang mit der Unterrichtszeit	-5,391***	-1,69	10
Qualität der Organisation	-3,34**	-0,87	10
Umgang mit Fehlern	-2,66*	-2,17*	10
Motivierungsfähigkeit	-2,77**	-1,06	10
Pacing	-2,65*	-2,32*	10
Lehrperson als Mediator	-2,03*	-2,38*	10
Sprachliche Qualität	-2,67*	-1,57	10
Gesprächsführung	-3,22**	-3,72**	10
Zielklarheit	-2,36*	-1,72	10

Tabelle 11-11: t-Werte der Stichprobe über die Messpunkte; p ≤ 5% *; p ≤ 1% **; p ≤ 0,1% ***

In der Betrachtung der Signifikanzen fällt auf, dass im Verhältnis des ersten und zweiten Messpunktes die Mittelwertvergleiche der Dimensionen durchgehend signifikant ausfallen. Dabei sind drei Subskalen hoch signifikant und eine höchst signifikant. Anders verhält es sich mit dem Vergleich von zweitem und drittem Messpunkt. Hier zeigen sich beim Mittelwertvergleich „nur“ drei Subskalen signifikant und eine hochsignifikant. Für nur zwei Dimensionen liegen die Signifikanzen im Vergleich des zweiten und dritten Messpunktes höher als im Vergleich des ersten und zweiten Messpunktes. Das betrifft die Subskalen „Lehrperson als Mediator“ und „Gesprächsführung“. Umgekehrt fallen die Signifikanzen im Vergleich des ersten und zweiten Messpunktes in sieben Fällen höher aus als im Vergleich des zweiten und dritten Messpunktes. Der höchste Wert betrifft die Subskala „Umgang mit der

Unterrichtszeit“ (1. MP-2. MP), während der niedrigste Wert für die Skala „Qualität der Organisation“ (2. MP-3. MP) vergeben wird.

11.9 Diskussion der Signifikanzwerte über die Messzeitpunkte

Im Vergleich der Entwicklungen zwischen dem ersten Messpunkt (Tagespraktikum) und dem zweiten Messpunkt (Beginn des Praxisjahres) kann eine signifikante Veränderung über alle Skalen beobachtet werden. Folglich sind zwischen dem Tagespraktikum und einer intensiven, wenn auch vergleichsweise kurzen Praxiserfahrung signifikante Unterschiede in einer hoch-inferenten Einschätzung festzustellen.
Dieses Ergebnis mag nicht überraschen. Im Tagespraktikum erfahren viele Studierende das erste Mal den Rollentausch vom Schüler zum Lehrer, was die niedrigeren Einschätzungen zu diesem Messpunkt erklärt. Auch wenn eine stetige Verbesserung der Mittelwerte über alle Dimensionen und Messpunkte festgestellt werden kann, ist allerdings überraschend, dass zwischen dem ersten und zweiten Messpunkt größere Entwicklungen aufscheinen als zwischen dem zweiten und dritten Messpunkt, da man einen stärkeren Effekt mit längerer Verweildauer in der Praxis hätte annehmen können.
Ausnahmen bilden die Dimensionen „Lehrperson als Mediator“ und „Gesprächsführung“. Nimmt man an, dass die zeitliche Verweildauer in Praxislernsituationen ein Indikator für die Kompetenzentwicklung der Studierenden sein kann, so handelt es sich anscheinend um die Dimensionen, die sich durch eine besondere Zeitintensität in der Aneignung auszeichnen. Inhaltlich ist das plausibel, weil die Dimensionen angesprochen sind, die am stärksten auf individuelle Entwicklungen schließen lassen und ein vertieftes didaktisches Verständnis der Lehrerrolle im Umgang mit Schülern erfordern. Davon abgesehen kann man schlussfolgern, dass sich viele der Konstrukte nach relativ kurzer Praxisbegegnung signifikant verbessern.
Eingedenk der Signifikanzen nach einer nur kurzen Praxisbegegnung und der relativ gesehen weniger signifikanten Verbesserung über die Dauer des Praxisjahres ist auf einen starken Sozialisationseffekt zu schließen. Der Praxis der Lehrerbildung gelingt es relativ schnell unterrichtliche Kompetenzen zu vermitteln, die zu einer signifikanten Verbesserung führen. Damit werden den Studierenden Handlungselemente zur Verfügung gestellt, die eine schnelle Routinisierung komplexer Handlungsdimensionen erlauben. Die stärksten Effekte treten in den Subskalen „Umgang mit der Unterrichtszeit“, „Qualität der Organisation“ und „Gesprächsführung“ auf. Damit sind Kompetenzen angesprochen, die gewissermaßen das „Rückrad“ unterrichtlichen Handelns bilden und dafür sorgen, dass der Unterricht sich routinisiert konstituiert. Insofern ist nachvollziehbar, weshalb die Effekte einer Kurzzeitsozialisation hier am stärksten wirken. Am geringsten fallen die Effekte für die Subskalen „Lehrperson als Mediator“ und „Zielklarheit“ aus. Hier sind Kompetenzen angesprochen sind, die sich offenbar weniger effektiv in nur kurzen Praxisbegegnungen entwickeln.
Betrachtet man die Entwicklungen über das Praxisjahr, so sind die geringsten Effekte an den Subskalen „Qualität der Organisation“, „Motivierungsfähigkeit“ und

„Sprachliche Qualität“ auszumachen. Diese nicht signifikanten Entwicklungen lassen sich dahingehend erklären, dass sich die Qualität der Organisation, zumindest auf einer Beobachtungsebene, ab einem gewissen Zeitpunkt nur noch schwerlich verbessern lässt. Für eine differenziertere Betrachtung von Unterschieden wäre eine fachdidaktische Analyse notwendig, eine Perspektive, die diese Studie nicht leistet. Ähnliches gilt für die Sprachqualität, die zudem stark von den Eingangswerten der Studierenden abhängig ist.

Die Motivierungsfähigkeit speist sich wiederum aus Personeneigenschaften, von denen man z.T. eine zeitliche Stabilität erwarten kann. Darüber hinaus kann man die These aufstellen, dass verhältnismäßig klare Muster über einen motivierenden Unterricht bestehen, weshalb extreme Gestaltungsformen, z.B. ein Unterrichtverhalten, das durch ein völliges Desinteresse den Inhalten oder den Schülern gegenüber gekennzeichnet ist bzw. Motivationsversuche, die aus Beobachtungswarte als sehr übertrieben aufgefasst werden müssen, unterbleiben.[94]

In der genaueren Betrachtung der Entwicklung von Dimensionen über die Messzeitpunkte gelangt man zu dem Schluss, dass zu Beginn einer verlängerten Praxisphase die Routinisierung von Kompetenzen schnell gelingt und sich dieser Verlauf über die Zeit verstetigt. Für die weitere Entwicklung über längere Praxisphasen kann man vermuten, dass sich der Schwerpunkt auf eine reflexive Elaborierung der Kompetenzen verschiebt. Damit sind Entwicklungsdimensionen angesprochen, die sich nicht mehr alleine in einer allgemeinpädagogischen Perspektive untersuchen lassen. Eine fachwissenschaftliche oder fachdidaktische längsschnittliche Untersuchung könnte diesbezüglich zu weiterführenden Ergebnissen kommen.

Methodisch kommen Beobachtungsstudien an Grenzen, wenn es um die Untersuchung kognitiv-reflexiver Entwicklungen geht. Auch hoch-inferente Zugänge können im Hinblick auf diese Fragestellung nur mittelbare Ergebnisse liefern. Diese Einschränkung ist bei der Interpretation der Ergebnisse zu bedenken.

[94] Weiteren Aufschluss könnte ein Vergleich mit den Daten aus dem Forschungsprojekt „Standarderreichung in der Lehrerinnen und Lehrerbildung“ (Baer et al. 2007) liefern. In diesem Projekt werden Studierende an Schweizer Pädagogischen Hochschulen bzw. Studierende, die das Lehramtsstudium an der Pädagogischen Hochschule Weingarten absolvieren, mit demselben Instrumentarium untersucht. Zudem wurden neben den Studierenden auch Ausbildungslehrer videographiert, womit interessante Experten-Novizen-Vergleichsmöglichkeiten gegeben sind. Die Daten liegen zum jetzigen Zeitpunkt allerdings noch nicht vollständig vor.

12 Zusammenfassende Betrachtung und Diskussion

Die Daten der niedrig- bzw. hoch-inferenten Beobachtungsverfahren liefern unterschiedliche Ergebnisse, die auf den ersten Blick inkonsistent erscheinen. Während über niedrig-inferente Zugänge keine Veränderungen (bei allerdings unspezifisch streuenden Ergebnissen über die Probanden und im Längsschnitt) über die Messzeitpunkte gefunden werden können, zeugen die hoch-inferenten Verfahren von signifikanten Veränderungen über die Zeit.
Die inhaltlichen Forschungsfragen der Studie (vgl. Kap. 9) über aufzeigbare Entwicklungen (1), eine Signifikanz über die Messzeitpunkte (2) und deren Richtungen (3) lassen sich nur mit einer hoch-inferenten Methodik, die eine Entwicklungsdynamik nachzeichnen kann, beantworten, während eine niedrig-inferente Methodik diesbezüglich eine Stagnation impliziert.
Bezogen auf die hoch-inferent feststellbaren Entwicklungstendenzen fällt auf, dass die Veränderungen zwischen dem ersten allgemeinen Tagespraktikum und einem relativ frühen Zeitpunkt im Praxisjahr größer ausfallen als zwischen Beginn und Ende des Praxisjahres. Die positive Richtung der Entwicklung bleibt über alle Subskalen und Messzeitpunkte erhalten, während die Entwicklungsdynamik erheblich differiert, sowohl was die absoluten Ausprägungen als auch die relativen Veränderungen über die Messzeitpunkte angeht. Auf der Basis dieser Daten sind Rückschlüsse auf den Schwierigkeitsgrad für die Studierenden zur Bewältigung der modellierten Kompetenzkonstrukte möglich, welche sich auch konsistent erklären lassen (vgl. Kap. 11.9).
Wenn auch in der Entwicklung der beiden Instrumente keine explizite wechselseitige Bezogenheit der Konstrukte vorgenommen wurde, so war prinzipiell die Absicht damit verbunden, die Ergebnisse inhaltlich für Validierungszwecke aufeinander beziehen zu können. Durch die unterschiedlichen Entwicklungsverläufe ist dies nur eingeschränkt möglich. Diese Differenzen zwischen den methodischen Beobachtungszugängen bzw. daraus ableitbare Schlussfolgerungen können auf drei Ebenen interpretiert werden:

(a) Methode
(b) längsschnittlicher Verlauf
(c) Rückschlüsse auf Lerngelegenheiten in verlängerten Praxisphasen

Auf einer methodischen Ebene (a) stellt sich die Frage, was von den Beobachtungsverfahren erfasst wird bzw. erfasst werden kann. Die an sich klar anmutende Unterscheidung zwischen Beobachtungsverfahren unterschiedlicher Inferenz erweist sich nur dann als hilfreich, wenn diese Inferenzdifferenzierung auch inhaltliche Äquivalenzen zu dem Beobachtungsfeld aufweist, die in dieser Richtung sinnvoll interpretiert werden können. Problematisch dabei ist allerdings, dass Beobachtungsschwierigkeiten, die weit reichende Schlussfolgerungen für den Rekonstruktionsprozess vom Beobachter verlangen, nicht gleichzeitig immer den Grad der Schwierigkeit abbilden, der bewältigt werden muss, um diese Handlungen auf der Akteursseite zu realisieren. Schwierigkeitsgrad der Beobachtung und der dafür notwendige Grad an Expertise korrespondieren nicht notwendigerweise mit den dahinter liegenden Entwicklungen, die auf Kompetenzen schließen lassen.

Maßgebend ist die Unterscheidung zwischen Schwierigkeiten, die sich aus der Beobachtungsmethodik ergeben – genuin ein methodisches Problem – und Schwierigkeiten, die sich aus der Komplexität der entwickelten Konstrukte erklären, vor allem ein domänenspezifisches Problem. Insbesondere im letzten Fall kann aus dem Inferenzgrad auf dahinter liegende Entwicklungssprünge geschlossen werden. Der Inferenzgrad ist also sowohl von dem einzuschätzenden Beobachtungsgegenstand als auch von der Beobachtungsschwierigkeit beeinflusst. Zudem kann man die Frage aufwerfen, über welche Einschätzungsleistungen der Beobachter verfügen muss, um – wie in der vorliegenden Studie – unterrichtliche Kompetenzen „kompetent“ einschätzen zu können. Muss also der Beobachter kompetenter als der zu Beobachtende in der einzuschätzenden Domäne sein oder muss er zumindest über einen ähnlichen Erfahrungshintergrund verfügen? Beide Fragen lassen sich aus meiner Sicht nicht eindeutig beantworten. Ohne Zweifel ist ein bestimmter domänenspezifischer Expertisegrad eine notwendige Voraussetzung, um Einschätzungen sachadäquat vornehmen zu können. Unterscheidbare Expertiseausprägungen zwischen Beobachtern können aber auch die Basis für systematische Beobachtungsfehler sein, die sich z.B. daraus ergeben, dass unwillkürlich Vergleiche zwischen der zu beobachtenden und der eigenen Person hergestellt werden. Ähnlich verhält es sich mit dem angesprochenen Erfahrungshintergrund. Die Lösung dieser Fragen liegt deshalb meiner Einschätzung nach in einer streng kriteriengeleiteten Herangehensweise, wobei die Kriterien diskursiv bearbeitet werden müssen. In diesem diskursiven Prozess kann es hilfreich sein, wenn die Beteiligten über verschiedene Expertiseausprägungen und Erfahrungshintergründe verfügen, um eine multiperspektivische Annäherung zu ermöglichen.

Die erste genannte Einschränkung, Schwierigkeiten der Beobachtung, bezieht sich sowohl auf niedrig- als auch auf hoch-inferente Beobachtungen. Daraus ergibt sich eine Überbewertung dessen, was beobachtet werden kann, insofern es schwieriger anmutet Rückschlüsse aus Sachverhalten zu ziehen, die nicht beobachtet werden können. Man kann z.B. Zweifel daran hegen, ob ein Unterricht, in dem keine Störungen beobachtet werden können (was je nach Definition prinzipiell einfach, also niedrig-inferent, zu leisten sein sollte), Rückschlüsse auf komplexe Verhaltensweisen zur Prävention von Disziplinschwierigkeiten (prinzipiell eine hoch-inferent zu treffende Entscheidung) erlaubt. An diesem Beispiel soll illustriert werden, welche Herausforderungen mit Beobachtungsstudien verbunden sind. Die Differenzierung von Inferenzstufen liefert auf einer inhaltlichen Ebene vor allem für Extremwerte eindeutige Unterscheidungen. Für andere Beobachtungs-Konstrukte kann die Zuordnung zu einem Inferenzniveau nur annäherungsweise erfolgen. Das Problem einer ungenauen Inferenzzuordnung lässt sich über die Formulierung von Beobachtungsmanualen, in denen genaue Beobachtungskriterien definiert werden, teilweise lösen. Damit wird die Tendenz einer nur ausschnitthaften Berücksichtigung der gesamten Unterrichtskomplexität, ohnehin eine Herausforderung für methodisch geleitete Beobachtungsstudien, weiter verschärft.

Für die zukünftige Methodik von Videostudien erscheint mir eine stärkere Berücksichtigung von hoch-inferenten Beobachtungszugängen (eingedenk der oben vorgenommenen Einschränkungen) viel versprechender als der Einsatz von niedrig-inferenten Zugängen, die häufig nur basale Aussagen über die vorgefundenen

Unterrichtsrealitäten machen können und deshalb im Hinblick auf ihren Aussagewert Ergebnisse liefern, die nur im Rückgriff auf übergreifende Theoriepositionen eingeordnet und interpretiert werden können. Zudem stellt sich die Frage nach der zugrunde gelegten Intervallgröße, so man denn die gesamte Unterrichtsstunde als Zeitstichprobe zu untersuchen beabsichtigt.
Die Wahl der Untersuchungseinheit stellt einen grundlegenden Einflussfaktor für die Analyse niedrig-inferenter Beobachtungen dar. Der Einsatz von Zehn-Sekunden Intervallen überzeugt auf den ersten Blick, weil damit eine Untersuchungseinheit angezeigt wird, die wegen ihrer relativen Kürze einen Analysezugang zu komplexen Unterrichtsmustern verspricht, während die Intervallgröße immerhin so angelegt ist, dass aussagekräftige Entscheidungen getroffen werden können, wenn man ein disjunktes Kodiersystem zugrunde legt.
Je kleiner die Untersuchungseinheit gewählt wird, desto größer fallen tendenziell Einflussfaktoren in das Gewicht, die nicht unstrittig als Ergebnis der Aufnahmesituation oder im Fall der eigenen Studie als Ergebnis der besonderen Situation von Praktikanten im Feld Schule bewertet werden können. Auch wenn die eigenen Bemühungen nicht aussagekräftig darauf verweisen, so bringt der Einsatz einer verhältnismäßig kleinen Untersuchungseinheit nicht nur Vorteile mit sich. Vermutlich bedarf es, in Abhängigkeit der ausgewählten Konstrukte und der dahinter stehenden theoretischen Prämissen, jeweils spezifischer Intervallgrößen.
Für die Präferenz hoch-inferenter Beobachtungszugänge muss allerdings auch der mit Videostudien verbundene hohe Aufwand angesprochen werden, der mit einer solchen Herangehensweise noch ansteigt, speziell dann, wenn man dem von Wirtz & Caspar (2002) vorgetragenen Argument folgt, mit hoch-inferenten Videostudien verbundenen Reliabilitätsproblemen über die Mittelung der Ergebnisse von mehreren Ratern entgegenzuwirken. Hier könnte ein Ansatz liegen, um das von Clausen, Reusser & Klieme (2003, S. 125) angesprochene Übereinstimmungs-Bedeutsamkeits-Dilemma im Bezug auf das Verhältnis von niedrig- und hoch-inferenten Beobachtungsverfahren anzugehen.

Betrachtet man den längsschnittlichen Verlauf der Daten über beide methodische Zugänge (b), so kann daraus geschlussfolgert werden, dass basale Choreographien des Unterrichts sich auf einer Sichtebene sehr stark ähneln können und Qualitätseinschätzungen trotzdem zu unterschiedlichen Urteilen kommen. Damit ist ein bekannter Sachverhalt benannt, nach dem man Dinge auf sehr ähnliche Weise, aber trotzdem unterschiedlich gut angehen kann. So gesehen ist die differenzierende Ergebnislage niedrig- und hoch-inferenter Beobachtungsverfahren inhaltlich gut interpretierbar. Dies gilt umso mehr, wenn man berücksichtigt, dass in der Entwicklung des Rater-Inventars Kompetenzdimensionen berücksichtigt wurden, die in Anlehnung an das Angebot-Nutzen-Modell von Helmke (2003) die Schaffung von Lerngelegenheiten thematisieren und damit auf einer vorwiegend hoch-inferent einzuschätzenden Latenzdimension angesiedelt sind.
Die auf einer niedrig-inferenten Beobachtung basierende weitgehende Stabilität der Ergebnisse kann auf der Ebene einer subjektbezogenen Rekonstruktion dahingehend gedeutet werden, dass im Rahmen und als Ergebnis der Praxisbegegnungen als Lerngelegenheiten keine Modifizierung der subjektiven Theorien der Studierenden als handlungsleitende Kognitionen erfolgte, zumindest

was die makrodidaktischen Strukturen angeht. Zudem kann man vermuten, dass in die Unterrichtsgestaltung der Studierenden Einflüsse eingegangen sind, die mit ihren eigenen schulischen Erfahrungen korrespondieren. Damit lässt sich der teilweise recht konservativ anmutende Unterrichtsstil mit einem Schwerpunkt auf das fragend entwickelnde Unterrichtsgespräch und der Bearbeitung von Aufgabenblättern durch die Schüler nachvollziehen. Gleiches gilt für das fast vollständige Fehlen von Differenzierungen, Strukturierungen und offenen Lehr-Lernformen, wenn man auch einräumen muss, dass bezüglich der Offenheit des Unterrichts ein Einfluss durch das Aufnahmesetting anzunehmen ist.

Nimmt man dagegen eine Perspektive ein, Auswirkungen der Sozialisationsinstanz Schule als Feld der Praxisbegegnung zu rekonstruieren, so wirken Praxisphasen bzw. ihre Verlängerung nicht per se als Korrektiv zu im erziehungswissenschaftlichen Diskurs für überholt geltenden Lehr-Lernformen. Mehr Praxis bedeutet nicht per se eine Maximierung von Lernmöglichkeiten oder deren Nutzung. Wie viele diesbezügliche Lernangebote notwendig sind, lässt sich nicht generalisierend beantworten. Die Daten der Studie lassen sich in eine Richtung deuten, dass grundlegende Unterrichtstechniken auf einer Inszenierungsebene schnell erlernbar sind oder am Modell gelernt werden können, während übergreifende didaktische Strategien, die wesentlich an persönliche Lerneinsichten gekoppelt sind, längere Zeiträume benötigen, um überzeitlich habitualisiert werden zu können. Hier kann ein Vorteil von verlängerten Praxisphasen innerhalb des Studiums liegen, die die Möglichkeit liefern, Handlungsnotwendigkeiten der praktischen Berufsausübung kennenzulernen und vor dem Hintergrund der erziehungswissenschaftlichen Fundierung im weiteren Verlauf des Studiums bewerten und revidieren zu können (wobei hier nicht die Frage geklärt werden kann, ob die Studierenden diese Chance auch nutzen).

Wertet man die dokumentierten Unterrichtsstunden der Studierenden u.a. unter der Annahme des Lernens am Modell, so lassen sich aus dem längsschnittlichen Verlauf vorsichtig Rückschlüsse über die Lerngelegenheiten in verlängerten Praxisphasen an den betroffenen Schulen ziehen. Das weitgehende Fehlen von metakognitiven Adressierungen und Differenzierungen bzw. die Dominanz des fragend entwickelnden Unterrichtsgesprächs und der Bearbeitung von Aufgabenblättern lassen die These zu, dass ein solcher Unterrichtsstil auch nicht selten bei den betreuenden Mentoren anzutreffen ist, obwohl er nicht dem Stand des gegenwärtigen erziehungswissenschaftlichen Diskurses entspricht, in dem individualisierende und auf Selbststeuerung abzielende Unterrichtsstrategien stark präferiert werden. Dieses Ergebnis deutet auf eine mögliche Differenz zwischen der erziehungswissenschaftlichen Theoriebildung an den Hochschulen und der Unterrichtspraxis an deutschen Schulen hin.

Zumindest decken sich die Ergebnisse niedrig-inferenter Kodierungen weitgehend mit den Ergebnissen von Hoeltje et al. (2002, 2004), die durch schriftliche Befragungen von Studierenden im Anschluss an ein Halbjahrespraktikum die Dominanz lehrerzentrierter Lehr-Lernformen feststellten. Hier scheint das viel zitierte Theorie-Praxis-Problem (vgl. Kap. 6) durch, das unter den Aspekten der strukturellen Unähnlichkeit der auf die Konstituierung von Theorie bzw. Praxis bezogenen Logiken und dementsprechender Realisierungsmöglichkeiten in der Koppelung von Theorie und Praxis diskutiert wird.

Wenn es auch gute Gründe dafür gibt, den von Dann et al. (1978) thematisierten Praxisschock und sich daraus ergebende, vermeintlich zwingend erscheinende Rigiditäten als Notwendigkeiten für die Bewältigung von Anforderungen der Praxis nicht als selbstverständliche Bestandteile der Lehrerbildung anzusehen, sondern eher als Herausforderung für flankierende Beratungsangebote (Larcher-Klee 2005), so weisen die Ergebnisse in eine Richtung, die eine eher konservativ geprägte Berufssozialisation durch verlängerte Praxislernphasen vermuten lassen. Der längsschnittliche Verlauf der Datenbasis liefert somit zumindest mittelbare Rückschlussmöglichkeiten auf vorfindbare Praxislernangebote (c).

Die methodischen Fragestellungen der Arbeit nach einer für Validierungszwecke wechselseitigen Bezugnahme der gewählten Beobachtungsmethodiken (4) oder nach einer evtl. differierenden Ergebnislage durch diese (5) (vgl. Kap. 9) führen zu Antworten, die als Impuls für eine veränderte Forschungsmethodik gewertet werden können. Eine genuin methodische Bezugnahme der beiden Verfahren wurde in dieser Studie nicht geleistet, weil es beabsichtigt war, die aufgenommenen Stunden möglichst umfassend im Sinne der Kompetenzthematik zu betrachten und nicht vorschnell eine Einschränkung auf spezifische Fragen z.B. der Methodik vorzunehmen. Für den niedrig-inferenten Bereich wurden viele Konstrukte getestet, die aber nach negativen Reliabilitätsüberprüfungen nicht zum Einsatz kamen. Übrig blieben damit Konstrukte, die eher übergreifend die Unterrichtsrealität untersuchen, womit ein weiteres Argument geliefert ist, warum diesbezüglich keine Entwicklungen festgestellt werden können.
Für zukünftige Vorhaben kann es sinnvoll sein niedrig-inferente in den Dienst hoch-inferenter Verfahren zu stellen, um spezifische Unterrichtsanteile zu identifizieren und hoch-inferent näher zu analysieren. Weiterhin ist es prinzipiell von Interesse Mikrostrukturen des Unterrichts in den Blick zu nehmen und niedrig-inferent als Zeit- oder Handlungsstichprobe zu kodieren. Aus eigener Erfahrung kann man dabei aber Reliabilitätsprobleme nicht ausschließen, womit sich die Frage stellt, ob das Inferenzniveau dann tatsächlich noch als niedrig bezeichnet werden kann oder ob es sich mit Blick auf Beobachtungsschwierigkeiten nicht doch um eine hoch-inferente Methodik handelt. Des Weiteren gewinnen bei einer mikrostrukturellen Untersuchung konversationsanalytische Aspekte an Bedeutung, so dass man nicht mehr unbedingt von Beobachtungsverfahren, die wesentlich auf eine Videoinformation angewiesen sind, sprechen kann.
Die Frage der Studie (5), in welche Richtung Entwicklungen über die Methodik aufgehellt werden und ob Unterschiede in den längsschnittlichen Ergebnissen der Beobachtungsverfahren gezeigt werden können, ist bereits oben beantwortet worden. Ursächlich für diese Fragestellung war die Überlegung, dass Entwicklungslinien (so denn auffindbar) unter Umständen nicht linear verlaufen. Sie konnte im hoch-inferenten Zugang nicht bestätigt werden. Vielmehr lässt sich – ohne Ausnahme – eine lineare Entwicklung über die Daten aufzeigen, wobei allerdings die Entwicklungsdynamik differiert. Ein so klares Ergebnis war nicht vermutet worden, genau so wenig wie die klare Stabilität der Ergebnisse im niedrig-inferenten Modus.

Abschließend muss die Frage beantwortet werden, inwieweit Videostudien der nach wie vor aktuellen Kompetenzthematik und den damit konnotierten

Begriffsdimensionen gerecht werden können. In der vorliegenden Arbeit wurde im theoretischen Teil der Kompetenzthematik viel Raum gewährt, weil damit gegenwärtig eine Begriffsfolie vorliegt, über die ein breiter Konsens besteht, wenn die Modellierung von Leistungsunterschieden empirisch angestrebt wird. Die Kompetenzdebatte ist mittlerweile derart ausdifferenziert, dass ein methodischer Zugang alleine kaum mehr ausreicht, um dem Begriff gerecht zu werden und ihn sinnvoll verwenden zu können. Videostudien zeichnen sich dadurch aus, dass sie den Prozesscharakter des Unterrichts ansatzweise abbilden und damit im Unterschied zu anderen methodischen Verfahren eine Ganzheitlichkeit in der Erhebung versprechen, die jedoch nicht in Gänze eingelöst werden kann.
Nimmt man die von Weinert (2001) vorgeschlagene Definition als Referenz, wie es mittlerweile üblich geworden ist, so stechen vier Dimensionen hervor:

- Kognitive Dimension
- Handlungsdimension
- Entscheidungsdimension
- Soziale Dimension

Videostudien legen den Fokus auf die Handlungsdimension, mit der automatisch auch die Entscheidungsdimension erfasst wird, denn die Akteure präsentieren das Ergebnis von Entscheidungsprozessen, die im Vorfeld abgelaufen sind. Gleiches gilt für die Konstituierung von Kompetenzen im sozialen Prozess, wenn auch Interaktionsprozesse für Videostudien eine Herausforderung darstellen. Man kann argumentieren, dass sich die kognitive Dimension ebenfalls auf der Handlungsebene erfassen lässt, immerhin ist sie das Ergebnis vorangegangener Planungsüberlegungen und deren Adaption im sozialen Prozess, wobei auch hier immer wieder situative Entscheidungen notwendig sind. In diesem Punkt liegt meiner Einschätzung nach aber das größte Gegenargument für die Verwendung des Kompetenzbegriffes im Rahmen von Videostudien. Kognitive Prozesse lassen sich am besten erfassen, indem man Probanden die Möglichkeit gibt, über Alternativen zu reflektieren und Begründungen dafür zu liefern, warum bestimmte Entscheidungen getroffen worden sind. Dieser Herausforderung werden Videostudien nicht gerecht, wenn sie sich darauf beschränken, die Unterrichtsrealitäten lediglich abzubilden und über verschiedene Beobachtungszugänge zu bewerten.
Deshalb sollten Videostudien durch andere Zugänge ergänzt werden. Innerhalb der Methodik von Videostudien liegt es nahe mit dem Material zu arbeiten und über die Methode des „stimulated recall“ die Probanden mit dem Videomaterial zu konfrontieren, um Begründungen oder Alternativoptionen des gezeigten Handelns darlegen zu können. Weitere Möglichkeiten liegen in Interviewverfahren oder in verschriftlichten Problemdarstellungen, die Abwägungsprozesse transparent machen. Das Gesamtdesign des Projektes trägt dieser Problematik Rechnung (vgl. Dörr & Küster 2006). Die vorliegende Studie rekurriert aber ausschließlich auf die Daten der Videostudie, um Grenzen und Möglichkeiten dieser Methodik auszuloten. Damit war beabsichtigt die Erkenntnismöglichkeiten, die sich aus dem Einsatz von Videostudien vor dem Hintergrund der Kompetenzdiskussion ergeben, darzustellen.

Literatur

Abell foundation (2001a): Teacher Certification Reconsidered. Stumbling for Quality. Baltimore: Abell. Verfügbar unter URL: http://www.abell.org/pubsitems/ed_cert_1101.pdf (Zugriff am 25.08.2006).

Abell foundation (2001b): Teacher Certification Reconsidered. Stumbling for Quality - Appendix. Review of Research Teacher Certification and Effective Teaching. Baltimore: Abell. Verfügbar unter URL: http://www.abell.org/pubsitems/ed_cert_appendix_1101.pdf (Zugriff am 25.08.2006).

Abs, H.J. (2006): Zur Bildung diagnostischer Kompetenz in der zweiten Phase der Lehrerbildung. In: Zeitschrift für Pädagogik, 51. Beiheft, S. 217-234.

Adorno, T.W. (1962): Theorie der Halbbildung. In: Horkheimer, M./Adorno, T.W.: Sociologica II. Frankfurt/M.: Suhrkamp, S. 168-192.

Aebli, H. (1977): Grundformen des Lehrens. Eine allgemeine Didaktik auf kognitionspsychologischer Grundlage. Stuttgart: Klett.

Altrichter, H. (2000): Handlung und Reflexion bei Donald Schön. In: Neuweg, G.H. (Hrsg.): Wissen – Können – Reflexion. Innsbruck: Studienverlag, S. 201-221.

Altrichter H./Hascher, T. (2005): Editorial. In: Journal für Lehrerinnen- und Lehrerbildung, 5 (1), S. 4-7.

Altrichter, H./Posch, P. (1998): Lehrer erforschen ihren Unterricht. Eine Einführung in die Methoden der Aktionsforschung. Bad Heilbrunn: Klinkhardt.

Amelang, M./Bartussek, D. (2001): Differenzielle Psychologie und Persönlichkeitsforschung. Standards Psychologie. Stuttgart: Kohlhammer.

Anderson, J.R. (1982): Acquisition of Cognitive Skill. In: Psychological Review, 89, pp. 369-406.

Anderson, J.R. (1983). The Architecture of Cognition. Cambridge, MA: Harvard University Press.

Anderson, J.R. (1987): Skill Acquisition: Compilation of Weak-Method Problem Solutions. In: Psychological Review, 94, pp. 192-210.

Anderson, L.W./Ryan, D.W./Shapiro, B.J. (Eds.) (1989): The IEA Classroom Environment Study. Oxford: Pergamon Press.

Arnim, F. (2000): Kompetenzorientierung der Lehrerbildung. In: Bayer, M./Bohnsack, F./Koch-Priewe, B./Wildt, J. (Hrsg.): Lehrerin und Lehrer werden ohne Kompetenz? Bad Heilbrunn: Klinkhardt, S. 302-315.

Arnold, K.-H./Sandfuchs, U./Wiechmann, J. (Hrsg.) (2006): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt.

Arnold, R. (2007): Grundlinien einer Ermöglichungsdidaktik. Ziel: Augsburg.

Arnold, R./Müller, H.-J. (2006): Das Konzept der Schlüsselqualifikationen – nicht länger ein „Fliegender Holländer“ im Meer des Wissens? In: Dies. (Hrsg.): Kompetenzentwicklung durch Schlüsselqualifikations-Förderung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 7-13.

Ausubel, D.P./Novak, J.D./Hanesian, H. (1980): Psychologie des Unterrichts, Bd.1. Weinheim: Beltz.

Baacke, D. (1980): Kommunikation und Kompetenz. Grundlegung einer Didaktik der Kommunikation und ihrer Medien. München: Juventa.

Baer, M./Dörr, G./Fraefel, U./Kocher, M./Küster, O./Larcher, S./Müller, P./Sempert, W./Wyss, C. (2006): Standarderreichung in der Lehrerinnen- und Lehrerbildung: Analyse der Wirksamkeit der berufsorientierten Ausbildung. In: Seifried, J./Abel, J. (Hrsg.): Empirische Lehrerbildungsforschung. Münster: Waxmann, S. 141-160.

Baer, M./Dörr, G./Fraefel, U./Kocher, M./Küster, O./Larcher, S./Müller, P./Sempert, W./Wyss, C. (2007): Werden angehende Lehrpersonen kompetenter? – Kompetenzaufbau und Standarderreichung in der berufswissenschaftlichen Ausbildung an drei Pädagogischen Hochschulen in der Schweiz und in Deutschland. In: Unterrichtswissenschaft, 35 (1), S. 15-47.

Ballstaedt, S.P. (1997): Wissensvermittlung. Die Gestaltung von Lernmaterial. Weinheim: Beltz.

Barthold, R./Küster, O./Müller, K. et al. (2006): KOPRA – Kompetenzentwicklung im Praxisjahr. Wirksamkeit des Modellversuches „Praxisjahr Biberach“. In: Hilligus, A.H./Rinkens, H.-D. (Hrsg.): Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 383-389.

Bartussek, D. (1996): Faktorenanalytische Gesamtsysteme der Persönlichkeit. In: Birbaumer, N./Frey, D./Kuhl, J. /Prinz, W./Weinert, F.E. (Hrsg.): Enzyklopädie der Psychologie, Themenbereich C, Serie VIII, Bd. 3: Temperaments- und Persönlichkeitsunterschiede. Göttingen: Hogrefe, S. 51-105.

Bastian, J./Helsper, W. (2000): Professionalisierung im Lehrberuf – Bilanzierung und Perspektiven. In: Bastian, J./Helsper, W./Reh, S./Schelle, C. (Hrsg.): Professionalisierung im Lehrerberuf. Opladen: Leske + Budrich.

Bastian, J./Helsper, W./Reh, S./Schelle, C. (Hrsg.) (2000): Professionalisierung im Lehrerberuf. Opladen: Leske + Budrich.

Bätz, U./Oser, F. (1997): Evaluation Orientierungsschule Basel, Zwischenbericht. Freiburg: Pädagogisches Institut der Universität.

Bauer, K.-O. (2000): Pädagoge – Profession oder Nebenbeschäftigung? In: Jaumann-Graumann, O./Köhnlein, W. (Hrsg.): Lehrerprofessionalität – Lehrerprofessionalisierung. Jahrbuch Grundschulforschung, Bd. 3. Bad Heilbrunn/Obb.: Klinkhardt, S. 25-44.

Bauer, K.-O./Kopka, A./Brindt, S. (1999): Pädagogische Professionalität und Lehrerarbeit. Weinheim: Juventa.

Bauman, Z. (2000): Liquid Modernity. Cambridge: Polity.

Baumert, J. (1998): Physikleistungen am Ende der Schullaufbahn im inter- und intranationalen Vergleich. In: Didaktik und Physik. Regensburg: Deutsche Physikalische Gesellschaft, S. 22-32.

Baumert, J./Bos, W./Lehmann, R.H. (Hrsg.) (2000a): TIMMS/III. Dritte Internationale Mathematik- und Naturwissenschaftstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe, Bd. 2). Opladen: Leske + Budrich.

Baumert, J./Bos, W./Lehmann, R.H. (Hrsg.) (2000b): TIMMS/III. Dritte Internationale Mathematik- und Naturwissenschaftstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Mathematische und physikalische Kompetenzen am Ende der Pflichtschulzeit, Bd. 1). Opladen: Leske + Budrich.

Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K./Weiß, M. (Hrsg.) (2001): Pisa 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.

Baumert, J./ Köller. O. (1998): Nationale und internationale Schulleistungsstudien. Was können sie leisten, wo sind ihre Grenzen? In: Pädagogik, 50 (6), S. 12-18.

Baumert, J./Köller, O. (2000): Unterrichtsgestaltung, verständnisvolles Lernen und multiple Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): TIMMS/III. Dritte Internationale Mathematik- und Naturwissenschaftstudie. Mathematisch und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Opladen: Leske + Budrich, S. 271-315.

Baumert, J./Kunter, M. (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft, 9 (4), S. 469-519.

Baumert, J./Lehmann, R./Lehrke, M./Schmitz, B./Clausen, M./Hosenfeld, I./Köller, O./Neubrand, J. (1997): TIMMS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.

Baumert, J./Prenzel, M. (2002): Einleitung zum Themenheft PISA. In: Unterrichtswissenschaft, 30 (2), S. 98-99.

Bayer, M./Carle, U./Wildt, J. (Hrsg). (2001): Brennpunkt: Lehrerbildung. Strukturwandel und Innovation im europäischen Kontext. Opladen: Leske + Budrich.

Beck, E./Guldimann, T./Zutavern, M. (1994): Eigenständiges Lernen verstehen und fördern. In: Reusser, K./Reusser-Weyeneth, M. (Hrsg.): Verstehen: Psychologischer Prozess und didaktische Aufgabe. Bern: Huber, S. 207-225.

Beck, B./Klieme, E. (Hrsg.) (2007): Sprachliche Kompetenzen. Konzepte und Messung. Weinheim: Beltz.

Benner, D. (2002): Die Struktur der Allgemeinbildung im Kerncurriculum moderner Bildungssysteme. Ein Vorschlag zur bildungstheoretischen Rahmung von PISA. In: Zeitschrift für Pädagogik, 48 (1), S. 68-90.

Benner, D. (Hrsg.) (2007): Bildungsstandards. Paderborn: Schöningh.

Benner, D./Brüggen, F. (2004): Bildsamkeit/Bildung. In: Benner, D./Oelkers, J. (Hrsg.): Historisches Wörterbuch der Pädagogik. Weinheim: Beltz, S. 174-215.

Berliner, D.C. (1992): The Nature of Expertise and Teaching. In: Oser, F.K./Dick, A./Patry, J.L. (Eds.): Effective and Responsible Teaching. San Francisco: Jossey-Bass, pp. 127-248.

Bischoff, S./Brühwiler, C./Baer, M. (2005): Videotest zur Erfassung adaptiver Lehrkompetenz. In: Beiträge zur Lehrerbildung, 23 (3), S. 382-397.

Blömeke, S. (2004): Empirische Befunde zur Wirksamkeit der Lehrerbildung. In: Blömeke, S./Reinhold, P./Tulodziecki, G./Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn/Obb.: Klinkhardt, S. 59-92.

Blömeke, S. (2006a): „Fast fish – Loose Fish“: International-vergleichende Forschung zur Wirksamkeit der Lehrerausbildung. In: Hilligus, A.H./Rinkens, H.-D.: Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 189-214.

Blömeke, S. (2006b): KMK-Standards für die LehrerInnenbildung in Deutschland. Ein Kommentar. In: Journal für LehrerInnenbildung, 8 (1), S. 25-33.

Blömeke, S. (2007): Qualitativ – quantitativ; induktiv – deduktiv, Prozess – Produkt, national – international. In: Lüders, M./Wissinger, J. (Hrsg.): Forschung zur Lehrerbildung. Münster: Waxmann, S.13-36.

Blömeke, S./ Eichler, D./Müller, C. (2003): Rekonstruktion kognitiver Prozesse von Lehrpersonen als Herausforderung für die empirische Unterrichtsforschung. Theoretische und methodische Überlegungen zu Chancen und Grenzen von Videostudien. In: Unterrichtswissenschaft, 31 (2), S. 103-121.

Blömeke, S./Müller, C./Eichler, D. (2005): Handlungsmuster von Lehrerinnen und Lehrern beim Einsatz neuer Medien. Grundlagen eines Projektes zur empirischen Unterrichtsforschung. In: Jahrbuch Medienpädagogik 4. Wiesbaden: Verlag für Sozialwissenschaften, S. 229-254.

Bodensohn, R./Frey, A./Balzer, L. (2004): Diagnose und Rückmeldung von Handlungskompetenzen Studierender im Blockpraktikum – das Projekt VERBAL. In: Journal für LehrerInnenbildung, 4 (1), S. 30-36.

Bodensohn, R./Schneider, C. (2006): Weiterentwicklung der Evaluationskultur Schulpraktischer Studien. In: Hilligus, A.H./Rinkens, H.-D.: Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 261-274.

Bodensohn, R./Schneider, C./Jäger, R.S. (2007): Studierende drängen in das Lehramt! Haben wir Anlass zu Kompetenzbeobachtung und Studierendenauswahl? Zentrum für empirische pädagogische Forschung: Universität Koblenz-Landau.

Bohl, T. (2004): Empirische Unterrichtsforschung und Allgemeine Didaktik. Ein prekäres Spannungsverhältnis und Konsequenzen im Kontext der PISA-Studie. In: Die Deutsche Schule, 96 (4), S. 414-425.

Böhle, F. (2005): Erfahrungswissen – die verborgenen Seiten professionellen Handelns. Eine Herausforderung für die berufliche Bildung. Vortrag auf der Fachtagung des Bundesinstituts für Berufsbildung am 14. Juni 2005 in Bonn.

Böhle, F. (2006): Typologie und strukturelle Probleme von Interaktionsarbeit. In: Böhle, F./Glaser, J. (Hrsg.): Arbeit in der Interaktion – Interaktion als Arbeitsorganisation und Interaktionsarbeit in der Dienstleistung. Wiesbaden: Verlag für Sozialwissenschaften, S. 325-347.

Böhle, F. (2007): Die Bewältigung der Unplanbarkeit durch erfahrungsgeleitetes Arbeiten. In: Tomaschek, N. (Hrsg.): Perspektiven systemischer Entwicklung und Beratung von Organisationen – ein Sammelband. Heidelberg: Carl Auer, S. 89-94.

Bohnsack, R. (2000): Rekonstruktive Sozialforschung – Einführung in Methodologie und Praxis qualitativer Forschung. Opladen: Leske + Budrich.

Bortz, J. (2005): Statistik für Human- und Sozialwissenschaftler. Berlin: Springer.

Bortz, J./Döring, N. (2006): Forschungsmethoden und Evaluation. Berlin: Springer.

Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Walter, G./Valtin, R. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.

Bos, W./Tarnai, C. (1999): Content Analysis in Empirical Social Research. In: International Journal of Educational Research, 21, pp. 659-671.

Bourdieu, P. (1994): Zur Soziologie der symbolischen Formen. Frankfurt/M.: Suhrkamp.

Bräu, K. (2006): Gesprächsanalytische Untersuchung der Lehrer-Schüler-Kommunikation bei der Betreuung individualisierten Lernens. In: Mammes, I./Rahm, S./Schratz, M. (Hrsg.): Schulpädagogische Forschung. Innsbruck: Studienverlag, S. 15-25.

Bräuer, G. (2006): Reflexive Praxis in der Lehrerausbildung – Portfolio als Grundlage für die Aneignung von Kompetenzen. In: Hilligus, A.H./Rinkens, H.-D.: Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 343-349.

Brezinka, W. (1971): Von der Pädagogik zur Erziehungswissenschaft. Weinheim: Beltz.

Brockhaus (2002): Leipzig: Brockhaus.

Brödel, R. (1998): Lebenslanges Lernen – lebensbegleitende Bildung. Neuwied: Luchterhand.

Brödel, R. (2002): Relationierungen zur Kompetenzdebatte. In: Literatur- und Forschungsreport Weiterbildung, Nr. 49, S. 39-47.

Bromme, R. (1992): Der Lehrer als Experte. Zur Psychologie des professionellen Wissens. Göttingen: Hogrefe.

Bromme, R. (1997): Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In: Enzyklopädie der Psychologie. D. I. 3. Göttingen: Hogrefe, S. 178-212.

Brophy, J.E. (2000): Teaching (Educational Practice Series, Vol. 1). Brussels: International Academy of Education & International Bureau of Education.

Brophy, J.E./Good, T.L. (1986): Teacher Behavior and Student Achievement. In: Wittrock, M. (Ed.): Handbook of Research on Teaching. New York: Mac Millan, pp. 328-375.

Brunner, H./Schratz, M. (2002): Was wissen wir über die Wirksamkeit von Offenem Unterricht? Evaluation von Unterrichtsqualität als zentrales Thema von Lehrerausbildung. In: Brunner, H./Mayr, E./Schratz, M./Wieser, I. (Hrsg.): Lehrerinnen- und Lehrerbildung braucht Qualität. Und wie!? Innsbruck: Studienverlag, S. 321-341.

Brunner, M./Krauss, S./Kunter, M./Baumert, J./Blum, W./Neubrand, M. (2006): Lassen sich das Fachwissen und das fachdidaktische Wissen von Lehrkräften empirisch trennen? Eine Untersuchung der Dimensionalität des Professionswissens von Lehrkräften. Vortrag auf der 3. Göttinger Fachtagung des ZEUS am 05. September 2006, Göttingen.

Brunner, M./Kunter, M./Krauss, S./Baumert, J./Blum, W./Dubberke, T./Jordan, A./Klusmann, U./Tsai, Y.-M./Neubrand, M. (2006): Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? In: Zeitschrift für Erziehungswissenschaft, 9 (4), S. 521-544.

Carroll, J.B. (1985): The Model of School Learning: Progress of an Idea. In: Fischer, C.W./Berliner, D. (Eds.): Perspectives on Instructional Time. London: Longman, pp. 29-58.

Chi, M.T.H. (1978): Knowledge Structure and Memory Development. In: Siegler, R.S. (Ed.): Children's Thinking: What Develops? Hillsdale, NJ: Erlbaum, pp. 73-96.

Chi, M.T.H./Feltovich, P.J./Glaser, R. (1981): Categorization and Representation of Physics Problems by Experts and Novices. In: Cognitive Science, 5, pp. 121-152.

Chi, M.T.H./Glaser, R./ Farr, M.J. (Eds.) (1988): The Nature of Expertise. Hillsdale, N.J.: Erlbaum.

Chi, M.T.H./Glaser, R./Rees, E. (1982): Expertise in Problem Solving. In: Sternberg, R.J. (Ed.): Advances in the Psychology of Human Intelligence, Vol 1. Hillsdale, NJ: Erlbaum, pp. 7-75.

Chomsky, N. (1970): Aspekte der Syntaxtheorie. Berlin: Akademie-Verlag.

Ciompi, L. (1988): Außenwelt – Innenwelt. Die Entstehung von Zeit, Raum und psychischen Strukturen. Göttingen: Vandenhoeck und Ruprecht.

Ciompi, L. (1997): Die emotionalen Grundlagen des Denkens. Entwurf einer fraktalen Affektlogik. Göttingen: Vandenhoeck und Ruprecht.

Clausen, M. (2002): Qualität von Unterricht – Eine Frage der Perspektive? Münster: Waxmann.

Clausen, M./Reusser, K./Klieme, E. (2003): Unterrichtsqualität auf der Basis hoch-inferenter Unterrichtsbeurteilungen. In: Unterrichtswissenschaft, 31 (2), S. 122-141.

Cognition and Technology Group at Vanderbilt (1993a): Toward Integrated Curricula: Possibilities from Anchored Instruction. In: Rabinowitz, M. (Ed.): Cognitive Science Foundation of Instruction. Hillsdale, NJ: Erlbaum, pp. 33-55.

Cognition and Technology Group at Vanderbilt (1993b): Anchored Instruction and Situated Cognition Revisited. In: Educational Technology, 3, pp. 52-70.

Cognition and Technology Group at Vanderbilt (1996): Multimedia Environments for Enhancing Learning in Mathematics. In: Vosniadou, S./De Corte, E./Glaser, R./Mandl, H. (Eds.): International Perspectives on the Design of Technology-Supported Learning Environments. Mahwah, NJ: Erlbaum, pp. 285-305.

Collins, A./Brown, J.S./Newman, S.E. (1989): Cognitive Apprenticeship: Teaching the Crafts of Reading, Writing and Mathematics. In: Resnick, L.B. (Ed.): Knowing, Learning and Instructions: Essays in Honour of Robert Glaser. Hillsdale, NJ: Erlbaum, S. 453-494.

Combe, A./Helsper, W. (2002): Professionalität. In: Otto, H.U./Rauschenbach, T./Vogel, P. (Hrsg.): Erziehungswissenschaft: Professionalität und Kompetenz. Opladen: Leske + Budrich, S. 29-47.

Combe, A./Kolbe, F.-U. (2004): In: Helsper, W./Böhme, J. (Hrsg.): Handbuch der Schulforschung. Wiesbaden: Verlag für Sozialwissenschaften, S. 833-851.

Criblez, L. (2001): Die Ausbildnerinnen und Ausbildner – Hauptakteure der Lehrerbildung. In: Oser, F./Oelkers, J. (Hrsg.): Die Wirksamkeit der Lehrerbildungssysteme. Zürich: Rüegger, S. 437-494.

Dann, H.-D./Cloetta, B./Müller-Fohrbrodt, G./Helmrich, R. (1978): Umweltbedingungen innovativer Kompetenz: Eine Längsschnittuntersuchung zur Sozialisation von Lehrern in Ausbildung und Beruf. Stuttgart: Klett-Cotta.

Darling-Hammond, L. (2000): Teacher Quality and Student Achievement. A Review of State Policy Evidence. In: Education Police Analysis Archives, 8 (1). Verfügbar unter URL: http://epaa.asu.edu/epaa/v8n1/ (Zugriff am 14.03.2006).

Darling Hammond, L. (2006): Powerful Teacher Education. San Francisco: Jossey-Bass.

Deci, E.L./Ryan, R.M. (1985): Intrinsic Motivation and Self-Determination in Human Behaviour. New York: Plenum Press.

De Groot, A.D. (1965): Thought and Choice in Chess. The Hague: Mouton.

Dewe, B. (1997): Grenzen der Didaktik: Über den Hiatus zwischen Lehrerwissen und Lehrerkönnen. In: Keuffer, J./Meyer, M.A. (Hrsg.): Didaktik und kultureller Wandel. Aktuelle Problemlagen und Veränderungsperspektiven. Weinheim: Deutscher Studienverlag, S. 220-248.

Dewe, B. (1998): Zur Relevanz der Professionstheorie für pädagogisches Handeln. In: Schulz, W.K. (Hrsg.): Expertenwissen. Soziologische, psychologische und pädagogische Perspektiven. Opladen: Leske + Budrich, S. 67-86.

Dewe, B./Ferchhoff, W./Radtke, F.O. (1992): Das Professionswissen von Pädagogen. In: Dies. (Hrsg.): Erziehen als Profession. Zur Logik professionellen Handelns in pädagogischen Feldern. Opladen: Leske + Budrich, S. 70-91.

Dewe, B./Radtke, F.-O. (1991): Was wissen Pädagogen über ihr Können? Professionstheoretische Überlegungen zum Theorie-Praxis-Problem in der Pädagogik. In: Zeitschrift für Pädagogik, 27. Beiheft, S. 143-162.

Dieterich, R. (1983): Konzepte der Pädagogischen Handlungskompetenz – Überblick über das Forschungsgebiet. In: Ders. (Hrsg.): Pädagogische Handlungskompetenz. Paderborn/München/Wien/Zürich: Schöningh, S. 13-33.

Ditton, H. (2002): Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspektiven. In: Unterrichtswissenschaft, 30 (3), S. 197-212.

Ditton, H./Krecker, L. (1995): Qualität von Schule und Unterricht – Empirische Befunde zu Fragestellungen und Aufgaben der Forschung. In: Zeitschrift für Pädagogik, Jg. 41, S. 507-529.

Dörig, R. (2006): Schlüsselqualifikationen aus kognitionspsychologischer Sicht. In: Arnold, R./Müller, H.-J. (Hrsg.): Kompetenzentwicklung durch Schlüsselqualifikations-Förderung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 55-71.

Dorlöchter, H./Krüger, U./Stiller, E./Wiebusch, D. (2005): Schau in den Spiegel. Unterrichtsvideos im Vorbereitungsdienst. In: Journal für LehrerInnenbildung, 5 (2), S. 31-35.

Dörner, D. (1979): Problemlösen als Informationsverarbeitung. Stuttgart: Kohlhammer.

Dörner, D. (2007): Die Logik des Misslingens. Strategisches Denken in komplexen Situationen. Reinbek: Rowohlt.

Dörr, G. (1991a): Plausibilitätserzeugung im Bereich mathematischer Sachverhalte – Experimentelle Analysen des Lösens semantischer Paradoxien mit Hilfe von Analogien. Saarbrücken: Univ. Diss.

Dörr, G. (1991b): Analogiegeleitete Schemainduktion – Experimentelle Überprüfung transfererleichternder Maßnahmen. In: Unterrichtswissenschaft, Jg. 19, S. 121-134.

Dörr, G./Küster, O. (2006): Erforschung berufsfeldbezogener Kompetenzen – Entwicklung eines Forschungsdesigns. In: Hilligus, A.H./Rinkens, H.-D. (Hrsg.): Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 223-229.

Doyle, W. (1986): Classroom Organization and Management. In: Wittrock, M. (Ed.): Handbook of Research on Teaching. New York: Macmillan, pp. 392-431.

Drerup, H. (1987): Wissenschaftliche Erkenntnis und gesellschaftliche Praxis. Anwendungsprobleme der Erziehungswissenschaft in unterschiedlichen Praxisfeldern. Weinheim: Deutscher Studien Verlag.

Dreyfuss, H.L./Dreyfuss, S.E. (1986): Künstliche Intelligenz. Von den Grenzen der Denkmaschine und dem Wert der Intuition. Reinbek: Rowohlt.

Ebke, H./Kliemann, P. (2006): Ethos und Lehrerbildung. Pädagogische Trends und Problemstellungen. In: Grunder, H.-U./Schweitzer, F. (Hrsg.): Gemeinschaft – Ethos – Schule. Eine praxisnahe Einführung für Ausbildung und Fortbildung. Weinheim: Beltz, S. 148-168.

Eder, F. (1996): Schul- und Klassenklima. Ausprägung, Determinanten und Wirkungen des Klimas an weiterführenden Schulen. Innsbruck: Studienverlag.

Eder, F. (2002): Unterrichtsklima und Unterrichtsqualität. In: Unterrichtswissenschaft, 30 (3), S. 213-229.

Ehlich, K./Rehbein, J. (1977): Wissen, kommunikatives Handeln und die Schule. In: Goeppert, H.C. (Hrsg.): Sprachverhalten im Unterricht. Zur Kommunikation von Lehrer und Schüler in der Unterrichtssituation. München: Fink, S. 36-114.

Ehlich, K./Rehbein, J. (1979): Sprachliche Handlungsmuster. In: Soeffner, H.-G. (Hrsg.): Interpretative Verfahren in den Sozial- und Textwissenschaften. Stuttgart: Metzler, S. 243-274.

Ehrenspeck , Y. (2004): Bildung. In: Krüger, H.-H./Grunert, C. (Hrsg.): Wörterbuch Erziehungswissenschaft. Wiesbaden: Verlag für Sozialwissenschaften, S. 64-71.

Einsiedler, W. (2002): Das Konzept „Unterrichtsqualität“. In: Unterrichtswissenschaft, 30 (3), S. 194-196.

Erpenbeck, J. /Rosenstiel, L. (Hrsg.) (2003): Handbuch Kompetenzmessung: Erkennen, Verstehen und Bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis. Stuttgart: Schäffer-Poeschel.

Erpenbeck, J./Sauer, J. (2001): Das Forschungs- und Entwicklungsprogramm „Lernkultur-Kompetenz-Entwicklung“. QUEM, Arbeitsgemeinschaft betriebliche Weiterbildungsforschung e.V. Berlin, H. 63.

Faulstich, P. (2002): Verteidigung von „Bildung“ gegen die Gebildeten unter ihren Verächtern. In: Literatur- und Forschungsreport Weiterbildung, Nr. 49, S. 15-25.

Fend, H. (1977): Schulklima. Soziale Einflussprozesse in der Schule. Weinheim: Beltz.

Fend, H. (1982): Gesamtschule im Vergleich. Weinheim: Beltz.

Fend, H. (1988): Die Wiederentdeckung der Schule als pädagogische Gestaltungsebene. In: Neue Sammlung, 28, S. 537-547.

Fend, H. (1998): Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim: Juventa.

Fend, H. (2000): Qualität und Qualitätssicherung im Bildungswesen – Wohlfahrtsstaatliche Modelle und Marktmodelle. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule. In: Zeitschrift für Pädagogik, 41. Beiheft. Weinheim: Beltz, S. 55-72.

Fend, H. (2002): Mikro- und Makroverfahren eines Angebot-Nutzungsmodells von Schulleistungen. Zum Stellenwert der pädagogischen Psychologie bei der Erklärung von Schulleistungsunterschieden verschiedener Länder. In: Zeitschrift für Pädagogische Psychologie, 16 (3/4), S. 141-149.

Fenstermacher, G. (1994): The Knower and the Known. The Nature of Knowledge in Research on Teaching. In: Darling-Hammond, L. (Ed.): Review of Research in Education, Vol. 20. Washington, DC: American Educational Research Association, pp. 2-56.

Fischer, D. (2006): Didaktische Gestaltungsmuster des Religionsunterrichts. Vergleichende Fallstudie mit Videoaufzeichnungen. In: Rahm, S./Mammes, I./Schratz, M. (Hrsg.): Schulpädagogische Forschung. Innsbruck: Studienverlag, S. 27-39.

Fischer, D./Schratz, M. (Hrsg.) (2005): Videos in der LehrerInnenbildung. In: Journal für LehrerInnenbildung, 5 (2), S. 4-7.

Forneck, H.J. (2002): Selbstgesteuertes Lernen und Modernisierungsimperative in der Erwachsenen- und Weiterbildung. In: Zeitschrift für Pädagogik, Jg. 48, S. 243-261.

Franzke, M. (1996): Transfer kognitiver Fertigkeiten. In: Hoffmann, J./Kintsch, W. (Hrsg.): Lernen. Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung. Ser. 2, Kognition, Bd. 7. Göttingen: Hogrefe, S. 355-387.

Frey, A. (2002): Berufliche Handlungskompetenz – Kompetenzentwicklung und Kompetenzvorstellungen in der Erzieherinnenausbildung. In: Empirische Pädagogik, 16 (2), S. 139-156.

Frey, A. (2004): Die Kompetenzstruktur von Studierenden des Lehrerberufs. Eine internationale Studie. In: Zeitschrift für Pädagogik, 50 (6), S. 903-925.

Frey, A. (2006): Strukturierung und Methoden zur Erfassung von Kompetenz. In: Bildung und Erziehung, 59 (2), S. 125-145.

Frey, A./Balzer, L. (2003): Beurteilungsbogen zu sozialen und methodischen Kompetenzen – smk99. In: Erpenbeck, J./Rosenstiel, L. (Hrsg.): Handbuch Kompetenzmessung. Erkennen, Verstehen und Bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis. Stuttgart: Schäffer-Poeschel, S. 323-335.

Friebertshäuser, B. (1997): Feldforschung und teilnehmende Beobachtung. In: Friebertshäuser, B./Prengel, A. (Hrsg.): Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim: Juventa, S. 503-534.

Friedrich, H.F./Mandel, H. (1992): Lern- und Denkstrategien – ein Problemaufriss. In: Dies.: Lern- und Denkstrategien. Analyse und Intervention. Göttingen: Hogrefe, S. 3-54.

Gage, N.L. (Ed.) (1963): Handbook of Research on Teaching. Chicago: Rand McNally.

Gage, N.L./Berliner, D.C. (1996): Pädagogische Psychologie. Weinheim: Beltz.

Gage, N.L./Needels, M.C. (1989): Process-Product Research on Teaching. A Review of Criticisms. In: The Elementary School Journal, pp. 253-300.

Gagné, R.M. (1975): Die Bedingungen des menschlichen Lernens. Hannover: Schroedel.

Geißler, K.A. (1974): Berufserziehung und kritische Kompetenz. Ansätze einer Interaktionspädagogik. München: Reinhardt.

Geißler, K.A./Orthey, F.M. (2002): Kompetenz: Ein Begriff für das Verwertbare Ungefähre. In: Literatur- und Forschungsreport Weiterbildung, Nr. 49, S. 69-79.

Gerber, B./ Knierim, B./Labudde, P. (2006): Lernoperationen im Physikunterricht: Videoanalysen von Lehrsequenzen. In: Pitton, A. (Hrsg): Lehren und Lernen mit neuen Medien. Münster: LIT Verlag, S. 228-300.

Gieseke, W. (1992): Durch berufliche Sozialisation zur Professionalität? In: Pädagogische Arbeitsstelle des Deutschen Volkshochschulverbandes (Hrsg.): Studienbibliothek für Erwachsenenbildung, Bd. 1: Gesellschaftliche Voraussetzungen der Erwachsenenbildung. Frankfurt/M.: PAS des DVV, S. 159-172.

Girmes, R. (2006a): Lehrprofessionalität in einer demokratischen Gesellschaft. Über Kompetenzen und Standards in einer erziehungswissenschaftlich fundierten Lehrerbildung. In: Zeitschrift für Pädagogik, 51. Beiheft, S. 14-29.

Girmes, R. (2006b): Lehrerprofessionalität: Über Kompetenzen und Standards einer fundierten Lehreraus- und -fortbildung. In: Hilligus, A.H./Rinkens, H.-D. (Hrsg.): Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 63-70.

Good, T.L./Biddle, B.J./Brophy, J.E. (1975): Teachers Make a Difference. New York: Holt, Rinehart & Winston.

Grell, J. (1974): Techniken des Lehrerverhaltens. Weinheim: Beltz.

Grell, J. (2000): Direktes Unterrichten. Ein umstrittenes Unterrichtsmodell. In: Wiechmann, J. (Hrsg.): Zwölf Unterrichtsmethoden. Vielfalt für die Praxis. Weinheim: Beltz, S. 35-49.

Greve, W./Wentura, D. (1997): Wissenschaftliche Beobachtungen. Eine Einführung. Weinheim: Psychologie Verlags Union.

Groeben, N.,/Wahl, D./Schlee, J./Scheele, B. (1988): Das Forschungsprogramm Subjektive Theorien. Einführung in die Psychologie des reflexiven Subjekts. Tübingen: Francke.

Gruber, H. (1997): Wege zum Können. Determinanten des Kompetenzerwerbs. Bern: Huber.

Gruber, H. (1998): Expertise. In: Rost, D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: Beltz, S. 126-129.

Gruber, H. (1999): Erfahrung als Grundlage kompetenten Handelns. Bern: Huber.

Gruber, H. (2000): Lehr-Lern-Forschung: Den Erwerb komplexer Kompetenzen müssen wir lehren und lernen. In: Unterrichtswissenschaft, Jg. 28, S. 38-43.

Gruber, H./Harteis, C./Rehrl, M. (2006): Professional Learning: Erfahrung als Grundlage von Handlungskompetenz. In: Bildung und Erziehung, 59 (2), S. 193-203.

Gruber, H./Mandl, H. (1996): Das Entstehen von Expertise. In: Enzyklopädie der Psychologie, Bd.7: Lernen. Göttingen: Hogrefe, S. 583-615.

Gruber, H./Rehrl, M. (2005): Praktikum statt Theorie? – Eine Analyse relevanten Wissens zum Aufbau pädagogischer Handlungskompetenz. In: Journal für LehrerInnenbildung, 5 (1), S. 8-16.

Gruber, H./Renkl, A. (2000): Die Kluft zwischen Wissen und Handeln: Das Problem des trägen Wissens. In: Neuweg, G.H. (Hrsg.): Wissen, Können, Reflexion. Ausgewählte Bestimmungen. Innsbruck: Studienverlag, S. 155-174.

Gruehn, S. (2000): Unterricht und schulisches Lernen. Münster: Waxmann.

Gruschka, A. (1995): Wie Schüler Erzieher werden. Wetzlar: Büchse der Pandorra.

Gruschka, A./Herrmann, U./Radtke, F.-O./Rauin, U./Ruhloff, J./Rumpf, H./Winkler, M. (2005): Das Bildungswesen ist kein Wirtschafts-Betrieb. Fünf Einsprüche gegen die technokratische Umsteuerung des Bildungswesens. In: Frankfurter Rundschau. Verfügbar unter URL: http://www.frankfurter-rundschau.de/fr_home/startseite/?cnt=738484 (Zugriff am 28.10.05).

Gudjons, H. (2003): Frontalunterricht – neu entdeckt. Integration in offene Unterrichtsformen. Bad Heilbrunn: Klinkhardt.

Gudjons, H. (2006): Neue Unterrichtskultur – veränderte Lehrerrolle. Bad Heilbrunn: Klinkhardt.

Habermas, J. (1987a): Theorie des kommunikativen Handelns, Bd. 1: Handlungsrationalität und gesellschaftliche Rationalisierung. Frankfurt/M.: Suhrkamp.

Habermas, J. (1987b): Theorie des kommunikativen Handelns, Bd. 2: Zur Kritik der funktionalistischen Vernunft. Frankfurt/M.: Suhrkamp.

Habermas, J. (1990): Strukturwandel der Öffentlichkeit. Frankfurt/M: Suhrkamp.

Häcker, T./Rihm, T. (2005): Professionelles Lehrer(innen)handeln – Plädoyer für eine situationsbezogene Wende. In: v. Carlsburg, G.-B./Meisteikene, I. (Hrsg.): Bildungsreform als Lebensform. Frankfurt/M: Lang, S. 359-380.

Hall, R. (2000): Videorecording as Theory. In: Kelly, A./Lesh, R. (Eds.): Handbook of Research Design in Mathematics and Science Education. Mahwah, NJ: Erlbaum, pp. 647-664.

Hascher, T. (2003): Diagnose als Voraussetzung für gelingende Lernprozesse. In: Journal für LehrerInnenbildung, 3 (2), S. 25-30.

Hascher, T. (2004): Professionelle Entwicklung von Lehrpersonen. In: Journal für LehrerInnenbildung, 4 (1), S. 44-50.

Hascher, T. (2005): Die Erfahrungsfalle. In: Journal für Lehrerinnen- und Lehrerbildung, 5 (1), S. 40-46.

Hascher, T. (2006): Veränderungen im Praktikum – Veränderungen durch das Praktikum: Eine empirische Untersuchung zur Wirkung von schulpraktischen Studien in der Lehrerbildung. In: Zeitschrift für Pädagogik, 51. Beiheft, S. 130-148.

Hascher, T./Thonhauser, J. (2004): Editorial. In: Journal für LehrerInnenbildung, 4 (1), S. 4-9.

Hascher, T./Wehr, S. (2005): Offener Geografie-Unterricht. Wirkungen und Probleme aus der Sicht von Lehrpersonen. Projektbericht der Forschungsstelle für Schulpädagogik und Fachdidaktik an der Universität Bern.

Hedtke, R. (2000): Theorie-Praxis-Problem der Lehrerbildung am Exempel schulpraktischer Studien. In: sowi-onlinejournal, H.2, http://www.sowi-online.de/journal/lehrerbildung/hedtke.htm (Zugriff am 07.10.2006).

Heid, H. (2000): Qualität. Überlegungen zur Begründung einer pädagogischen Beurteilungskategorie. In: Zeitschrift für Pädagogik, 41. Beiheft, S. 41-72.

Heid, H./Harteis, C. (Hrsg.) (2005): Verwertbarkeit. Ein Qualitätskriterium (erziehungs-)wissenschaftlichen Wissens? Wiesbaden: Verlag für Sozialwissenschaften.

Heimann, P./Otto, G./Schulz, W. (1965): Unterricht. Analyse und Planung. Hannover: Schrödel.

Helmke, A. (2002): Kommentar: Unterrichtsqualität und Unterrichtsklima: Perspektiven und Sackgassen. In: Unterrichtswissenschaft, 30 (3), S. 261-277.

Helmke, A. (2003): Unterrichtsqualität erfassen, bewerten, verbessern. Seelze: Kallmeyer.

Helmke, A. (2006): Unterrichtsforschung. In: Arnold, K.-H./Sandfuchs, U./Wiechmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt, S. 56-65.

Helmke, A./Jäger, R.S. (Hrsg.) (2002): Die Studie Markus-Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik.

Helmke, A./ Schneider, W./Weinert, F.E. (1986): Quality of Instruction and Classroom Learning Outcomes – Results of the German Contribution to the Classroom Environment Study of the IEA. In: Teaching and Teacher Education, Vol. 2, pp. 1-18.

Helmke, A./Schrader, F.-W. (2001): Jenseits von TIMMS: Messungen sprachlicher Kompetenzen, komplexe Längsschnittstudien und kulturvergleichende Analysen. Ergebnisse und Perspektiven ausgewählter Längsschnittstudien. In: Weinert, F.E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim: Beltz, S. 237-250.

Helmke, A./Weinert, F.E. (1997): Psychologie des Unterrichts und der Schule. In: Enzyklopädie der Psychologie. D. I. 3. Göttingen: Hogrefe, S. 71-175.

Helsper, W. (1996): Antinomien des Lehrerhandelns in modernisierten pädagogischen Kulturen. Paradoxe Verwendungsweisen von Autonomie und Selbstverantwortlichkeit. In: Combe, A./Helsper, W. (Hrsg.): Pädagogische Professionalität. Frankfurt/M.: Suhrkamp, S. 521-569.

Helsper, W. (2000): Antinomien des Lehrerhandelns und die Bedeutung der Fallrekonstruktion – Überlegungen zu einer Professionalisierung im Rahmen universitärer Lehrerausbildung. In: Cloer, E./Klika, E./Kunert, H. (Hrsg.): Welche Lehrer braucht das Land? Weinheim/München: Juventa, S. 141-178.

Herzog, W. (2005): Müssen wir Standards wollen? Skepsis gegenüber einem theoretisch (zu) schwachen Konzept. In: Zeitschrift für Pädagogik, 51 (2), S. 252-258.

Hesse, M.B. (1970): Models and Analogies in Science. Notre Dame, Ind.: University of Notre Dame Press.

Heursen, G. (1989): Didaktik, allgemeine. In: Lenzen, D. (Hrsg.). Pädagogische Grundbegriffe. Reinbek: Rowohlt, S. 307-317.

Hilligus, A.H./Rinkens, H.D. (Hrsg.) (2005): Zentren für Lehrerbildung – Neue Wege im Bereich der Praxisphasen. Münster: Lit-Verlag.

Hoeltje, B./Oberliesen, R./Schwedes, H./Ziemer, T. (2002): Evaluation des Halbjahrespraktikums der Lehrerbildung der Universität Bremen. Bremen: Universität Bremen.

Hoeltje, B./Oberliesen, R./Schwedes, H./Ziemer, T. (2004): Das Halbjahrespraktikum in der Lehrerausbildung der Universität Bremen. Bremen: Universität Bremen.

Honnefelder, L. (2002): Sittlichkeit/Ethos. In: Düwell, M./Hübenthal, C./Werner, M. (Hrsg.): Handbuch Ethik. Stuttgart: Metzler, S. 491-496.

Hornecker, E. (2004): Videobasierte Interaktionsanalyse – der Blick durch die (Zeit-)Lupe auf das Interaktionsgeschehen kooperativer Arbeit. In: Boes, A./Pfeiffer, S. (Hrsg.): Informationsarbeit neu verstehen. Methoden zur Erfassung informatisierter Arbeit. München: ISF Forschungsberichte.

Huber, H.D. (2001): Interkontextualität und künstlerische Kompetenz: Eine kritische Auseinandersetzung. In: Bühler, M./Koch, A. (Hrsg.). Kunst und Interkontextualität: Materialien zum Symposium Schau-Vogel-Schau. Köln: Salon.

Hugener, I./Pauli, C./Reusser, K. (2007): Inszenierungsmuster, kognitive Aktivierung und Leistung im Mathematikunterricht. In: Lemmermöhle, D./Rothgangel, M./Bögeholz, S./Hasselhorn, M./Watermann, R. (Hrsg.): Professionell Lehren. Erfolgreich Lernen. Münster: Waxmann, S. 109-122.

Hugener, I./Rakoczy, K./Pauli, C./Reusser, K. (2006): Videobasierte Unterrichtsforschung: Integration verschiedener Methoden der Videoanalyse für eine differenzierte Sicht auf Lehr-Lernprozesse. In: Rahm, S./Mammes, I./Schratz, M. (Hrsg.): Schulpädagogische Forschung. Innsbruck: Studienverlag, S. 41-53.

Ingenkamp, K. (Hrsg.) (1970): Handbuch der Unterrichtsforschung, Bd. 1. Weinheim: Beltz.

Ingenkamp, K. (1977): Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte. Weinheim: Beltz.

Ingenkamp, K. (1988): Lehrbuch der pädagogischen Diagnostik. Weinheim: Beltz.

Jacobs, J.K./Kawanaka, T./Stigler, J. (1999): Integrating Qualitative and Quantitative Approaches to the Analysis of Video Data on Classroom Teaching. In: International Journal of Educational Research, 31, pp. 717-724.

Jank, W./Meyer, H. (2005): Didaktische Modelle. Berlin: Cornelsen.

Janssen, J./Laatz, W. (2003): Statistische Datenanalyse mit SPSS für Windows. Berlin: Springer.

Keller, H.-J. (2002): Standards in der Lehrerbildung. Blick in die USA und Ausblick auf die deutschsprachige Lehrpersonenausbildung. In: Journal für LehrerInnenbildung, 2 (1), S. 20-28.

Keller, J.A./Novak, F. (1993): Bildung. In: Dies.: Kleines Pädagogisches Wörterbuch. Freiburg: Herder, S. 63-65.

Kersten, B. (2001): Befragung der Lehrerinnen und Lehrer: In: Oser, F./Oelkers, J. (Hrsg.): Die Wirksamkeit der Lehrerbildungssysteme. Zürich: Rüegger, S. 399-435.

Kiel, E. (2005): Klassenführung. In: Apel, H.J./Sacher, W. (Hrsg.): Studienbuch Schulpädagogik. Bad Heilbrunn: Klinkhardt, S. 327-342.

Klafki, W. (1963): Das pädagogische Problem des Elementaren und die Theorie der kategorialen Bildung. Weinheim/Bergstr.: Beltz 1963.

Klieme, E. (2004): Was sind Kompetenzen und wie lassen sie sich messen? In: Pädagogik, H. 6, S. 10-13.

Klieme, E. (2006a): Leistungsstand und Leistungsentwicklung im Deutschen und im Englischen: Zentrale Befunde der DESI-Studie. Am 21.03.2006 gehaltener Vortrag im Rahmen des DGfE-Kongresses „Bildung. Macht. Gesellschaft“ an der Johann Wolfgang Goethe Universität, Frankfurt/M.

Klieme, E. (2006b): Ergebnisse der DESI-Studie. Vortrag auf dem ersten Wuppertaler Bildungsdialog am 10.11.2006. Verfügbar unter URL: http://www.zbl.uni-wuppertal.de/zentrum/assets/Vortrag_Klieme.pdf (Zugriff am 16.04.2007).

Klieme, E. et al. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Frankfurt/M.: DIPF.

Klieme, E. et al. (2007): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Hrsg. v. Bundesministerium für Bildung und Forschung. Bonn: BMBF.

Klieme, E./Reusser, K. (2003): Unterrichtsqualität und mathematisches Verständnis im internationalen Vergleich. In: Unterrichtswissenschaft, 31 (3), S. 194-205.

Kluge, F. (2002): Etymologisches Wörterbuch der deutschen Sprache. Berlin: de Gruyter.

KMK (2004): Standards für die Lehrerbildung: Bericht der Arbeitsgruppe. Verfügbar unter URL: http://www.kmk.org/Lehrerbildung-Bericht%20der%20AG.pdf (Zugriff am 24.03.05).

Kobarg, M./Seidel, T. (2007): Prozessorientierte Lernbegleitung – Videoanalysen im Physikunterricht der Sekundarstufe I. In: Unterrichtswissenschaft, 35 (2), S. 148-168.

Koch, L. (1999): Bildung. In: Reinhold, G./Pollak, G./Heim, H. (Hrsg.): Pädagogik-Lexikon. München: Oldenbourg, S. 78-84.

Koch-Priewe, B. (2002): Der routinierte Umgang mit Neuem. Wie die Professionalisierung von JunglehrerInnen gelingen kann. In: Beetz-Rahm, S./Denner, L./Riecke-Baulecke, Th. (Hrsg.): Jahrbuch für Lehrerforschung und Bildungsarbeit, Bd. 3. Weinheim/München: Juventa, S. 311-324. Verfügbar unter URL: http://www.uni-koeln.de/ew-fak/einr/koekola/ (Zugriff am 17.08.06).

Kocher, M./Wyss, C. (2006): Unterrichtsbezogene Kompetenzen in der Lehrerinnen- und Lehrerausbildung. Eine Videoanalyse. Zürich: Lizentiatsarbeit am Pädagogischen Institut der Universität Zürich.

Köck, P. (1997): Bildung. In: Köck, P./Ott, H.: Wörterbuch für Erziehung und Unterricht. Donauwörth: Auer, S. 98f.

Köck, P. (2005): Handbuch der Schulpädagogik. Donauwörth: Auer.

Kohler, B. (2002): Zur Rezeption von TIMMS durch Lehrerinnen und Lehrer. In: Unterrichtswissenschaft, 30 (2), S. 158-188.

Kolbe, F.-U. (2001): Konvergenzen in der Lehrerwissensforschung – ein Beitrag zu den Grundfragen allgemeiner Didaktik. In: Finkbeiner, C./Schnaitmann, G.W. (Hrsg.): Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik. Donauwörth: Auer, S. 184-206.

Kolbe, F.-U. (2004): Verhältnis von Wissen und Handeln. In: Blömeke, S./Reinhold, P./Tulodziecki, G./Wildt, J. (Hrsg). Handbuch Lehrerbildung. Bad Heilbrunn/Obb.: Klinkhardt, S. 206-232.

Kolbe, F.-U./Combe, A. (2004): Lehrerbildung. In: Helsper, W./Böhme, J. (Hrsg.): Handbuch der Schulforschung. Wiesbaden: Verlag für Sozialwissenschaften, S. 853-877.

Kommer, S./ Biermann, R. (2004): Video(technik) in der erziehungswissenschaftlichen Forschung. In: MedienPädagogik, S. 1-30. Verfügbar unter URL: http://www.medienpaed.com/04-1/kommer04-1.pdf (Zugriff am 13.01.2007).

Kondakov, N.J. (1983): Wörterbuch der Logik. Leipzig: VEB.

Konsortium Bildungsberichterstattung (Hrsg.) (2006): Bildung in Deutschland. Bielefeld: Bertelsmann.

Kösel, E./Scherer, H. (2006): „Schlüsselqualifikationen“ – Die Illusion einer Realität? In: Arnold, R./Müller, H.-J. (Hrsg.): Kompetenzentwicklung durch Schlüsselqualifikations-Förderung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 141-154.

Kounin, J.S. (2006): Techniken der Klassenführung. Münster: Waxmann.

Krammer, K./Hugener, I. (2005): Netzbasierte Reflexion von Unterrichtsvideos in der Ausbildung von Lehrpersonen – eine Explorationsstudie. In: Beiträge zur Lehrerbildung, 23 (1), S. 51-61.

Krammer, K./Reusser, K. (2005): Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. In: Beiträge zur Lehrerbildung, 23 (1), S. 35-50.

Kucharz, D./Wagener, M. (2006): Ein altes Verfahren neu entdeckt. Die Pädagogische Tatsachenforschung. In: Rahm, S./Mammes, I./Schratz, M. (Hrsg.): Schulpädagogische Forschung. Innsbruck: Studienverlag, S. 97-106.

Kunter, M. (2005): Multiple Ziele im Mathematikunterricht. Münster: Waxmann.

Langewand, A. (2002): Bildung. In: Lenzen, D. (Hrsg.): Erziehungswissenschaft. Reinbek: Rowohlt, S. 69-98.

Lang-von Wins, T. (2003): Die Kompetenzhaltigkeit von Methoden moderner psychologischer Diagnostik-, Personalauswahl- und Arbeitsanalyseverfahren sowie aktueller Managementdiagnostikansätze. In Erpenbeck, J./Rosenstiel, L. v. (Hrsg.): Handbuch Kompetenzmessung. Stuttgart: Schäffer-Poeschel, S. 585-618.

Larcher Klee, S. (2005): Einstieg in den Lehrerberuf. Untersuchungen zur Identitätsentwicklung von Lehrerinnen und Lehrern im ersten Berufsjahr. Bern: Haupt.

Larcher, S./Oelkers, J. (2004): Deutsche Lehrerbildung im internationalen Vergleich. In: Blömeke, S./Reinhold, P./Tulodziecki, G./Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn/Obb.: Klinkhardt, S. 128-150.

Lass, U./Lüer, G. (1990): Psychologische Problemlöseforschung. In: Unterrichtswissenschaft, 18 (4), S. 295-312.

Lemmermöhle, D./Schellack, A./Brümmer, F. (2006): „Lehrerbildung auf dem Wege“ – ein Projekt an der Universität Göttingen. In: Hilligus, A.H./Rinkens, H.-D.: Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 231-237.

Leuchtner, M./Pauli, C./Reusser, K./Lipowsky, F. (2006): Unterrichtsbezogene Überzeugungen und handlungsleitende Kognitionen von Lehrpersonen. In: Zeitschrift für Erziehungswissenschaft, 9 (4), S. 562-579.

Leutner, D. (1992): Adaptive Lehrsysteme. Weinheim: Psychologie Verlags Union.

Lipowski, F. (2006): Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In: Zeitschrift für Pädagogik, 51. Beiheft: Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf, S. 47-70.

Lipowsky, F./Rakoczy, K./Pauli, C./Reusser, K./Klieme, E. (2007): Gleicher Unterricht – gleiche Chancen für alle? Die Verteilung von Schülerbeiträgen im Klassenunterricht. In: Unterrichtswissenschaft, 35 (2), S. 125-147.

Lipowski, F./Thußbas, C./Klieme, E./Reusser, K./Pauli, C. (2003): Professionelles Lehrerwissen, selbstbezogene Kognitionen und wahrgenommene Schulumwelt. In: Unterrichtswissenschaft, 31 (3), S. 206-237.

Lissmann, U. (2002): Beurteilung und Beurteilungsprobleme bei Portfolios. In: Jäger, R.S. (Hrsg.): Von der Beobachtung zur Notengebung – ein Lehrbuch. Landau: Verlag Empirische Pädagogik, S. 299-345.

Lortie, D.C. (1975): Schoolteacher. A Sociological Study. Chicago: University of Chicago Press.

Luhmann, N. (1990): Die Homogenisierung des Anfangs. Zur Ausdifferenzierung der Schulerziehung. In: Luhmann, N./Schorr, K.E. (Hrsg.): Zwischen Anfang und Ende. Fragen an die Pädagogik. Frankfurt/M.: Suhrkamp, S. 73-111.

Luhmann, N. (2002): Das Erziehungssystem der Gesellschaft. Frankfurt/M.: Suhrkamp.

Luhmann, N./Schorr, K.E. (1988): Reflexionsprobleme im Erziehungssystem. Frankfurt/M.: Suhrkamp.

Maag Merki, K./Grob, U. (2003): Überfachliche Kompetenzen: Zur Validierung eines Indikatorensystems. In: Empirische Didaktik, 17 (2), S. 123-147.

Mandl, H./Gruber, H./Renkl, A. (1992): Prozesse der Wissensanwendung beim komplexen Problemlösen in einer kooperativen Situation. In: Achtenhagen, F./John, E.G. (Hrsg.). Mehrdimensionale Lehr-Lern-Arrangements – Innovationen in der kaufmännischen Aus- und Weiterbildung. Wiesbaden: Gabler, S. 478-490.

Marcel, A. (1983): Conscious and Unconscious Perception: Experiments in Visual Masking and Word Recognition. Cognitive Psychology, 15, pp. 197-237.

Mayr, J. (2002): Sich Standards aneignen – Befunde zur Bedeutung der Lernwege und der Bearbeitungstiefe. In: Journal für LehrerInnenbildung, 2 (1), S. 29-37.

Mayr, J. (2003): Persönlichkeitsfragebögen in der Lehrerforschung und Lehrerberatung. In: Samac, K.: Empirisches Arbeiten in der Arbeitsgemeinschaft der Bewegungserzieherinnen und Bewegungserzieher an Pädagogischen Akademien. In: Theorie & Praxis – Texte zur LehrerInnenbildung, Heft 20. Wien: bm:bwk, S. 79-89. Verfügbar unter URL: http://www.bmbwk.gv.at/medien/11287_TP_20.pdf (Zugriff am 28.10.05).

Mayr, J. (2005): Wie sich Lehrpersonen Kompetenzen aneignen. Zur Nutzung und Wirksamkeit institutionalisierter und informeller „Opportunities to Learn“. Vortrag auf der 67. Tagung der Arbeitsgruppe für empirische pädagogische Forschung (AEPF): Qualität durch Standards? 19.09. bis 21.09.2005, Salzburg.

Mayr, J. (2006a): Persönlichkeitsentwicklung im Studium. Eine Pilotstudie zum Wirkungspotenzial von Lehrerbildung. In: Hilligus, A.H./Rinkens, H.-D.: Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 249-260.

Mayr, J. (2006b): Theorie + Übung + Praxis = Kompetenz? Empirisch begründete Rückfragen zu den „Standards in der Lehrerbildung“. In: Zeitschrift für Pädagogik, 51. Beiheft, S. 149-163.

Mayr, J./Paseka, A. (2002): „Lehrerpersönlichkeit“. In: Journal für LehrerInnenbildung, 2 (2), S. 50-55.

Merkens, H. (Hrsg.) (2003): Lehrerbildung in der Diskussion. Opladen: Leske + Budrich 2003.

Merkens, H. (Hrsg.) (2005): Lehrerbildung: Zentren für Lehrerbildung. Wiesbaden: Verlag für Sozialwissenschaften.

Mertens, D. (1974): Schlüsselqualifikationen. Thesen zur Schulung für eine moderne Gesellschaft. In: MittAB, H. 1, S. 36-43.

Mertens, D. (1977): Schlüsselqualifikationen. In: Siebert, H.: Begründungen gegenwärtiger Erwachsenenbildung. Braunschweig: Westermann, S. 99-121.

Mertens, D. (1988): Das Konzept der Schlüsselqualifikation als Flexibilisierungsinstrument. In: Literatur- und Forschungsreport Weiterbildung, Nr. 22, S. 33-46.

Messner, H. (1978): Wissen und Anwenden. Zur Problematik des Transfers im Unterricht. Stuttgart: Klett-Cotta.

Messner, R. (2001): Szenarien zur Bearbeitung des Theorie-Praxis-Problems in der Lehrerbildung. In: Journal für LehrerInnenbildung, 1 (2), S. 10-19.

Meyer, H. (1980): Rezeptionsprobleme der Didaktik oder wie Lehrer lernen. In: Adl-Amini, B./Künzli, R. (Hrsg.): Didaktische Modelle und Unterrichtsplanung. München: Juventa, S. 88-118.

Meyer, H. (2001): Türklinkendidaktik. Aufsätze zur Didaktik, Methodik und Schulentwicklung. Berlin: Cornelsen.

Meyer, H. (2004): Was ist guter Unterricht? Berlin: Cornelsen.

Meyers Lexikonverlag (2007): Online-Lexikon. Verfügbar unter URL: http://lexikon.meyers.de/meyers/Kontingenz (Zugriff am 03.08.2007).

Mintrop, H. (2006): Einen qualifizierten Lehrer für jede Klasse. Neuere Ansätze der Lehrerausbildung in den Vereinigten Staaten von Amerika im Zuge der Standard-Bewegung. In: Hilligus, A.H./Rinkens, H.-D.: Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 89-116.

Mortimore, P./Sammons, P./Stoll, L./Lewis, D./Ecob, R. (1988): School matters. The Junior Years. Sommerset: Open Books.

Mortimore, P./Sammons, P./Stoll, L./Lewis, D./Ecob, R. (1989): A Study of Effective Junior Schools. In: International Journal of Educational Research, 13, pp. 753-768.

Mühlhausen, U. (2005): Multimediale Unterrichtsdokumente. In: Journal für LehrerInnenbildung, 5 (2), S. 19-25.

Müller, C./Eichler, D./Blömeke, S. (2006): Chancen und Grenzen von Videostudien in der Unterrichtsforschung. In: Rahm, S./Mammes, I./Schratz, M. (Hrsg.): Schulpädagogische Forschung. Innsbruck: Studienverlag, S. 125-138.

Müller, K. (i.V.): Entwicklungsverlauf professioneller Kompetenzen von Studierenden in einem Lehrerbildungskonzept mit integrierter Praxisphase – Empirische Überprüfung der Wirksamkeit des Modellversuches „Kompetenzentwicklung im Praxisjahr (KOPRA) anhand einer Fragebogen- und Interviewstudie. Arbeitstitel einer geplanten Dissertation an der Pädagogischen Hochschule Weingarten.

Murnane, R. (1985): Comparisons of Private and Public Schools: Lessons from the Uproar. In: Journal of Human Resources, 20, 263-267.

Murnane, R.J./Phillips, B.R. (1981): What do Effective Teachers of Inner-City Children have in Common? In: Social Science Research (Orlando), 10, pp. 83-100.

Nakamura, Y./Böckelmann, C./Tröhler, D. (Hrsg.) (2006): Theorie versus Praxis? Perspektiven auf ein Missverständnis. Zürich: Pestalozzianum.

Neuweg, G.H. (2000): Können und Wissen. Eine alltagssprachphilosophische Verhältnisbestimmung. In: Ders. (Hrsg.): Wissen – Können – Reflexion. Innsbruck/Wien: Studienverlag, S. 65-82.

Neuweg, G.H. (2005): Implizites Wissen als Forschungsgegenstand. In: Rauner, F. (Hrsg.): Handbuch der Berufsbildungsforschung. Bielefeld: Bertelsmann, S. 581-588.

Neuweg, G.H. (2006): Das Schweigen der Könner: Strukturen und Grenzen des Erfahrungswissens. Linz: Trauner.

Nölle, K. (2002): Probleme der Form und des Erwerbs unterrichtsrelevanten pädagogischen Wissens. In: Zeitschrift für Pädagogik, 48 (1), 48-67.

Obolenski, A./Meyer, H. (2003): Forschendes Lernen. Theorie und Praxis einer professionellen LehrerInnenausbildung. Bad Heilbrunn: Klinkhardt.

OECD (2007): OECD-Veröffentlichung „Bildung auf einem Blick“. Wesentliche Aussagen in der Ausgabe 2007. Verfügbar unter URL: http://www.bmbf.de/pub/bildung_auf_einen_blick_07_wesentliche_aussagen.pdf (Zugriff am 05.10.2007).

Oelkers, J. (1999): Studium als Praktikum? Illusionen und Aussichten der Lehrerbildung. In: Radtke, F.-O. (Hrsg.): Lehrerbildung an der Universität. Zur Wissensbasis pädagogischer Professionalität. Frankfurt/M.: Fachbereich Erziehungswiss. der Johann-Wolfgang-Goethe-Univ.

Oelkers, J. (2002): Die Entwicklung von Standards in der Lehrerbildung. Vortrag im Studienseminar Göttingen für das Lehramt an Gymnasien am 10. Dezember 2002. Verfügbar unter URL: http://www.paed.unizh.ch/ap/downloads/oelkers/Vortraege/087_GoettingenStudienseminar.pdf (Zugriff am 13.05.05).

Oelkers, J. (2004a): Zum Problem von Standards aus historischer Sicht. In: Neue Sammlung, Jg. 44, S. 179-200.

Oelkers, J. (2004b): Was müssen gute Lehrerinnen und Lehrer wissen und können? Vortrag vor Praktikumslehrerinnen und Praktikumslehrern in der Pädagogischen Hochschule Zürich am 10. März 2004.

Oerter, R. (2000): Implizites Lernen beim Sprechen, Lesen und Schreiben. In: Unterrichtswissenschaft, 28 (3), S. 239-256.

Oevermann, U. (1996): Theoretische Skizze einer revidierten Theorie professionalisierten Handelns. In: Combe, A./Helsper, W. (Hrsg.): Pädagogische Professionalität. Frankfurt/M.: Suhrkamp, S. 70-182.

Oser, F. (1997a): Standards in der Lehrerbildung, Teil 1. In: Beiträge zur Lehrerbildung, 15 (1), S. 26-37.

Oser, F. (1997b): Standards in der Lehrerbildung, Teil 2. In: Beiträge zur Lehrerbildung, 15 (2), S. 220-230.

Oser, F. (1998): Ethos – die Vermenschlichung des Erfolgs. Zur Psychologie der Berufsmoral von Lehrpersonen. Opladen: Leske + Budrich.

Oser, F. (2001): Standards. Kompetenzen von Lehrpersonen. In: Oser, F./Oelkers, J. (Hrsg.): Die Wirksamkeit der Lehrerbildungssysteme. Zürich: Rüegger, S. 215-342.

Oser, F. (2004): Standardbasierte Evaluation der Lehrerbildung. In: Blömeke, S./Reinhold, P./Tulodziecki, G./Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn/Obb.: Klinkhardt, S. 184-205.

Oser, F. (2005): Schrilles Theoriegezerre, oder warum Standards gewollt sein sollen. In: Zeitschrift für Pädagogik, 51 (2), S. 266-274.

Oser, F. (2006): Zugänge ermöglichen, Zugänge verwehren: Entwurf eines Ausleseverfahrens in der Lehrerbildung. Ein Essay. In: Beiträge zur Lehrerbildung, 24 (1) , S. 30-42.

Oser, F./Baeriswyl, F.J. (2001): Choreographies of Teaching: Bridging Instruction to Learning. In: Richardson, V. (Eds.): AERA Handbook of Research on Teaching. Washington: American Education Research Association, pp. 1031-1065.

Oser, F./Curcio, G.-P./Düggeli, A. (2007): Kompetenzmessung in der Lehrerbildung als Notwendigkeit – Fragen und Zugänge. In: Beiträge zur Lehrerbildung, 25 (1), S. 15-35.

Oser, F./Dick, Patry, J.-L. (Eds.) (1992): Effective and Responsible Teaching: The New Synthesis. San Francisco: Jossey-Bass.

Oser, F./Oelkers, J. (Hrsg.) (2001): Die Wirksamkeit der Lehrerbildungssysteme. Von der Allrounderbildung zur Ausbildung professioneller Standards. Chur/Zürich: Rüegger.

Oser, F./Patry, J.-L. (1990): Choreographien unterrichtlichen Lernens. Basismodelle des Unterrichts. Berichte zur Erziehungswissenschaft Nr. 89. Universität Freiburg (Schweiz): Pädagogisches Institut.

Oser, F./Patry, J.-L. (1994): Sichtstruktur und Basismodelle des Unterrichts: Über den Zusammenhang von Lehren und Lernen unter dem Gesichtspunkt psychologischer Lernverläufe. In: Olechowski, R./Rollett, B. (Hrsg.): Theorie und Praxis. Aspekte empirisch-pädagogischer Forschung – quantitative und qualitative Methoden. Frankfurt/M.: Lang, S. 138-146.

Oser, F./Renold, U. (2005): Kompetenzen von Lehrpersonen – über das Auffinden von Standards und ihre Messung. In: Zeitschrift für Erziehungswissenschaft, 4. Beiheft, S. 119-140.

Oser, F./Sarasin, S. (1995): Basismodelle des Unterrichts: Von der Sequenzierung als Lernerleichterung. Verfügbar unter URL: http://opus.kobv.de/ubp/volltexte/2005/469/ (Zugriff am 14.07.2007).

Parsons, T. (1964): Die akademischen Berufe und die Sozialstruktur. In: Ders. (Hrsg.): Beiträge zur soziologischen Theorie. Neuwied: Luchterhand, S. 160-178.

Patry, J.-L. (2005): Zum Problem der Theoriefeindlichkeit der Praktiker. In: Heid, H./Harteis, C. (Hrsg.): Verwertbarkeit. Ein Qualitätskriterium (erziehungs-) wissenschaftlichen Wissens? Wiesbaden: Verlag für Sozialwissenschaften, S. 143-161.

Pätzold, G./Wortmann, E. (2006): Didaktische Handlungsmöglichkeiten zur Ausbildung von Schlüsselqualifikationen. In: Arnold, R./Müller, H.-J. (Hrsg.): Kompetenzentwicklung durch Schlüsselqualifikations-Förderung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 155-173.

Pauli, C./Lipowsky, F. (2007): Mitmachen oder zuhören? Mündliche Schülerinnen- und Schülerbeteiligung im Mathematikunterricht. In: Unterrichtswissenschaft, 35 (2), S. 101-124.

Pauli, C./Reusser, K. (2003): Unterrichtsskripts im schweizerischen und im deutschen Mathematikunterricht. In: Unterrichtswissenschaft, 31 (3), S. 238-272.

Pauli, C./Reusser, K. (2006): Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. In: Zeitschrift für Pädagogik, 52 (6), S. 774-798.

Perrig, W.J. (1990): Implizites Wissen. Eine Herausforderung für die Kognitionspsychologie. In: Schweizerische Zeitschrift für Psychologie, 49, S. 234-249.

Perrig, W.J. (1996): Implizites Lernen. In: Hoffmann, J./Kintsch, W. (Hrsg.): Lernen. Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung. Ser. 2, Kognition, Bd. 7. Göttingen: Hogrefe, S. 203-234.

Perrig, W.J. (2006): Theorie und Praxis in der Lehrerbildung im Kontext impliziter und expliziter Erfahrung. In: Nakamura, Y./ Böckelmann, C./Tröhler, D. (Hrsg.): Theorie versus Praxis? Perspektiven auf ein Missverständnis. Zürich: Pestalozzianum, S. 53-71.

Perrig, W.J./Hofer, D. (1989): Sensory and Conceptional Representations in Memory: Motor Images which cannot Imaged. In: Psychological Research, 51, pp. 201-207.

Petko, D./Waldis, M./Pauli, C./Reusser, K. (2003): Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik. In: Zeitschrift für die Didaktik der Mathematik, 35 (6), S. 265-280.

Plöger, W. (2006): Was ist Kompetenz? – Ein theoretischer Rahmen mit Blick auf die beruflichen Fähigkeiten von Lehrerinnen und Lehrern. In: Pädagogische Rundschau, 60 (3), S. 255-270.

Polanyi, M. (1985): Implizites Wissen. Frankfurt/M.: Suhrkamp.

Pollack, G./Heid, H. (Hrsg.) (1994): Von der Erziehungswissenschaft zur Pädagogik. Weinheim: Deutscher Studien Verlag.

Popper, K. (1935/2005): Logik der Forschung. Tübingen: Mohr-Siebeck.

Popper, K. (2000): Vermutungen und Widerlegungen. Ausgabe in einem Band. Tübingen: Mohr-Siebeck.

Posner, M.I. (1988): Introduction: What is it to be an Expert? In: Chi, M.T.H./Glaser, R./Farr, M.J. (Eds.): The Nature of Expertise. Hillsdale, NJ: Erlbaum, pp. xxix-xxxvi.

Preckel, F./Holling, H. (2006): Die Rolle von Intelligenz und Begabung für Handlungskompetenz am Beispiel beruflicher Hochbegabung. In: Bildung und Erziehung, 59 (2), S. 167-178.

Prenzel, M./ Mandl, H. (1992): Transfer of Learning from a Constructivist Perspective. In: Research Report, No. 6. München: Ludwig-Maximilians-Universität.

Proske, M. (2006): Die Innovierung der empirischen Unterrichtsforschung und das Problem der Kontingenz. Zur Reichweite neuerer theoretischer und methodologischer Ansätze. In: Rahm, S./Mammes, I./Schratz, M. (Hrsg.): Schulpädagogische Forschung. Innsbruck: Studienverlag, S. 139-152.

Radtke, F.-O. (1996): Wissen und Können. Die Rolle der Erziehungswissenschaft in der Erziehung. Opladen: Leske + Budrich.

Ravitch, D. (1995): National Standards in American Education. A Citizen's Guide. Washington, D.C.: Brookings Institution Press.

Ravitch, D. (2000): Left Back: A Century of Failed School Reform. New York: Simon & Schuster.

Reber, A.S./Walkenfeld, F.F./Hernstadt, R. (1991): Implicit and Explicit Learning. Individual Differences and IQ. In: Journal of Experimental Psychology: Learning, Memory and Cognition, 5, pp. 888-896.

Reetz, L. (2006): Schlüsselqualifikationen aus bildungstheoretischer Sicht – in der berufs- und wirtschaftspädagogischen Diskussion. In: Arnold, R./Müller, H.-J. (Hrsg.): Kompetenzentwicklung durch Schlüsselqualifikations-Förderung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 39-54.

Regenbrecht, A. (2005): Sichern Bildungsstandards die Bildungsaufgabe der Schule? In: Pädagogisches Forum, H. 1, S. 16-22.

Reh, S. (2004): Abschied von der Profession, von Professionalität oder vom Professionellen? In: Zeitschrift für Pädagogik, 50 (3), S. 358-372.

Reinhoffer, B./Barthold, R./Küster, O. (2007): Modellversuch Praxisjahr Biberach (GHS). Eine neue Form der Theorie-Praxis-Verknüpfung in der Lehrerbildung. In: Flagmeyer, D./Rotermund, M. (Hrsg.): Mehr Praxis in der Lehrerbildung – aber wie? Leipzig: Universitätsverlag, S. 77-94.

Reinmann-Rothmeier, G./Mandl, H. (1997): Lehren im Erwachsenenalter. Auffassungen vom Lehren und Lernen, Prinzipien und Methoden. In: Weinert, F.E./Mandl, H. (Hrsg.): Enzyklopädie der Psychologie, Themenbereich D, Praxisgebiete, Serie 1, Pädagogische Psychologie, Bd. 4: Psychologie der Erwachsenenbildung. Göttingen: Hogrefe, S. 355-403.

Reinmann-Rothmeier, G./Mandl, H. (2001): Unterrichten und Lernumgebungen gestalten. In: Krapp, A./Weidenmann, B. (Hrsg.): Pädagogische Psychologie. Weinheim: Beltz, S. 601-646.

Renkl, A./Nückles, M. (2006): Träge Kompetenzen? – Gründe für die Kontextgebundenheit von beruflichen Handlungskompetenzen. In: Bildung und Erziehung, 59 (2), S. 179-191.

Reusser, K. (2005a): Problemorientiertes Lernen – Tiefenstruktur, Gestaltungsformen, Wirkung. In: Beiträge zur Lehrerbildung, 23 (2), S. 159-182.

Reusser, K. (2005b): Situiertes Lernen mit Unterrichtsvideos. In: Journal für LehrerInnenbildung, 5 (2), S. 8-18.

Reusser, K. (2006): Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In: Baer, M./Fuchs, M./Füglister, P./Reusser, K./Wyss, H. (Hrsg.): Didaktik auf psychologischer Grundlage. Bern: h.e.p. Verlag, S. 151-168.

Richter, D. (2006): Civic literacy, reading literacy – gibt es auch eine „politische Lesekompetenz“? In: Gesellschaft für Politikdidaktik und politische Jugend- und Erwachsenenbildung (Hrsg.): Standards der Theoriebildung und empirischen Forschung in der politischen Bildung. Schwalbach/Ts.: Wochenschau Verlag, S. 57-70.

Rinkens, H.-D. (Hrsg.) (1999): Zentrum für Lehrerbildung – fünf Jahre Unterstützung und Weiterentwicklung der Lehrerausbildung. Münster: Lit-Verlag.

Rose, H. (1992): Erfahrungsgeleitete Arbeit als Innovationskonzept für Arbeitsgestaltung und Technikentwicklung. In: Zeitschrift für Arbeitswissenschaft, 46 (18), S. 145-149.

Rosenthal, R./Jacobson, L. (1968): Pygmalion in the Classroom. New York: Holt, Rinehart & Winston.

Rost, D. (2005b): Interpretation und Bewertung pädagogisch-psychologischer Studien. Weinheim/Basel: Beltz.

Rost, J. (2005a): Psychometrische Modelle zur Überprüfung von Bildungsstandards anhand von Kompetenzmodellen. In: Zeitschrift für Pädagogik, 50 (5), S. 662-678.

Rotermund, M. (Hrsg.) (2006): Schulpraktische Studien – Evaluationsergebnisse und neue Wege der Lehrerbildung. Leipzig: Universitätsverlag.

Roth, H. (1966): Pädagogische Anthropologie, Bd. 1. Bildsamkeit und Bestimmung. Dortmund: Schroedel.

Roth, H. (1971): Pädagogische Anthropologie, Bd. 2. Entwicklung und Erziehung. Grundlagen einer Entwicklungspädagogik. Hannover: Schroedel.

Rupp, A.A./Leucht, M./Hartung, R. (2006): ‚Die Kompetenzbrille aufsetzen'. Verfahren zur multiplen Klassifikation von Lernenden für Kompetenzdiagnostik in Unterricht und Testung. In: Unterrichtswissenschaft, 34 (3), S. 195-219.

Rutter, M./Maughan, B./Mortimore, P./Ouston, J. (1980): Fünfzehntausend Stunden: Schulen und ihre Wirkung auf die Kinder. Weinheim: Beltz.

Rychen, D.S./Salganik, L.H. (Hrsg.) (2000): INES GENERAL ASSEMBLY 2000. A Contribution of the OECD Program Definition and Selection of Competencies: Theoretical and Conceptual Foundations. Defining and Selection of Key Competencies. Neuenburg: SFSO, OECD, pp. 1-13.

Schaarschmidt, U./Fischer, A.W. (2001): Bewältigungsmuster im Beruf. Persönlichkeitsunterschiede in der Auseinandersetzung mit der Arbeitsbelastung. Göttingen: Vandenhoeck & Ruprecht.

Schaub, H./Zenke, K.G. (2000): Bildung. In: Dies.: Wörterbuch Pädagogik. München: Deutscher Taschenbuch Verlag, S. 95f.

Scheerens, J./Bosker, R. (1997): The Foundations of Educational Effectiveness. Oxford: Elsevier Science.

Scheunpflug, A. (2004): Das Technologiedefizit. Nachdenken über Unterricht aus systemtheoretischer Perspektive. In: Lenzen, D. (Hrsg.): Irritationen des

Erziehungssystems. Pädagogische Resonanzen auf Niklas Luhmann. Frankfurt/M.: Suhrkamp, S. 65-87.

Schiefele, U./Pekrun, R. (1996): Psychologische Modelle des fremdgesteuerten und selbstgesteuerten Lernens. In: Weinert, F.E. (Hrsg.): Psychologie des Lernens und der Instruktion, Bd. 2. Göttingen: Hogrefe, S. 249-278.

Schmid, C. (2006): Lernen und Transfer: Kritik der didaktischen Steuerung. Bern: h.e.p. Verlag.

Schön, D. (1983): The Reflective Practitioner. New York: Basic Books.

Schorb, A.O. (1975): Bildung: In: Ders.: Pädagogisches Taschenlexikon. Bochum: Kamp, S. 41f.

Schröder, H. (1992): Bildung. In: Ders.: Grundwortschatz Erziehungswissenschaft. München: Ehrenwirth, S. 36-38.

Schulze, H. (2001): Erfahrungsgeleitete Arbeit in der industriellen Produktion. Berlin: Edition Sigma.

Schwenk, B. (1989): Bildung. In: Lenzen, D. (Hrsg.): Pädagogische Grundbegriffe, Bd. 1. Reinbek: Rowohlt, S. 208-221.

Seidel, T. (2003a): Lehr-Lernskripts im Unterricht. Münster: Waxmann.

Seidel, T. (2003b): Videobasierte Kodierverfahren in der IPN Videostudie Physik – ein methodischer Überblick. In: Seidel, T./Prenzel, M./Duit, R./Lehrke, M. (Hrsg.): Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“. Kiel: IPN, S. 99-111.

Seidel, T./Prenzel, M./Duit, R. et al. (2002): „Jetzt bitte alle nach vorne schauen!“ – Lehr-Lernskripts im Physikunterricht und damit verbundene Bedingungen für individuelle Lernprozesse. In: Unterrichtswissenschaft, 30 (1), S. 1-10.

Seipp, B. (2003): Standards in der Lehrerbildung: eine Befragung zur Vermittlung der OSERschen Standards in der ersten Phase der Lehramtsausbildung. Bochum: Projekt Verlag.

Shulman, L.S. (1986): Those who Understand. Knowledge Growth in Teaching. In: Educational Researcher, Vol. 15l, No. 2, pp 4-14.

Shulman, L.S. (1987): Knowledge and Teaching: Foundations of the New Reform. In: Harvard Educational Review, Vol. 57, No. 1, pp 1-22.

Siebert, H. (2006): Schlüsselqualifikationen aus erwachsenenpädagogischer Sicht. In: Arnold, R./Müller, H.-J. (Hrsg.): Kompetenzentwicklung durch

Schlüsselqualifikations-Förderung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 31-38.

Slavin, R.E. (1989): A Theory of School and Classroom Organization. In: Slavin, R.E. (Ed.): School and Classroom Organization. Hillsdale, NJ: Erlbaum, pp. 3-24.

Spinner, K.H. (2003): Lesekompetenz – was untersucht IGLU und was nicht? In: Grundschulverband aktuell, H. 82, S. 10-12.

Spranger, E. (1958): Der geborene Erzieher. Heidelberg: Quelle & Meyer.

Stadler, H. (2003): Videos als Mittel zur Qualitätsverbesserung von Unterricht. In: Brunner, E.J./Noack, P./Scholz, G./Scholl, I. (Hrsg.): Diagnose und Intervention in schulischen Handlungsfeldern. Münster: Waxmann, S. 175-193.

Stark, R./Graf, M./Renkl, A./Gruber, H./Mandl, H. (1995): Förderung von Handlungskompetenz durch geleitetes Problemlösen und multiple Lernkontexte. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 27 (4), S. 289-312.

Staub, C. (2005): Videos im Fachspezifisch-Pädagogischen Coaching. In: Journal für LehrerInnenbildung, 5 (2), S. 26-30.

Staub, C. (2006): Allgemeine Didaktik und Lernpsychologie: Zur Dynamisierung eines schwierigen Verhältnisses. In: Baer, M./Fuchs, M./Füglister, P./Reusser, K./Wyss, H. (Hrsg.): Didaktik auf psychologischer Grundlage. Bern: h.e.p. Verlag, S. 169-179.

Stern, E. (2006): Kommentar zu Hans Aeblis „Psychologische Didaktik“ aus der Sicht einer lern-lehrtheoretischen Didaktik. In: Baer, M./Fuchs, M./Füglister, P./Reusser, K./Wyss, H. (Hrsg.): Didaktik auf psychologischer Grundlage. Bern: h.e.p. Verlag, S. 139-142.

Stigler, J.W./Gallimore, R./Hiebert, J. (2000): Using Video Surveys to Compare Classrooms and Teaching across Cultures: Examples and Lessons from the TIMMS Video Studies. In: Educational Psychologist, 35 (2), pp. 87-100.

Stigler, J.W./Gonzales, P./Kawanaka, T./Knoll, S./Serrano, A. (1999): The TIMMS Videotape Classroom Study. Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan and the United States. Washington, D.C.: U.S. Department of Education.

Stigler, J.W./Hiebert, J. (1997): Understanding and Improving Classroom Mathematics Instruction: an Overview of the TIMMS Video Study. In: Phi Delta Kappan, 79 (1), pp. 14-21.

Tenorth, H.-E. (1997): „Bildung“ – Thematisierungsformen und Bedeutung in der Erziehungswissenschaft. In: Zeitschrift für Pädagogik, H. 6, S. 969-984.

Tenorth, H.-E. (2006): Professionalität im Lehrerberuf. In: Zeitschrift für Erziehungswissenschaft, 9 (4), S. 580-597.

Terhart, E. (2000): Perspektiven der Lehrerbildung in Deutschland. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Kommission. Materialband zum Abschlussbericht der KMK. Weinheim/Basel: Beltz.

Terhart, E. (2002): Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz. ZKL-Texte Nr. 24. Universität Münster: Zentrale Koordination Lehrerbildung.

Terhart, E. (2004): Wirkungen von Lehrerbildung: Perspektiven einer an Standards orientierten Evaluation. In: Beckmann, U./Brandt, H./Wagner, H. (Hrsg.): Ein neues Bild vom Lehrerberuf? Pädagogische Professionalität nach PISA. Weinheim: Beltz, S. 87-98.

Terhart, E. (2005a): Über Traditionen und Innovationen oder: Wie geht es weiter mit der Allgemeinen Didaktik? In: Zeitschrift für Pädagogik, 51 (1), S. 1-13.

Terhart, E. (2005b): Standards für die Lehrerbildung – ein Kommentar. In: Zeitschrift für Pädagogik, 51 (2), S. 275-279.

Terhart, E. (2005c): Die Lehre in den Zeiten der Modularisierung. In: Teichler, U./Tippelt, R. (Hrsg.): Hochschullandschaft im Wandel. In: Zeitschrift für Pädagogik, 50. Beiheft. Weinheim: Beltz, S. 87-102.

Terhart, E. (2006): Standards und Kompetenzen in der Lehrerbildung. In: Hilligus, A.H./Rinkens, H.-D. (Hrsg.): Standards und Kompetenzen – neue Qualität in der Lehrerausbildung. Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit-Verlag, S. 29-42.

Terhart, E. (2007): Erfassung und Beurteilung der beruflichen Kompetenz von Lehrkräften. In: Lüders, M./Wissinger, J. (Hrsg.): Forschung zur Lehrerbildung. Münster: Waxmann, S. 37-62.

Tietgens, H. (1992): Reflexionen zur Erwachsenendidaktik. Bad Heilbrunn: Klinkhardt.

Tippelt, R. (Hrsg.) (2002): Handbuch Bildungsforschung. Opladen: Leske + Budrich.

Thonhauser, J. (2005): Videos in der Lehrerbildung – ein historischer Überblick. In: Journal für LehrerInnenbildung, 5 (2), S. 44-48.

Trendel, G./Wackermann, R./Fischer, H.E. (2007): Lernprozessorientierte Lehrerfortbildung in Physik. In: Zeitschrift für die Didaktik der Naturwissenschaften, Jg. 13, S. 9-31.

Tulodziecki, G. (2004): Bildungsstandards im erziehungswissenschaftlichen Studium für Lehrämter. In: Hilligus, A. (Hrsg.): Zur Entwicklung von Standards für die Lehrerausbildung (1). Paderborn: PLAZ, S. 29-56.

Tulodziecki, G./Grafe, S. (2006): Stellenwert und Kritik von Standards für die Lehrerbildung. Internationale Vergleiche und Einschätzungen zur Situation. In: Journal für LehrerInnenbildung, 8 (1), S. 34-44.

Twardella, J. (2004): Rollenförmig oder als „ganze Person“? Ein Beitrag zur Diskussion über die Professionalisierungstheorie und die Struktur pädagogischen Handelns. In: Pädagogische Korrespondenz, Jg. 33, S. 65-74.

Übersax, J.S. (2001): Statistical Methods for Rater Agreement. Verfügbar unter URL: http://ourworld.compuserve.com/homepages/jsuebersax (Zugriff am 26.11.2005).

Ulich, K. (1983): Schüler und Lehrer im Schulalltag. Eine Sozialpsychologie der Schule. Weinheim: Beltz.

Urban, W. (1984): Persönlichkeitsstruktur und Unterrichtskompetenz. Wien: Österreichischer Bundesverlag.

Urban, W. (2005): Theorie der personalen Ressourcen für den Umgang, die Bewältigung und Lösung komplexer Probleme in sozialen Situationen. Manuskript.

Vanderstraeten, R. (2004): Erziehung als Kommunikation – Doppelte Kontingenz als systemtheoretischer Grundbegriff. In: Lenzen, D. (Hrsg.): Irritationen des Erziehungssystems. Pädagogische Resonanzen auf Niklas Luhmann. Frankfurt/M.: Suhrkamp, S. 37-64.

Von Hentig, H. (2006): Die Schule neu denken. Eine Übung in pädagogischer Vernunft. Weinheim: Beltz.

Vonken, M. (2001): Von der Bildung zur Kompetenz. Die Entwicklung erwachsenenpädagogischer Begriffe oder die Rückkehr zur Bildung? In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 97 (4), S. 503-522.

Vonken, M. (2005): Handlung und Kompetenz. Theoretische Perspektiven für die Erwachsenen- und Berufspädagogik. Wiesbaden: Verlag für Sozialwissenschaften.

Von Prondczynsky, A. (in Vorbereitung): Bildungstheorie – Bildungskritik – Bildungsforschung. Zum Wandel der Bildungssemantik. In: Wigger, L. (Hrsg.): Wie ist Bildung möglich? Bad Heilbrunn: Klinkhardt.

Wahl, D. (1991): Handeln unter Druck. Der weite Weg vom Wissen zum Handeln bei Lehrern, Hochschullehrern und Erwachsenenbildnern. Weinheim: Deutscher Studien Verlag.

Wahl, D. (2005): Lernumgebungen erfolgreich gestalten. Wirksame Wege vom trägen Wissen zum kompetenten Handeln in der Erwachsenenbildung, Hochschuldidaktik und Unterricht. Bad Heilbrunn: Klinkhardt.

Wahl, D./Weinert, F.E./Huber, G.L. (2001): Psychologie für die Schulpraxis. München: Kösel.

Waldis, M./Gautschi, P./Hodel, J./Reusser, K. (2006): Die Erfassung von Sichtstrukturen und Qualitätsmerkmalen im Geschichtsunterricht. Methodologische Überlegungen am Beispiel der Videostudie „Geschichte und Politik im Unterricht". In: Günter-Arndt, H./Sauer, M. (Hrsg.): Geschichtsdidaktik empirisch. Untersuchungen zum historischen Denken und Lernen. Berlin: Lit-Verlag, S. 155-188.

Watzlawick, P./Beavin, J.H./Jackson, D.D. (1990): Menschliche Kommunikation. Formen, Störungen, Paradoxien. Bern: Huber.

Wayne A.J. / Youngs, P. (2006): Die Art der Ausbildung von Lehrern und die Lerngewinne ihrer Schüler. Eine Übersicht über aktuelle empirische Forschung. In: Zeitschrift für Pädagogik, 51. Beiheft, S. 71-96.

Wechsler, D. (1956): Die Messung der Intelligenz Erwachsener. Bern: Huber.

Weidenmann, B. (2001): Lernen mit Medien. In: Krapp, A./Weidenmann, B. (Hrsg.): Pädagogische Psychologie. Ein Lehrbuch. Weinheim: Beltz, S. 417-465.

Weinert, F.E. (1996): Für und Wider die neuen „Lerntheorien" als Grundlagen pädagogisch-psychologischer Forschung. In: Zeitschrift für Pädagogische Psychologie, Jg. 10, S. 1-12.

Weinert, F.E. (1998): Guter Unterricht ist ein Unterricht, in dem mehr gelernt, als gelehrt wird. In: Freund, J./Gruber, H./Weidinger, W. (Hrsg.): Guter Unterricht – Was ist das? Aspekte von Unterrichtsqualität. Wien: ÖBV Pädagogischer Verlag, S. 7-18.

Weinert, F.E. (1999): Konzepte der Kompetenz. Paris: OECD.

Weinert F.E. (2000): Lehr-Lernforschung an einer kalendarischen Zeitwende. Im alten Trott weiter oder Aufbruch zu neuen wissenschaftlichen Horizonten? In: Unterrichtswissenschaft, 28 (1), S. 44-48.

Weinert, F.E. (2001): Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: Weinert, F.E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim/Basel: Beltz, S. 17-32.

Weinert, F.E./Schrader, F.-W. (1986): Diagnose des Lehrers als Diagnostiker. In: Pettilon, H./Wagner, J.W.L./Wolf, B. (Hrsg.): Schülergerechte Diagnose. Weinheim: Beltz, S. 11-29.

Weinert, F.E./Schrader, F.-W. (1997): Lernen als psychologisches Problem. In: Weinert, F.E./Mandl, H. (Hrsg.): Psychologie der Erwachsenenbildung. Göttingen: Hogrefe, S. 295-335.

Weinert, F.E./Schrader, F.W./Helmke, A. (1989): Quality of Instruction and Achievement Outcomes. In: International Journal of Educational Research, 13, pp. 895-914.

Weinert, F.E./Schrader, F.W./Helmke, A. (1990): Unterrichtsexpertise – ein Konzept zur Verringerung der Kluft zwischen zwei theoretischen Paradigmen. In: Alisch, L.M./Baumert, J./Beck, K. (Hrsg.): Professionswissen und Professionalisierung. Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft, Bd. 28. Braunschweig: Copy-Center Colmsee, S. 173-206.

Weiser, B. (2005): Vom Skilltraining zum Videoportfolio. In: Journal für LehrerInnenbildung, 5 (2), S. 36-43.

Wellenreuther, M. (2004): Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen. Baltmannsweiler: Schneider Verlag Hohengehren.

Wernet, A. (2003): Pädagogische Permissivität. Schulische Sozialisation und pädagogisches Handeln jenseits der Professionalisierungsfrage. Opladen: Leske + Budrich.

Wiechmann, J. (Hrsg.) (2000): Zwölf Unterrichtsmethoden. Vielfalt für die Praxis. Weinheim: Beltz.

Wigger, L. (2004): Bildungstheorie und Bildungsforschung in der Gegenwart. Versuch einer Lagebeschreibung. In: Vierteljahrsschrift für wissenschaftliche Pädagogik, 80 (4), S. 478-494.

Wild, K.P. (2003): Videoanalysen als neuer Impulsgeber für eine praxisnahe prozessorientierte empirische Unterrichtsforschung. In: Unterrichtswissenschaft, 31 (2), S. 98-102.

Wild-Näf, M. (2001): Die Ausbildungen für Lehrkräfte der Deutschschweiz im Urteil der Studierenden: Ein Strukturmodell des Zusammenhangs von Person, Organisation und Ausbildungsprozess. In: Oser, F./Oelkers, J. (Hrsg.): Die Wirksamkeit der Lehrerbildungssysteme. Zürich: Rüegger, S. 141-214.

Wildt, J. (1996): Reflexive Lernprozesse. In: Hänsel, D./Huber, L. (Hrsg.): Lehrerbildung neu denken und gestalten. Weinheim/Basel: Beltz, S. 91-107.

Wildt, J. (2003): Reflexives Lernen in der Lehrerbildung – ein Mehrebenenmodell in hochschuldidaktischer Perspektive. In: Obolenski, A./Meyer, H. (Hrsg.): Forschendes Lernen. Theorie und Praxis einer professionellen Lehrerbildung. Bad Heilbrunn: Klinkhardt, S. 71-84.

Winkler, M. (2003): Wie vorbereitet ist die Jugend auf Arbeit und Beruf? Ergebnisse und Perspektiven aus der Kompetenzforschung. Vortrag anlässlich der 4. Fachtagung des Programms „Schule – Wirtschaft/ Arbeitsleben“ am 18.09.2003 in Neukirchen – Pleiße mit dem Thema „Kompetenzen für die Berufsorientierung nach PISA – Auf welche Kompetenzen kommt es an?“

Wirtz, M./Caspar, F. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen: Hogrefe.

Wollersheim, H.-W. (1993): Kompetenzerziehung. Befähigung zur Bewältigung. Frankfurt/M.: Lang.

Zedler, P. (2002): Erziehungswissenschaftliche Bildungsforschung. In: Tippelt, R. (Hrsg.): Handbuch Bildungsforschung. Opladen: Leske + Budrich, S. 21-39.

Zeichner, K. (2006): Konzepte von Lehrerexpertise und Lehrerausbildung in den Vereinigten Staaten. In: Zeitschrift für Pädagogik, 51. Beiheft, S. 97-115.

Verzeichnis der Abbildungen[95]

Abbildung 3-1:	Hypothetisches Modell beruflicher Kompetenz (Frey 2006)	S. 34
Abbildung 4-1:	Angebot-Nutzungs-Modell der Wirkungsweise des Unterrichts (Helmke 2003, S. 42)	S. 51
Abbildung 10-1:	Vorgehen bei der Entwicklung von Kategoriensystemen (Seidel 2003, S. 104)	S. 172
Abbildung 11-1:	Unterrichtsstatus	S. 189
Abbildung 11-2:	Unterrichtliche Arbeitsformen	S. 191
Abbildung 11-3:	Allgemeindidaktische Unterrichtsphasen	S. 192
Abbildung 11-4:	Aktivitäten im Klassenunterricht	S. 193
Abbildung 11-5:	Unterrichtskommunikation	S. 195
Abbildung 11-6:	Strukturierung	S. 196
Abbildung 11-7:	Differenzierung	S. 197
Abbildung 11-8:	Umgang mit der Unterrichtszeit	S. 204
Abbildung 11-9:	Qualität der Organisation	S. 204
Abbildung 11-10:	Umgang mit Fehlern	S. 205
Abbildung 11-11:	Motivierungsfähigkeit	S. 205
Abbildung 11-12:	Pacing	S. 205
Abbildung 11-13:	Lehrperson als Mediator	S. 206
Abbildung 11-14:	Sprachliche Qualität	S. 206
Abbildung 11-15:	Gesprächsführung	S. 206
Abbildung 11-16:	Zielklarheit	S. 207

[95] Die Titel der Abbildungen und Tabellen sind teilweise leicht verkürzt wiedergegeben.

Verzeichnis der Tabellen

Tabelle 5-1:	Übersicht der thematischen Standardhaupt- und -untergruppen	S. 82
Tabelle 11-1:	Unterrichtsstatus	S. 189
Tabelle 11-2:	Unterrichtliche Arbeitsformen	S. 190
Tabelle 11-3:	Allgemeindidaktische Unterrichtsphasen	S. 191
Tabelle 11-4:	Aktivitäten im Klassenunterricht	S. 192
Tabelle 11-5:	Unterrichtskommunikation	S. 194
Tabelle 11-6:	Strukturierung	S. 195
Tabelle 11-7:	Differenzierung	S. 196
Tabelle 11-8:	Koeffizienten der Intraklassenkorrelation	S. 202
Tabelle 11-9:	Reliabilitäten der Subskalen (Cronbachs Alpha)	S. 203
Tabelle 11-10:	Mittelwerte und Standardabweichungen der Stichprobe	S. 203
Tabelle 11-11:	t-Werte der Stichprobe über die Messpunkte	S. 208

Beschreibung des Modellversuches[96]

Mit Beginn des Schuljahres 2004/2005 startete der Modellversuch „Praxisjahr Biberach“. Von März bis September 2004 lief eine Pilotphase. Jeweils bis zu 16 Studierende absolvieren ab dem dritten Semester ein Praxisjahr an einer der beiden im Modellversuch involvierten Schulen. Dabei werden sie jeweils einer Mentorin/ einem Mentor zugeordnet (Lehrkräfte der Schulen). Darüber hinaus werden sie von Lehrenden der PH Weingarten als auch von Lehrbeauftragten des Staatlichen Seminars Laupheim betreut. Hierbei werden sowohl der von den Studierenden gehaltene Unterricht besprochen als auch bedarfsgerecht Seminarinhalte thematisiert.
Ein halbes Jahr dieses Praxisjahres umfasst die im Studium vorgesehenen beiden Blockpraktika und die fachpraktischen Tagespraktika. Die zweite Hälfte umfasst das erste halbe Jahr des Referendariats. Entsprechend ändert sich der Status der Studierenden.
Ungeachtet der formalen Strukturierung sind über das ganze Praxisjahr hinweg sowohl das schulpraktische Seminar Laupheim als auch die PH Weingarten in die Betreuung einbezogen. Die Studierenden sind in einem Umfang von 20 Unterrichtsstunden pro Woche an den Schulen.
Die Laufzeit des Modellversuchs beträgt fünf Jahre, so dass fünf Durchgänge möglich sind, d.h. fünf Studierendengruppen von 16 Personen können an dem Praxisjahr teilnehmen.

[96] Die Beschreibung stammt aus einem internen Antragspapier an das Rektorat der Pädagogischen Hochschule Weingarten.

Überblick über das Kodiersystem

(in Anlehnung an Seidel 2001 / 2003)

Allgemeine Grundregeln

- Die Kodierung erfolgt in 10-Sekunden-Abschnitten (im Weiteren als „Segmente“ erwähnt).
- In jeder Facette wird grundsätzlich die Zuordnung zu einer Kategorie gefordert (disjunktes Kategoriensystem).
- Wenn innerhalb eines Segmentes zwei Kategorien einer Facette vorkommen, wird die Kategorie gewählt, welche mehr Gewicht (5 oder mehr Sekunden) hat.
- Zu Beginn und am Schluss einer Lektion wird die Kategorie „Unterricht“ auch dann kodiert, wenn während weniger als 5 Sekunden innerhalb des Segmentes Unterricht stattfindet. Diese Regel gilt dementsprechend auch für alle anderen Facetten.
- Wenn eine Kategorie nur einen kleinen Teil des Segmentes ausfüllt (jedoch mindestens 3 Sekunden), wird diese Kategorie trotzdem kodiert und nicht auf „Keine“ gewechselt (z.B. Facette „Strukturierung“, Facette „Unterrichtliche Arbeitsformen“ / Kategorie „Übergang“).
- Die Kategorie 0 wird nur gewählt, wenn keine andere Kategorie innerhalb des Segmentes vorkommt.

Überblick

A Beobachtung von Sichtstrukturen	**B Lernorganisation während des Klassenunterrichts**
1 Unterrichtsstatus	1 Strukturierung
2 Unterrichtliche Arbeitsformen	2 Differenzierung
3 Allgemeindidaktische Unterrichtsphasen	
4 Aktivitäten im Klassenunterricht	
5 Kommunikation im Unterricht	

A Beobachtung von Sichtstrukturen

Facette 1: Unterrichtsstatus		
0	**Kein Unterricht / Unterbrechung**	Betrifft Segmente, in denen Unterbrechungen des Unterrichts durch Dritte vorkommen (z.B. Durchsagen per Lautsprecher, Anklopfen, Probleme mit Mikrophon usw.). Wird nur dann kodiert, wenn die S nicht selbstständig arbeiten und der Klassenunterricht wirklich gestört wird (erkennbar am Verhalten der LP und der S). *Spezifische Kodierungsregeln:* Unterbricht die LP ihren Unterricht selbst, weil sie z.B. etwas sucht und nicht findet, wird Unterricht kodiert.
1	**Vor Unterrichts-beginn**	*Inhaltliche Bestimmung:* Es werden alle Segmente kodiert, in denen noch kein Unterricht stattfindet. *Beschreibung auf der Beobachtungsebene:* Z.B. S kommen nach „Gongschlag“ ins Klassenzimmer, packen ihre Sachen aus, warten auf die LP, bis diese ihre Sachen geordnet hat und der Unterricht beginnen kann. *Spezifische Kodierungsregeln:* Unterscheidung zu Unterricht: „Vor Unterrichtsbeginn“ wird solange kodiert, bis die LP explizit äußert, dass der Unterricht nun beginnen soll.
2	**Unterricht**	*Inhaltliche Bestimmung:* Es werden alle Segmente kodiert, in denen Unterricht stattfindet. Ihre Anzahl belegt, wie viel Zeit tatsächlich für den Unterricht verwendet wird. *Beschreibung auf der Beobachtungsebene:* Keine spezifische Beschreibung notwendig: Unterricht findet statt. *Spezifische Kodierungsregeln:* Explizite Äußerung, dass der Unterricht nun beginnen soll. Es wird auch dann „Unterricht“ kodiert, wenn die Aussage den kleineren Teil des Segmentes in Anspruch nimmt. „Unterricht“ wird bis zur letzten öffentlichen Äußerung der LP kodiert. Auch Begrüßungs- resp. Abschlussrituale (singen etc.) gehören zum Unterricht. *Beispiel:* *„So, nun wollen wir anfangen.“*
3	**Nach Unterrichtsende**	*Inhaltliche Bestimmung:* Es werden alle Segmente nach dem Unterrichtsende kodiert. Das Unterrichtsende wird von der LP bestimmt. *Beschreibung auf der Beobachtungsebene:* Die LP beendet z.B. explizit die Stunde. *Spezifische Kodierungsregeln:* Unterscheidung zu Unterricht: auf die explizite Äußerung der LP achten. Wichtig: auch wenn die Stunde schon zu Ende ist, die LP aber noch weiter unterrichtet, wird „Unterricht“ kodiert. Erst wenn die LP selbst die Stunde beendet, wird danach „Nach Unterrichtsende“ kodiert. Dazu gehören auch organisatorische Aussagen, wie z.B. „Öffnet bitte noch die Fenster.“ „Unterricht“ wird auch dann noch kodiert, wenn die Aussage den kleineren Teil des Segmentes in Anspruch nimmt. Findet keine explizite Äußerung der LP zum Unterrichtsende statt, so muss das Ende der Stunde intentional am Verhalten der LP und der S erschlossen werden.

Facette 2: Unterrichtliche Arbeitsformen Wird nicht Kategorie 1 „Klassenunterricht“ kodiert, ist bei Facette 4 „Aktivitäten im Klassenunterricht“ stets die Kategorie 0 zu kodieren. Für die offenen Unterrichtsformen (Kategorien 2, 3, 4 und 6) gilt in erster Priorität die Aussage der LP (*Bsp.: „Diese Aufgabe löst ihr jetzt alleine.“ = Kategorie 2).* Liegt keine Aussage der LP vor, erfolgt die Kodierung gemäß Handlung in der Filmaufnahme.		
0	**Keine**	Wird immer dann kodiert, wenn keine Arbeitsformen stattfinden.

1	**Klassen-unterricht**	*Inhaltliche Bestimmung:* Im Klassenunterricht arbeitet die LP mit der ganzen Klasse. Die Hauptinteraktion ist öffentlich. *Beschreibung auf der Beobachtungsebene:* Der Klassenunterricht beginnt, wenn die LP explizit äußert, dass die Klasse ihre Aufmerksamkeit auf sie richten soll. Der Klassenunterricht kann vor allem zu Beginn einer Stunde, nach Gruppen- und Stillarbeitsphasen oder nach einer Übergangsphase vorkommen. Typisches Frage-Antwort-Verhalten von LP und S. *Beispiele:* *Beginn der Stunde: „So, wollen wir anfangen… Heute soll es um die Frage gehen, wohin zeigt der…“* *Am Ende der Still- / Gruppenarbeit: „Könntet ihr das Lesen des Textes beenden und mir die Aufmerksamkeit schenken…“* *Am Ende der Übergangsphase: „Nachdem ihr jetzt aufgeräumt habt, wollen wir noch einmal zusammentragen…“* *Spezifische Kodierungsregeln:* Keine.
2	**Stillarbeit / Einzelarbeit**	*Inhaltliche Bestimmung:* Diese Kategorie bezieht sich auf Unterrichtsphasen, in denen S alleine Arbeitsaufträge bearbeiten. *Beschreibung auf der Beobachtungsebene:* S bearbeiten alleine Arbeitsaufträge. Dazu zählen: Texte lesen, Aufgaben lösen, Tafelbild übertragen etc. Meist keine oder wenig Kommunikation. *Spezifische Kodierungsregeln:* Stillarbeit beginnt, sobald die LP alle Instruktionen gegeben hat und alles Material verteilt ist.
3	**Partnerarbeit**	*Inhaltliche Bestimmung:* Kategorie 3 bezieht sich auf eine Unterrichtsform, die durch die Organisation von 2er-Arbeitsgruppen gekennzeichnet ist. *Beschreibung auf der Beobachtungsebene:* S bearbeiten zu zweit einen Arbeitsauftrag. Dazu zählen: Texte lesen, Aufgaben lösen, Experimente durchführen etc. *Spezifische Kodierungsregeln:* Partnerarbeit beginnt, sobald die LP alle Instruktionen gegeben hat und alles Material verteilt ist.
4	**Gruppenarbeit**	*Inhaltliche Bestimmung:* Kategorie 4 bezieht sich auf eine Unterrichtsform, die durch die Organisation von Arbeitsgruppen gekennzeichnet ist. *Beschreibung auf der Beobachtungsebene:* S bearbeiten einen Arbeitsauftrag in Gruppen. Dazu zählen: Aufgaben lösen, Gruppendiskussionen, Dialoge lesen etc. *Spezifische Kodierungsregeln:* Gruppenarbeit beginnt, sobald die LP alle Instruktionen gegeben hat und alles Material verteilt ist.
5	**Übergang**	*Inhaltliche Bestimmung:* Diese Kategorie stellt eine Art „Zwischenphase“ dar, wenn zwischen den anderen genannten Kategorien gewechselt wird. Sie betrifft häufig organisatorische Aktivitäten für die Klasse, wie z.B. Blätter austeilen bzw. einsammeln, Arbeitsmaterialien wegräumen, Arbeitsort wechseln etc. *Beschreibung auf der Beobachtungsebene:* Die LP gibt in dieser Phase keine Anweisungen an die S. Sobald die LP keine öffentliche Aussage mehr macht, beginnt „Übergang“. Das Ende dieser Phase ist angezeigt, sobald die LP erneut zur ganzen Klasse spricht oder alle S im Blickfeld an den Aufgaben arbeiten. *Spezifische Kodierungsregeln:* Kategorie für „Zwischenzustände“ im Unterricht, ohne öffentliche LP-Äußerung. Wenn „Übergang“ aus einer Still-, Partner- oder Gruppenarbeit erfolgt, ohne dass dazu die LP noch einen speziellen Kommentar gibt, so wird „Übergang“ kodiert, sobald das erste Blatt eingesammelt oder ausgeteilt wird oder der erste S sich vom Platz erhebt, um bspw. ein Blatt an der WT zu befestigen.
6	**Mehrere**	*Inhaltliche Bestimmung:*

	Arbeitsformen gleichzeitig	In dieser Kategorie werden Unterrichtsphasen kodiert, in denen mehrere Arbeitsformen gleichzeitig stattfinden. Bspw. gibt die LP die Option in Einzel- oder Partnerformen zu arbeiten. *Beispiel:* „Ihr könnt das alleine oder zu zweit machen." *Beschreibung auf der Beobachtungsebene:* Die LP arbeitet mit einer Gruppe von S im Klassenunterricht, während die anderen S an einer Stillarbeit bzw. Gruppen- oder Partnerarbeit arbeiten. *Spezifische Kodierungsregeln:* Wenn beobachtet werden kann, dass die Mehrheit der S von den Instruktionen der LP abweicht, so wird das Beobachtungsergebnis kodiert und nicht die Intention der LP (z.B. Schüler arbeiten nur in Einzelarbeit, obwohl auch Partnerarbeit möglich wäre).
7	**Störung durch die LP**	*Inhaltliche Bestimmung:* In dieser Kategorie werden Unterrichtsphasen kodiert, in denen die LP eine Stillarbeit, Partner- bzw. Gruppenarbeit durch eine Wortmeldung unterbricht. Eine „Störung" wird nur dann kodiert, wenn die LP eine öffentliche Mitteilung macht, die von allen S gehört wird. Spricht die LP nur mit einzelnen S, so dass die anderen S ungestört weiterarbeiten können, so wird dies nicht als „Störung" kodiert. *Beschreibung auf der Beobachtungsebene:* Die LP hat eine Stillarbeit bzw. Gruppen- oder Partnerarbeit initiiert und die S haben bereits angefangen zu arbeiten. Wird die angefangene Arbeit durch öffentliche Wortmeldungen der LP gestört, die eine Äußerung ohne Bezug zur momentanen Arbeit enthält oder nicht von der LP angekündigt wurde, so wird dies hier kodiert. *Spezifische Kodierungsregeln:* Unterbricht die LP die Stillarbeit bzw. Gruppen- oder Partnerarbeit willentlich mit einer entsprechenden Aussage wie bspw. „Achtung, hört nochmals zu", so wird keine „Störung" kodiert insofern die Aussage in Bezug zur momentanen Arbeit steht. Hier wird auf Kategorie 1 „Klassen-unterricht" gewechselt und die entsprechenden Kodes aus Facette 4 „Aktivitäten im Klassenunterricht" werden gewählt. Gibt die LP weitere Informationen zur Strukturierung, so wird dies zusätzlich unter „Lernorganisation während des Klassenunterrichts" / Facette 1 „Strukturierung" kodiert.
8	**Andere**	Diese Kategorie wird vergeben, wenn eine Zuordnung in die Kategorien 0 – 7 nicht möglich ist.

Facette 3: Allgemeindidaktische Unterrichtsphasen		
0	**Keine**	Betrifft Videosequenzen vor oder nach dem Unterricht bzw., wenn „Kein Unterricht / Unterbrechung" kodiert wurde.
1	**Einführung / Hinführung**	*Inhaltliche Bestimmung:* Kategorie 1 umfasst die Einführung in eine neue Unterrichtsstunde sowie Unterrichtsphasen, in denen explizit neue Inhalte erarbeitet werden. *Spezifische Kodierungsregeln:* Keine.
2	**Arbeits-, Anwendungs- und Vertiefungs-phase**	*Inhaltliche Bestimmung:* Diese Kategorie umfasst Unterrichtsphasen, in denen erarbeitete Inhalte gesichert, eingeübt und auf neue Aufgabenstellungen / Bereiche übertragen und / oder Inhalte vertieft werden. *Beschreibung auf der Beobachtungsebene:* Die Kategorie soll den didaktischen Übergang zu der S-Arbeitsphase markieren und wird erst dann kodiert, wenn die S nach der Instruktion der LP in schriftlicher oder experimenteller Form Aufgabenstellungen bearbeiten oder Fragen in Gruppenarbeitsformen diskutieren mit dem Ziel sie schriftlich zu fixieren oder sie mündlich an die Klasse rückzumelden. *Spezifische Kodierungsregeln*: Der fragend-entwickelnde Unterricht wird hier nicht kodiert. Eine Kodierung wird jedoch vorgenommen, wenn ein einzelner oder mehrere S stellvertretend für die Klasse ein Experiment durchführen oder Lösungen von Aufgabenstellungen der Klasse erklären.

3	**Zusammen-fassung / Rückschau**	*Inhaltliche Bestimmung:* Diese Kategorie umfasst Unterrichtsphasen, in denen zum Abschluss einer Unterrichtsstunde oder -einheit Inhalte strukturiert und organisiert werden. Ziel und Zweck dieser Phase ist die inhaltliche Organisation gelernter Inhalte. *Beschreibung auf der Beobachtungsebene:* Bspw. Zusammenfassung der besprochenen Inhalte und Experimente im Klassengespräch (dies im Sinne einer Wiederholung von Themen, die in der Stunde neu besprochen wurden), Aufschreiben der wichtigsten Inhalte, Hefteinträge sowie Lesen eines zusammenfassenden Textes im Buch. Auf das übergeordnete Ziel achten und weniger auf einzelne Aktivitäten! *Spezifische Kodierungsregeln:* Die Kategorie ist primär auf Inhalte ausgerichtet. So gilt „Zusammenfassen" auch, wenn innerhalb der Stunde die LP bestimmte Inhalte noch einmal strukturiert und zusammenfasst bspw. zwischen zwei Experimenten. Gibt die LP am Ende der Stunde eine „Rückschau" bzw. „Feedback" über den Lernprozess, so wird ebenfalls „Zusammenfassung / Rückschau" kodiert. Auf den Zusammenhang bei der Facette 1 „Lernorganisation während des Klassenunterrichts / Strukturierung" achten. Wird dort Kategorie 4 oder 5 kodiert, dann muss auch hier „Zusammenfassung / Rückschau" kodiert werden.
4	**Prüfen / Leistungs-kontrolle / Haus-aufgaben-kontrolle / Repetition**	*Inhaltliche Bestimmung:* Unterrichtsphasen, die der Kontrolle der S-Leistungen dienen. Ziel der Leistungskontrolle ist eine Rückmeldung über den Leistungsstand der S für die LP. Dazu muss beim Kodieren eine deutliche Äußerung der LP gegeben sein, dass sie Rückmeldung über den Leistungsstand der S erhalten möchte. Wird ebenfalls kodiert, wenn die LP den Inhalt der letzten Stunde mit den S repetiert. *Beschreibung auf der Beobachtungsebene:* Ausfragen, Tests, Kontrollieren der Schulhefte usw. *Spezifische Kodierungsregeln:* Inhaltliche Besprechungen von vorangegangenen Arbeits-, Anwendungs- und Vertiefungsphasen werden nicht kodiert. *Beispiele:* *Bei den Hausaufgaben: „Wer konnte die Hausaufgaben lösen?"* *Bei Abfragen: „Wiederholen wir den Stoff der letzten Stunde. M. soll uns erzählen, was wir in der letzten Stunde gemacht haben."*
5	**Andere**	*Inhaltliche Bestimmung:* Betrifft Videosequenzen, bei denen während des Unterrichts keine Zuordnung zu anderen unterrichtlichen Phasen möglich ist. *Spezifische Kodierungsregeln:* Die Kategorie „Andere" wird bei längeren Übergangsphasen (Aufbau von Experimenten usw.) kodiert, wenn die übergeordnete Zielsetzung / Unterrichtsphase noch nicht erkennbar ist. Kombination: „Übergang" (bei Arbeitsformen) und „Andere" (bei Unterrichtsphasen). Das Erteilen der Hausaufgaben wird hier kodiert.

Facette 4: Aktivitäten im Klassenunterricht Die Aktivitäten im Klassenunterricht beziehen sich auf die Arbeit mit fachlichen Inhalten. Sie werden nur kodiert, wenn „Klassenunterricht" (Kategorie 1) in der Facette 2 „Unterrichtliche Arbeitsformen" kodiert ist. *Spezifische Kodierungsregeln:* Wenn im jeweiligen Segment Strukturierungselemente kodiert werden (B / 1), so wird unter „Aktivitäten im Klassenunterricht" generell Kategorie 12 „Andere" vergeben. Eine Ausnahme innerhalb der Facette „Strukturierung" stellt die Kategorie 4 „Zusammenfassen" dar, weil es dort zentral um die Vermittlung von Inhalten geht; diese Vermittlung findet im Regelfall in Form eines LP-Vortrags oder im Unterrichtsgespräch statt.		
0	**Keine**	Wird immer dann kodiert, wenn keine Aktivitäten stattfinden und wenn in Facette 2 nicht Kategorie 1 „Klassenunterricht" kodiert wird.
1	**LP-Vortrag**	*Inhaltliche Bestimmung:* Die LP spricht über einen längeren Zeitraum (mind. 2 x 10-sec-Segmente) mit der Intention, Inhalte zu vermitteln.

		Beschreibung auf der Beobachtungsebene: Die LP steht vor der Klasse und vermittelt Inhalte. Es kann vorkommen, dass die LP Inhalte bspw. mit Hilfe des Overheadprojektors erläutert. *Spezifische Kodierungsregeln:* Wichtig: Die Abgrenzung zu „Unterrichtsgespräch“: Häufig unterbrechen LP ihre Monologe durch Fragen an S, erhalten eine Antwort und reden wieder über einen längeren Zeitraum alleine. Kodierungsregel: Auf die Hauptintention achten. Wenn die LP vorwiegend darauf abzielt, Inhalte im Vortrag zu vermitteln, „LP-Vortrag“ kodieren. Wenn die LP vom Vortrag in ein Unterrichtsgespräch wechselt, zur Kodierung „Unterrichtsgespräch“ übergehen und umgekehrt. Spezialfall: Bei kurzen Vorträgen (z.B. zusammenfassenden Kommentaren, die nur 10 Sekunden dauern) bei Kodierung „Unterrichtsgespräch“ bleiben! Achtung: Die Kategorien von Facette „Strukturierung“ werden nicht als „LP-Vortrag“ kodiert.
2	**S-Vortrag**	*Inhaltliche Bestimmung:* S spricht über einen längeren Zeitraum (mindestens 2 x 10-sec-Segmente) mit der Intention, Inhalte zu vermitteln. *Beschreibung auf der Beobachtungsebene:* S erläutert für die ganze Klasse Inhalte. *Spezifische Kodierungsregeln:* In diese Kategorie fallen Phasen des Unterrichts, in denen S Texte zusammenfassen, Vorträge halten oder Resultate einer vorangegangenen Arbeitsphase vortragen. Abgrenzung zum „Unterrichtsgespräch“: Beantwortet S eine Frage der LP, wird dies nicht als „S-Vortrag“ kodiert, auch wenn die Ausführung länger dauert.
3	**Unterrichts-gespräch**	*Inhaltliche Bestimmung:* Die LP arbeitet mit der ganzen Klasse. Fragend-entwickelnd erarbeitet die LP mit den S bestimmte Inhalte (Konzepte und Begriffe). *Beschreibung auf der Beobachtungsebene:* Das Unterrichtsgespräch beginnt, wenn die LP explizit äußert, dass die Klasse ihre Aufmerksamkeit auf sie richten soll. *Spezifische Kodierungsregeln:* Typisches Frage-Antwort-Verhalten von LP und S. Das Unterrichtsgespräch ist auf fachliche Inhalte beschränkt. Das Nachfragen, ob bspw. alle fertig gelesen haben, wird nicht unter dieser Kategorie kodiert.
4	**Diktat / Abschreiben**	*Inhaltliche Bestimmung:* Entweder werden Texte von der LP diktiert oder die S schreiben sie ab. *Beschreibung auf der Beobachtungsebene:* LP diktiert einen Text, den S aufschreiben. Oder S schreiben einen Text z.B. von der WT oder aus dem Buch ab. *Spezifische Kodierungsregeln:* Keine.
5	**Diskussion**	*Inhaltliche Bestimmung:* Es findet ein Gespräch mit Diskussionscharakter statt. Gespräche dieser Kategorie sind von längerer Dauer (mind. 2 x 10-sec-Segmente). *Beschreibung auf der Beobachtungsebene:* Häufig findet diese Aktivität in einer speziellen Sitzordnung statt, z.B. Kreis. *Spezifische Kodierungsregeln:* Abgrenzung zu „Unterrichtsgespräch“: Es werden hier Meinungen über Inhalte ausgetauscht, wobei sich die Aussagen der S bzw. der LP aufeinander beziehen sollten. Es ist kein fragend-entwickelndes Unterrichtsgespräch, in dem fachliche Inhalte erarbeitet werden. Auf die Hauptintention der LP achten. Oft wird eine Diskussion durch eine entsprechende Aussage der LP eingeleitet: „Dieses Thema möchten wir nun miteinander diskutieren.“ *Beispiel:* *S: Also, zum Thema Gott: Ich glaube nicht unbedingt an Gott, aber etwas muss es ja eigentlich geben, dass es diese Welt überhaupt gibt.* *T: Bleiben wir mal hier. Sie glaubt nicht unbedingt dran, aber etwas muss es ja geben, weil die Welt existiert. Ja?* *SN: Ich habe …*

		T: Nein, du musst nicht lesen, was du abgeschrieben hast, sondern nimm einfach Stellung. Bist du mit ihrer Aussage einverstanden? *S: Nein, also ich – ich denke eher….*
6	**Film / Video**	*Inhaltliche Bestimmung:* Es wird im Klassenverband ein Filmausschnitt angeschaut. *Beschreibung auf der Beobachtungsebene:* S betrachten einen Film oder einen Videoausschnitt. *Spezifische Kodierungsregeln:* Keine.
7	**Lesen**	*Inhaltliche Bestimmung:* Die LP bzw. S lesen (=Vorlesen) einen Text der Klasse vor. *Beschreibung auf der Beobachtungsebene:* S bzw. die LP liest Text vor. Die Kategorie „Lesen“ wird auch kodiert, wenn es sich um einen eigenen Text handelt. *Spezifische Kodierungsregeln:* Abgrenzung zu „S-Vortrag“: Bei S-Vorträgen, die abgelesen werden, wird trotzdem „S-Vortrag“ kodiert.
8	**Spielformen**	*Inhaltliche Bestimmung:* In diese Kategorie gehören alle Spielformen, die im Klassenverband stattfinden. *Beschreibung auf der Beobachtungsebene:* Die Spielform ist eine spezielle Aktivität im Klassenunterricht, mit der Inhalte geübt oder vertieft werden. Dazu gehören auch Rollenspiele, Pantomime etc. *Spezifische Kodierungsregeln:* Keine.
9	**Mehrere Aktivitäten gleichzeitig**	*Inhaltliche Bestimmung:* In dieser Kategorie werden Unterrichtsphasen kodiert, in denen mehrere Aktivitäten gleichzeitig stattfinden. *Beschreibung auf der Beobachtungsebene:* Die LP ist mit einer Gruppe von S in einem Unterrichtsgespräch vertieft, währenddessen die anderen S bspw. mit einem Spiel beschäftigt sind. *Spezifische Kodierungsregeln:* Gibt die LP in dieser Phase einer Gruppe von S strukturierende Anweisungen, so wird dies unter „Lernorganisation während des Klassenunterrichts“ / Facette 1 „Strukturierung“ kodiert.
10	**S leitet Unterricht**	*Inhaltliche Bestimmung:* In dieser Kategorie werden Unterrichtsphasen kodiert, in denen S die Rolle der LP übernehmen (z.B. S stellt Fragen und nimmt Antworten entgegen und beurteilt diese). *Beschreibung auf der Beobachtungsebene:* Die LP übergibt die Leitung S, welche nun die Rolle der LP übernehmen. *Spezifische Kodierungsregeln:* Achtung: Hält S einen Vortrag oder liest etwas vor, so wird die entsprechende Kategorie 2 bzw. 7 gewählt.
11	**Abfrage von Resultaten**	*Inhaltliche Bestimmung:* Die LP befragt die S nach den Resultaten von Hausaufgaben oder Einzel- bzw. Gruppenarbeiten („Ping-Pong-Gesprächsführung“). Die Antworten der S dürfen allerdings nur maximal zwei Intervalle betragen, ansonsten wird „Lesen“ oder „S-Vortrag“ kodiert.
12	**Andere**	*Inhaltliche Bestimmung:* Diese Kategorie wird vergeben, wenn eine Zuordnung in die Kategorien 0 – 11 nicht möglich ist (z.B. Hausaufgaben erteilen).

Facette 5: Kommunikation im Unterricht		
0	**Keine**	Betrifft Videosequenzen vor und nach Unterrichtsbeginn.

1	**Niemand spricht**	*Inhaltliche Bestimmung:* Kommunikation bezieht sich nur auf verbale Kommunikation. Hier werden Intervalle kodiert, in denen weder die LP, noch S verbale Aussagen machen. Führt die LP Handlungen ohne Kommentar aus, wird auch „Niemand spricht“ kodiert. Die Kategorie „Niemand spricht“ wird als aktive und gleichwertige Kategorie von Kommunikation behandelt. Nimmt diese Kategorie im jeweiligen Intervall den größten Zeitwert ein, so wird sie auch kodiert, unabhängig davon ob auch LP- und / oder S-Aussagen im Intervall vorkommen. *Spezifische Kodierungsregeln:* „Niemand spricht“ kann nur kodiert werden, wenn überhaupt Unterricht stattfindet.
2	**LP spricht**	*Inhaltliche Bestimmung:* Die Funktionen der Lehreraussagen werden hier nicht unterschieden. Es geht lediglich um das Vorkommen verbaler Aussagen der LP. Die Zuordnung für die Lehreraussage erfolgt über den jeweils höheren Redeanteil im betreffenden Intervall. Die verbalen Aussagen müssen an die Klassenöffentlichkeit gerichtet sein. Beratungsgespräche mit einzelnen S während Gruppenarbeitsphasen werden hier nicht kodiert. Die Zuordnung zu den einzelnen Kategorien wird über den jeweils größeren zeitlichen Anteil vorgenommen. *Spezifische Kodierungsregeln:* Kann nur während des Klassenunterrichts kodiert werden. Sollten die Gesprächsanteile von LP und S im Intervall jeweils genau 5 Sekunden betragen, so richtet sich die Kodierung nach dem größeren Kommunikationsanteil im vorherigen oder nachfolgenden Intervall.
3	**S spricht**	*Inhaltliche Bestimmung:* Ein oder mehrere S machen verbale Aussagen oder stellen Fragen an die Klassenöffentlichkeit bzw. LP. Die Zuordnung für die Schüleraussage erfolgt über den jeweils höheren Redeanteil im betreffenden Intervall. Zwiegespräche unter einzelnen S während des Unterrichts werden hier nicht kodiert, es sei denn, sie finden intentional für den Unterricht nach Anweisung der LP in Gruppenarbeitsphasen statt. *Spezifische Kodierungsregeln:* Kann nur während des Klassenunterrichts kodiert werden. Sollten die Gesprächsanteile von LP und S im Intervall jeweils genau 5 Sekunden betragen, so richtet sich die Kodierung nach dem größeren Kommunikationsanteil im vorherigen oder nachfolgenden Intervall.
4	**Mischformen**	*Inhaltliche Beschreibung:* „Mischformen“ bezieht sich auf offene Unterrichtsformen bzw. die gleichzeitige Verbindung mehrerer Arbeitsformen oder Übergänge im Unterrichtsverlauf, in denen nicht genau unterschieden werden kann, ob die LP oder die S den jeweils größeren Redeanteil im Intervall haben. *Spezifische Kodierungsregeln:* Es kann beobachtet werden, dass S und LP bzw. mehrere S-Gruppen gleichzeitig reden.
5	**Andere**	*Inhaltliche Bestimmung:* Hier werden Intervalle kodiert, die nicht den anderen zugeordnet werden können. Beispiele sind Störungen des Unterrichts durch Dritte bzw. das Aufnahmeteam. *Spezifische Kodierungsregeln:* Die Kategorie „Andere“ kann nur vergeben werden, wenn „kein Unterricht / Unterbrechung“ kodiert wurde.

B Lernorganisation während des Klassenunterrichts

Facette 1: Strukturierung
Die Facette „Strukturierung“ hält fest, ob und wie die LP die Unterrichtsstunde strukturiert. Das Hauptaugenmerk liegt auf der Art der Strukturierung und nicht auf der damit verbundenen Aktivität.
In Abgrenzung zu den „Aktivitäten im Klassenunterricht“, wo fachliche Inhalte vermittelt werden, werden mit strukturierenden Äußerungen keine neuen Inhalte vermittelt,

sondern klärende Aussagen zur Lernorganisation gemacht. Achtung: Für die Kategorien 1 – 3 und 5 – 6 (nicht Kategorie 0 „Keine“ bzw. Kategorie 4 „Zusammenfassen“) der Facette „Strukturierung“ ist bei der Facette „Aktivitäten im Klassenunterricht“ generell Kategorie 12 „Andere“ zu kodieren. Eine Ausnahme innerhalb der Facette „Strukturierung“ stellt die Kategorie 4 „Zusammenfassen“ dar, weil es dort zentral um die Vermittlung von Inhalten geht; diese Vermittlung wird im Regelfall in Form des LP-Vortrags oder im Unterrichtsgespräch durchgeführt (vgl. spezifische Kodierungsregeln „Aktivitäten im Klassenunterricht“).		
0	**Keine**	Zuordnung in Kategorie 0, wenn keine strukturierenden Äußerungen gemacht werden.
1	**Ziel und Ablauf der Stunde**	*Inhaltliche Bestimmung:* „Ziel und Ablauf der Stunde“ wird kodiert, wenn der geplante Stundenablauf und die Ziele der Stunde transparent gemacht werden für die S und / oder wenn die Abfolge der Aktivitäten in der einzelnen Unterrichtsstunde den S vorgestellt wird. *Beschreibung auf der Beobachtungsebene:* Das Ziel und der Ablauf der Stunde werden von der LP genannt. Dies findet gewöhnlich bei Unterrichtsbeginn statt. Präzisierungen von Arbeitsaufträgen werden hier nicht kodiert. Die LP muss etwas für die S Neues nennen, ansonsten wird „Inhaltliche Arbeitsaufträge“ kodiert. *„Also werden wir schauen, wie haben es die Griechen gehalten. Was hatten sie für Vorstellungen und dann wollen wir diskutieren, was habt Ihr für Vorstellungen. Und hoffentlich sind die Vorstellungen unterschiedlich, dann kommt was raus.“* *Spezifische Kodierungsregeln:* Mündlicher Hinweis der LP auf den Verlauf notwendig. Der Hinweis muss sich auf die gesamte Stunde beziehen, nicht nur auf einzelne Arbeitsaufträge.
2	**Ziel und Ablauf weiterer Stunden**	*Inhaltliche Bestimmung:* Diese Kategorie umfasst Informationen über den Ablauf oder die Ziele künftiger Unterrichtsstunden. *Beschreibung auf der Beobachtungsebene:* Das Ziel der weiteren Stunden wird deutlich durch die Mitteilungen der LP an die S. *„Wenn wir im Frühjahr über den 2. Weltkrieg sprechen, werden wir uns…“* Die LP kann den Ablauf der nächsten Stunden z.B. mit folgender Äußerung erläutern: *„In der nächsten Stunde werden wir mit dem heutigen Thema weiterfahren. Wir werden dann die folgenden Dinge betrachten…“* *Spezifische Kodierungsregeln:* Keine.
3	**Inhaltliche Arbeitsaufträge**	*Inhaltliche Bestimmung:* Diese Kategorie umfasst Unterrichtsphasen, in denen Arbeitsaufträge angekündigt und erteilt werden. Zu den Arbeitsaufträgen gehören Aussagen über die Dauer, das Arbeitsmaterial und / oder den Arbeitsort. Es werden hier nur inhaltliche Arbeitsaufträge kodiert, die zu Einzel-, Partner- oder Gruppenarbeitsphasen hinführen. Arbeitsaufträge an die Klassenöffentlichkeit („Lest Euch das an der Tafel durch“) werden hier nicht kodiert. *Beschreibung auf der Beobachtungsebene:* Die LP vermittelt oder erklärt die Aufgaben der folgenden Arbeitsphase. *„Die Aufgabe ist jetzt hört zu und nehmt Stellung zum Inhalt. Nicht einfach Eures vorlesen, sondern Stellung nehmen.“* *Spezifische Kodierungsregeln:* Angaben über die zeitliche Dauer oder den Ort werden hier ebenfalls kodiert (auf Abgrenzung zu Ziel und Ablauf der Stunde achten). Gibt die LP innerhalb einer Stillarbeitsphase noch weitere Instruktionen, so werden diese nur als solche kodiert, wenn die LP die S ausdrücklich zum Zuhören auffordert (siehe dazu auch Facette 2 „Unterrichtliche Arbeitsformen“ Kategorie 7 „Störung durch die LP“). Stellen S während des Auftrages Fragen, die von der LP beantwortet werden, so wird „Inhaltliche Arbeitsaufträge“ gewählt. Späteres Nachfragen von Seiten der S zur Präzisierung des Arbeitsauftrages wird ebenfalls hier kodiert, sofern dies von der LP zum öffentlichen Thema gemacht wird.
4	**Zusammen-fassen**	*Inhaltliche Bestimmung:* Diese Kategorie umfasst Unterrichtsphasen, in denen zum Abschluss einer Unterrichtsstunde oder -einheit Inhalte zusammengefasst werden. Ziel und Zweck dieser Phase ist die inhaltliche Zusammenfassung gelernter Inhalte der erteilten Lektion.

		Beschreibung auf der Beobachtungsebene: Bspw. Zusammenfassung der besprochenen Inhalte und Experimente im Unterrichtsgespräch (als Wiederholung von Themen, die in der Stunde neu besprochen wurden), Aufschreiben der wichtigsten Inhalte, Hefteinträge sowie Lesen einer Zusammenfassung. Auf das übergeordnete Ziel achten und weniger auf einzelne Aktivitäten! *Spezifische Kodierungsregeln:* Achtung: Auf die Intention der LP achten. Es geht in dieser Kategorie um das strukturierende Zusammenfassen von Inhalten der Lektion. „Zusammenfassen" heisst auch, dass innerhalb der Stunde die LP bestimmte Inhalte noch einmal strukturiert und zusammenfasst. Das Lesen eines neuen Textes sowie ein allfälliges Zusammenfassen des gelesenen Textes gehören nicht in diese Kategorie. Kennzeichnend für diese Kategorie ist das übergeordnete Zusammenfassen auf einem höheren Abstraktionsniveau (im Vergleich zur „Rückschau", bei der ein gerichtetes Lernfeedback im Zentrum steht). „Zusammenfassen" wird dann kodiert, wenn funktional ersichtlich wird, dass die Aktivitäten der LP auch den Charakter einer Zusammenfassung für die S haben können, indem z.B. die Inhalte auf der Tafel oder dem Overheadprojektor am Ende der Stunde oder als Abschluss einer Arbeitsphase erläuternd angeschrieben werden. Auf rhetorische Besonderheiten achten (z.B. „Wir haben vorhin festgestellt…"), die verbal eine Zusammenfassung einläuten könnten, dies aber faktisch nicht tun, sondern vielmehr eine Form des lauten Denkens der LP darstellen. Wenn „Zusammenfassen" kodiert wird, muss auch Kategorie 3 „Zusammenfassung / Rückschau" der Facette „Allgemeindidaktische Unterrichtsphasen" kodiert werden.
5	**Rückschau**	*Inhaltliche Bestimmung:* Diese Kategorie bezieht sich auf die Arbeitsrückschau im Sinne Aeblis (2003). In dieser Phase gibt die LP den S eine Rückschau auf den individuellen bzw. gruppenbezogenen Lernverlauf in der Unterrichtsstunde. Ziel und Zweck ist es, den S Feedback über ihr Lernen zu geben (in Abgrenzung zur inhaltlichen Zusammenfassung des Gelernten). „Rückschau" wird auch dann kodiert, wenn die LP einzelne S bittet ein Feedback über den Lernverlauf anderer S zu geben. *Beschreibung auf der Beobachtungsebene:* Explizite Rückmeldungsrunde am Ende der Unterrichtsstunde. Beispiel: *„Ihr hattet ja am Anfang Probleme mit der Fragestellung, aber am Ende habt Ihr die Aufgaben gut gelöst."* *Spezifische Kodierungsregeln:* Abgrenzung zur inhaltlichen Zusammenfassung (z.B. Tafelbild, Hefteinträge usw.): Wichtig ist hier, dass die S explizit Rückmeldung über ihr Lernen in der Stunde erhalten (durch LP oder andere S). Stereotype Rückmeldungen (z.B. „sehr gut" oder „gut gemacht") reichen nicht aus, um „Rückschau" kodieren zu können. Die Kategorie „Rückschau" muss mindestens über zwei Intervalle kodiert werden können. Wenn „Rückschau" kodiert wird, muss auch Kategorie 3 der Facette „Allgemeindidaktische Unterrichtsphasen" kodiert werden.
6	**Andere**	Diese Kategorie wird vergeben, wenn eine Zuordnung in die Kategorien 0 – 5 nicht möglich ist. Anweisungen der LP, eine Still-, Partner- bzw. Gruppenarbeitsphase zu beenden, werden in diese Kategorie eingeteilt. Beispiel: *„Ich muss euch langsam unterbrechen. Ihr seid wie wahnsinnig am Schreiben. Macht ihr den Satz fertig? Den Rest könnt Ihr dann als Hausaufgabe fertig machen."* Wird eine Still-, Partner- bzw. Gruppenarbeitsphase allerdings durch die LP unterbrochen, ohne diese zu beenden, so gehört dies nicht in die Facette „Strukturierung", sondern wird unter Facette 4 „Unterrichtliche Arbeitsformen" Kategorie 7 „Störung der LP" kodiert.

Facette 2: Differenzierung Achtung: Es werden Segmente einer Still-, Partner- bzw. Gruppenarbeitsphase kodiert, in welchen eine Leistungsdifferenzierung stattfindet. Findet eine Differenzierung im Mehrklassenunterricht statt, so wird dies auch in den entsprechenden Kategorien 1 – 4 kodiert.

0	**Keine**	Wird immer dann kodiert, wenn keine Differenzierung stattfindet.
1	**Leistungs-differenzierung durch Selbst-einschätzung**	*Inhaltliche Bestimmung:* Die Kategorie „Leistungsdifferenzierung durch Selbsteinschätzung“ wird kodiert, wenn die S den Schwierigkeitsgrad der Aufgaben selber wählen können. Die Aufgabenwahl erfolgt durch Selbsteinschätzung. *Beschreibung auf der Beobachtungsebene:* Die S wählen aus verschiedenen Aufgaben aus. Die LP gibt eine Auswahl von Aufgaben mit verschiedenen Schwierigkeitsgraden vor, aus der die S selber auswählen können. *Spezifische Kodierungsregeln:* Keine.
2	**Leistungs-differenzierung durch LP**	*Inhaltliche Bestimmung:* Die Kategorie „Leistungsdifferenzierung durch LP“ wird kodiert, wenn die LP den S verschiedene Aufgaben gibt, die dem jeweiligen Leistungsvermögen der S entsprechen oder die gleiche Aufgabe verschiedenen Niveaugruppen zuordnet. Die Aufgabenwahl erfolgt durch die LP. *Beschreibung auf der Beobachtungsebene:* Die S erhalten individuelle Aufgaben, die ihrem Leistungsvermögen entsprechen. Die LP wählt dabei Aufgaben mit verschiedenen Schwierigkeitsgraden für die S aus. *Spezifische Kodierungsregeln:* Keine.
3	**Neigungs-differenzierung**	*Inhaltliche Bestimmung:* „Neigungsdifferenzierung“ wird kodiert, wenn die S Aufgaben nach ihren persönlichen Interessen wählen können. *Beschreibung auf der Beobachtungsebene:* Die S wählen aus verschiedenen Aufgaben aus. Die LP gibt eine Auswahl von Aufgaben mit verschiedenen Inhalten vor, aus der die S nach Interesse selber auswählen können. *Spezifische Kodierungsregeln:* Gibt die LP den Auftrag, dass die S in einer Werkstatt arbeiten dürfen, so wird dies unter „Neigungsdifferenzierung“ kodiert, so lange die LP keine spezifischen Hinweise gibt, die in Kategorie 1, 2 oder 4 kodiert werden können.
4	**Mischformen**	*Inhaltliche Bestimmung:* Die Differenzierung erfolgt in unterschiedlichen Formen. *Beschreibung auf der Beobachtungsebene:* Die Kategorie „Mischformen“ wird gewählt, wenn ein Teil der S die Aufgaben bspw. durch persönliche Neigung auswählen kann und ein Teil der S die Aufgaben durch die Einschätzung der LP erhält. *Spezifische Kodierungsregeln:* Keine.

Überblick über das Raterinventar

(in Anlehnung an Clausen, Reusser & Klieme 2003)

Allgemeine Grundregeln

- Das Rating wird vorgenommen auf einer sechsstufigen bipolaren Skala.
- Die Skalenbezeichnungen gehen von 1 („trifft überhaupt nicht zu“) bis 6 („trifft voll und ganz zu“). Als Ergebnis einer Vorstudie wird bewusst von der in Videostudien gängigen Praxis einer vierstufigen Skala abgewichen, um einen höheren Differenzierungsgrad zu erreichen.
- Sollten Bereiche nicht beobachtbar sein, so wird folgendermaßen vorgegangen:
    - beim Item „*Die LP geht kompetent mit Störungen um*“ wird eine 4 geratet
    - bei den Facetten „Umgang mit Fehlern“ und „Lehrperson als Mediator“ wird eine 3 geratet
- Die Videos werden erst kodiert und dann geratet.
- Zwischen dem Kodierverfahren und dem Rating liegt mindestens ein Zeitraum von einer Woche.
- Die Filme werden jeweils in einem Durchgang in voller Länge angesehen und dann anschließend in allen Dimensionen geratet.
- Es besteht die Möglichkeit, die Ratings in zwei Durchgängen vorzunehmen. Nachdem der Film in voller Länge angesehen wurde, werden die Facetten „Umgang mit Unterrichtszeit“, „Qualität der Organisation“, „Pacing“, „Sprachliche Qualität“ sowie „Zielklarheit“ geratet. Nachdem der Film erneut in voller Länge angesehen wurde, werden die restlichen Facetten geratet.
- Sollte ein Unterrichtsvideo zu zweit angeschaut werden, so müssen verbale und nonverbale Reaktionen unterbleiben.
- Die Unterrichtsvideos werden nach Möglichkeit von zwei Ratern getrennt bewertet.
- Unterschiedliche Bewertungen werden im Konsensverfahren geklärt.

Überblick

Dimensionen	Facetten	Items
A Instruktionseffizienz	Umgang mit der Unterrichtszeit	- *Die zeitliche Planung der Lehrperson ist erkennbar.* - *Die LP sorgt dafür, dass die S sich zu Beginn der Stunde ruhig verhalten und zur Mitarbeit bereit sind.* - *Die Wechsel zwischen den Arbeitsformen beanspruchen wenig Zeit.* - *Die LP geht kompetent mit Störungen um.*
	Qualität der Organisation	- *Das Arbeitsmaterial ist einsatzbereit.* - *Die LP setzt Medien (auch Materialien, Tafel und Visualisierungen) lernförderlich ein.* - *Die Organisation der offenen bzw. nicht lehrerzentrierten Unterrichtsformen erfolgt zweckmäßig.* - *Die LP variiert Methoden und Sozialformen in angemessener Weise.*
B Schülerorientierung	Umgang mit Fehlern	- *Die LP achtet darauf, dass Fehler nicht überbewertet werden.* - *Wegen fehlerhaften Verhaltens wird niemand bloßgestellt.* - *Bei dieser LP ist Fehlermachen nichts Schlimmes.*
	Motivierungsfähigkeit	- *Die LP unterrichtet tatkräftig und engagiert.* - *Die S arbeiten konzentriert und motiviert.* - *Die LP ermutigt einzelne S.* - *Die LP holt die Schüler/innen zur gemeinsamen Arbeit ab.* - *Die LP verhält sich den S gegenüber wohlwollend.*
C Kognitive Aktivierung	Pacing	- *Die LP geht im Stoff zügig voran, ohne die S zu überfordern.* - *Die S erhalten angemessen Zeit für die Bearbeitung einer Aufgabe.* - *Die LP unterstützt die S auch in den offenen Unterrichtsphasen.* - *Die LP passt ihren Unterricht den S an.*
	Lehrperson als Mediator	- *Die LP lässt den S Zeit.* - *Die LP unterstützt bei der Ausformulierung bzw. Ausführung von Ideen.* - *Die LP fragt – ohne zu bewerten – nach, wenn ein S unvollständige oder unklare Ideen formuliert.* - *Die LP unterstützt die S bei der Lösung von Fragen und Problemen.* - *Die LP gibt individuelle Rückmeldungen.* - *Die LP unterstützt mit Scaffolding, wo dies angebracht ist.* - *Die LP übernimmt die Funktion als Coach in offenen Unterrichtsphasen.*

D Klarheit und Struktu-riertheit	Sprachliche Qualität	- *Die LP spricht verständlich.* - *Die LP korrigiert eigene Fehler.* - *Die LP passt das Sprachniveau der Schulstufe an.*
	Gesprächsführung	- *Die LP vermeidet das permanente Wiederholen von S-Antworten („Lehrerecho“).* - *Die LP vermeidet stereotype Kommentare (z.B. „„sehr gut“, „okay“, „genau“).* - *Die LP verfügt über verschiedene Strategien mit S-Antworten umzugehen.* - *Die LP setzt Fragen didaktisch sinnvoll ein.* - *Die LP gibt aktivierende Impulse (verbal / nonverbal).*
	Zielklarheit	- *Die Ziele von Unterrichtsphasen oder der Lektion werden erkennbar.* - *Die Anweisungen und Aufträge der Lehrperson sind klar formuliert.* - *Der Zusammenhang zwischen Unterrichtsphasen ist erkennbar.*

A Instruktionseffizienz

Umgang mit der Unterrichtszeit
Die zeitliche Planung der LP ist erkennbar.
- Die LP behält die Initiative bei sich und die Zeit im Auge (z.B. wird nicht am Ende der Stunde ein neues Thema begonnen). - Die LP lässt sich durch Fragen der S nicht vom Thema abbringen.
Die LP sorgt dafür, dass die S sich zu Beginn der Stunde ruhig verhalten und zur Mitarbeit bereit sind.
- Der Unterrichtsbeginn erfolgt konzentriert und koordiniert. - Störungen behindern den Unterrichtsbeginn nicht. - Die S konzentrieren sich auf die LP. - Die LP benutzt Rituale effizient, um den Beginn der Stunde zu markieren.
Die Wechsel zwischen den Arbeitsformen beanspruchen wenig Zeit.
- Die LP führt den Wechsel zwischen Sozialformen rasch durch (z.B. Stuhlkreis zur Klassensitzordnung). - Die Materialien liegen bereit und die Verteilung erfolgt rasch. Die Übergänge sind kurz. - Während der Unterrichtsstunde finden keine aufwändigen Aufräumarbeiten statt. - Die LP teilt die Gruppen ein und lässt sich dabei nicht auf Diskussionen ein.
Die LP geht kompetent mit Störungen um. (Bezieht sich auf Störungen, die die Unterrichtsstunde offensichtlich stören.)
- Der Umgang mit Störungen beansprucht wenig Zeit. - Störungen bzw. Probleme werden in die entsprechende Instanz (z.B. Klassenratstunde) verlegt. - Die LP verfügt über ein großes Handlungsrepertoire (z.B. nicht immer die gleichen Reaktionen auf Störungen, nicht immer verbal). - Die LP interveniert bei Störungen und ignoriert diese nicht, wobei die Reaktionen angemessen sind.

Qualität der Organisation
Das Arbeitsmaterial ist einsatzbereit.
- Die Materialien liegen in ausreichender Anzahl und Qualität vor und sind einsatzbereit (z.B. Druckqualität, die Tafel ist gereinigt, Stifte sind angespitzt). - Das Material ist in einem Zustand, in dem man es benutzen kann. - Medien (z.B. Projektor) müssen nicht erst umständlich herbeigeholt werden.

Die LP setzt Medien (auch Materialien, Tafel, Visualisierungen) lernförderlich ein.
- Im Unterrichtsverlauf ist erkennbar, dass der Medieneinsatz didaktische Funktionen hat. - Medien werden eingesetzt, um z.B. eine Schüler-, Alltags-, Handlungs- oder Praxisorientierung herzustellen. - Medien werden lernförderlich eingesetzt, um durch Visualisierungseffekte Zusammenhänge und Funktionen transparenter zu machen. - Der Medieneinsatz ist im Kontext der Stunde eingebettet und steht nicht für sich alleine als Selbstzweck.
Die Organisation der offenen bzw. nicht lehrerzentrierten Unterrichtsformen erfolgt zweckmäßig.
- Die LP stellt Regeln für z.B. den Werkstattunterricht auf, welche die Abfolge der jeweiligen Phase der Werkstatt klären. - Laute und leise Stationen liegen nicht direkt nebeneinander. - Räumliche Gegebenheiten und Besonderheiten werden optimal ausgenutzt (z.B. akustische Besonderheiten). - Ergebnisse von offenen Unterrichtsphasen werden erst dann ins Reine übertragen bzw. eingeklebt, wenn sie von der LP überprüft worden sind.
Die LP variiert Methoden und Sozialformen in angemessener Weise.
- Die Methoden und Sozialformen sind den ausgewählten Inhalten angepasst. - Es ist eine angemessene Variation der Methoden und Sozialformen erkennbar; insofern erfolgt eine Adaptation an die Bedürfnisse der S. - Die Methoden und Sozialformen passen zueinander (z.B. nicht zu dritt ein Gedicht auswendig lernen lassen). - Die zeitliche Dimension (zu lang oder zu kurz) bzw. die Ortskomponente (wo wird gelernt) wird berücksichtigt. - Methoden und Sozialformen werden lernförderlich eingesetzt.

B Schülerorientierung

Umgang mit Fehlern In allen drei Bereichen geht es um eine positive Fehlerkultur, die von der LP aktiv vorangetrieben wird. Das Fehlen von diesbezüglich negativen Handlungselementen reicht nicht aus, um von einer positiven Fehlerkultur sprechen zu können.
Die LP achtet darauf, dass Fehler nicht überbewertet werden.
- Geben S falsche Antworten auf Fragen, bearbeiten Aufgabenstellungen in nicht korrekter Weise bzw. stellen Fragen, deren Antworten bereits gegeben worden sind oder die nicht in den Kontext der Stunde passen, so reagiert die LP trotzdem positiv verstärkend auf das S-Verhalten. - Die LP achtet darauf, dass Selbstvorwürfe von S nicht generalisiert werden und nimmt selbst keine negativen Generalisierungen vor. - Die LP nimmt aktiv Relativierungen bei Fehlern vor und macht dies auch vor den Aufgabenstellungen deutlich.
Wegen fehlerhaften Verhaltens wird niemand bloßgestellt.
- Die LP reagiert bei Fehlern der S nicht sarkastisch, zynisch oder lachend. - Wenn andere S ein solches Verhalten zeigen, nimmt sie Interventionen vor. - Die LP reagiert in unterstützender Weise, wenn ein/e S einer Aufgabenstellung vor der Klassenöffentlichkeit nicht gewachsen ist.
Bei dieser LP ist Fehlermachen nichts Schlimmes.
- Die LP sieht Fehlermachen als notwendigen Teil des Lernprozesses an. - Als Beobachterin bzw. Beobachter sieht man, dass offen mit Fehlern umgegangen wird. Niemand versucht Fehler zu vertuschen. - Wegen Fehlern muss sich niemand schämen. - S haben keine Angst davor im Unterricht mitzureden / mitzuarbeiten und dabei Fehler zu machen. - S, die Fehler vertuschen wollen, schwächen dabei ihre Aussagen im Voraus ab mit Äußerungen wie „sozusagen“, „irgendwie so“ oder „sag ich mal so“. - Sie orientieren sich auch sehr stark an den Erwartungen der LP, d.h. sie sind sehr sensitiv für Signale, die von der LP ausgehen. - Sie äußern Antworten wiederum als Fragen, um zu testen, ob ihre Antworten den Erwartungen der LP entsprechen.

Motivierungsfähigkeit
Die LP unterrichtet tatkräftig und engagiert.
- Die LP wirkt engagiert, tatkräftig und rege, hat dabei eine lebhafte Gestik und Stimmmodulation. - Die LP lässt erkennen, dass es ihr wichtig ist, den S etwas beizubringen. - Die LP hat ihren Unterricht genau vorbereitet. - Die LP führt keine Nebengespräche mit S, während ein anderer S etwas an der Tafel präsentiert. - Die LP lässt erkennen, dass sie sich für die Thematik oder die S interessiert.
Die S arbeiten konzentriert und motiviert.

- Es herrscht eine motivierte Stimmung. - Die S langweilen sich nicht. - Die S bearbeiten Aufgabenstellungen rasch und mit Einsatz und lenken sich nicht gegenseitig ab.
Die LP ermutigt einzelne S.
- Die LP motiviert S durch kompetenten Umgang mit Techniken des Lobes, der Ermutigung und der Bekräftigung. - Die LP gibt Beispiele aus der Erfahrungswelt der S, um deutlich zu machen, dass Probleme lösbar sind.
Die LP holt die S zur gemeinsamen Arbeit ab.
- Der Einstieg in die Lektion ist für alle S motivierend. - Nach Einzel-, Partner- oder Gruppenarbeiten werden die S aus der individuellen Arbeitsphase abgeholt. - Die Übergänge zwischen den Unterrichtsphasen verlaufen reibungslos, die Motivation nimmt dabei nicht ab. - Die Motivation hält durch die ganze Lektion hinweg an.
Die LP verhält sich den S gegenüber wohlwollend.
- Die LP ist den S zugewandt und verhält sich respektvoll, wird nicht ausfallend (z.B. Zynismus). - Auch auf wiederholte Fragen reagiert sie ruhig und freundlich. - Die LP beleidigt die S nicht. - Die LP gibt den S zu verstehen, dass sie ihnen die Lösung von Aufgaben zutraut (positive Verstärkung) und sich für die Meinungen der S interessiert (Nachfragen). - Mit „wohlwollendem Verhalten“ ist allerdings nicht eine infantilisierende Form der „Kuschelpädagogik“ gemeint, welche die S unterfordert.

C Kognitive Aktivierung

Pacing
Die LP geht im Stoff zügig voran, ohne die S zu überfordern.
- Die LP vermittelt die Unterrichtsinhalte in einem für die S passenden Tempo (der Unterrichtsverlauf ist weder zu träge, noch zu hektisch). - Es ist erkennbar, dass die S sich nicht langweilen bzw. keine Ängstlichkeitssymptome durch Überforderung zeigen (hektische Fragen an die LP oder den Sitzpartner, fragende Blicke).
Die S erhalten angemessen Zeit für die Bearbeitung einer Aufgabe.
- Die LP vermeidet unnötige Wiederholungen von Inhalten. - Die LP achtet darauf, dass nicht zu viel bzw. zu wenig Zeit für eine Unterrichtsphase verwendet wird.
Die LP unterstützt die S auch in den offenen Unterrichtsphasen.
- Die LP verfolgt die Arbeit der S mit Interesse, unterstützt mit neuen Inhalten und treibt den Lernprozess voran. - Die LP ist während der Still-, Partner- und Gruppenarbeitsphasen offen für Fragen und gibt konstruktive Antworten, um den Lernprozess zu unterstützen und voranzutreiben. - Es wird ein reibungsloser Unterrichtsablauf ermöglicht und der Lernprozess aktiv unterstützt.
Die LP passt ihren Unterricht den S an.
- Der Ablauf der Lektion erfolgt reibungslos bzw. „leicht von der Hand“ (bspw. ohne ständiges Ermahnen zur Arbeit der S durch die LP). - Die Lektion verläuft geschmeidig.

Lehrperson als Mediator In allen Bereichen geht es um eine aktive Unterstützung der S durch die LP. Das Fehlen von diesbezüglichen Handlungselementen muss demgemäß als negativ beurteilt werden.
Die LP lässt den S Zeit.
- Die LP verlangt nicht sofort eine Antwort auf eine Frage, sondern lässt den S Zeit, Ideen zu entwickeln und Antworten zu finden. - Die LP ermuntert die S, sich Zeit zu nehmen und nachzudenken. - Die LP bittet nicht immer den S, der sich als erstes meldet, um eine Antwort.
Die LP unterstützt bei der Ausformulierung bzw. Ausführung von Ideen.
- Die LP unterstützt die S entweder bei ihren Äußerungen im Klassenunterricht oder beim Lösen von Aufgabenstellungen in Gruppen-, Partner- oder Einzelarbeit. - Die LP vernetzt Ideen verschiedener S miteinander und gibt Fragen an andere S weiter. - Die LP ermuntert die S ihre Ideen weiterzuverfolgen.
Die LP fragt – ohne zu bewerten – nach, wenn ein S unvollständige oder unklare Ideen formuliert.
- Die LP liefert nicht sofort bei einer Schülerfrage eine Antwort oder Bewertung, sondern fragt erst einmal nach oder gibt die Frage an andere

S weiter. - Missverständliche, unvollständige oder unklare Äußerungen werden nicht ignoriert oder bloß kommentiert, sondern es wird nachgefragt. - Die LP reagiert wertneutral.
Die LP unterstützt die S bei der Lösung von Fragen und Problemen.
- Die LP nimmt die S ernst, hat Zeit für sie, weicht ihren Fragen nicht aus. - Sie bemüht sich um eine Antwort, ermuntert die S ihre Fragen zu stellen und mit anderen zu kommunizieren.
Die LP gibt individuelle Rückmeldungen.
- Als Hauptunterscheidung kann das Berücksichtigen der individuellen vs. der sozialen Bezugsnorm angenommen werden. - Individuelle Erfolge und Arbeitsergebnisse werden von der LP benannt. - Fragen werden individuell und nicht pauschal beantwortet. - Die LP gibt ein realitätskonformes Feedback, welches gemeinsam mit S reflektiert wird.
Die LP unterstützt mit Scaffolding, wo dies angebracht ist.
- Die LP macht auch in offenen Unterrichtsphasen didaktische Strukturierungen in Form instruktionaler Angebote („Lerngerüst“, „Scaffolding“) und elaboriert die Inhalte mit den S.
Die LP übernimmt die Funktion als Coach in offenen Unterrichtsphasen.
- Die LP steht während offenen Unterrichtsphasen für inhaltliche Fragen zur Verfügung. - Die LP beachtet gruppendynamische Prozesse und reagiert unterstützend darauf.

D Klarheit und Strukturiertheit

Sprachliche Qualität
Die LP spricht verständlich.
- Die LP spricht im Hinblick auf Deutlichkeit der Akzentuierung, Dialekt, Schnelligkeit und Lautstärke so, dass sie von allen S verstanden werden kann.
Die LP korrigiert eigene Fehler.
- Unterlaufen der LP bei der mündlichen oder schriftlichen Darstellung von Fachinhalten oder Aufgabenstellungen Fehler, so werden diese – entweder selbst aktiv handelnd oder nach Fragen von S – richtig gestellt. - Die LP spricht grammatikalisch korrekt bzw. korrigiert unterlaufene Fehler.
Die LP passt das Sprachniveau der Schulstufe an.
- Die LP passt sich dem Sprachniveau der S an, indem sie sich z.B. begrifflich oder in Form von Beispielen an der Lebenswelt der S orientiert. - Die LP erklärt neu eingeführte Begriffe und Fachtermini und macht deutlich, dass sie am diesbezüglichen Verständnis der S interessiert ist. - Bei Nachfragen von S erklärt sie den Sachverhalt auf einem einfacheren sprachlichen Niveau neu.

Gesprächsführung
Die LP vermeidet das permanente Wiederholen von S-Antworten.
- Sätze von S werden nicht wörtlich wiederholt („Lehrerecho“).
Die LP vermeidet stereotype Kommentare.
- Die Rückmeldungen der LP sind nicht wiederholend stereotyp und monoton (z.B. „sehr gut“, „okay“, „genau“).
Die LP verfügt über verschiedene Strategien mit Schülerantworten umzugehen.
- Im Unterrichtsverlauf ist erkennbar, dass die LP verschieden auf Schülerantworten reagiert und damit unterschiedliche Absichten für den Lehr-Lernprozess verfolgt. - Die LP verwendet Schülerantworten nicht nur als Stichwort für den weiteren Unterrichtsverlauf, sondern stellt z.B. Rückfragen, gibt die Fragen an andere S weiter, zeigt Zusammenhänge auf oder wiederholt z.B. eine experimentelle oder medial gestützte Erklärung.
Die LP setzt Fragen didaktisch sinnvoll ein.
- Aktivierender Einsatz von offenen bzw. geschlossenen Fragen. - Verwendung von echten Fragen (keine rhetorischen Fragen).

Die LP gibt aktivierende Impulse (verbal / nonverbal).
- Die LP gibt Aufforderungen und Befehle, durch welche die S zur Mitarbeit angeregt werden. - Anregender Einsatz von Ritualen (Triangel, Glocke, Gong). - Gestaltung von situativen Impulsen (z.B. Raumgestaltung). - Die LP konfrontiert die S mit verschiedenen Sichtweisen und Meinungen, um Widersprüche zu erzeugen.

Zielklarheit
Die Ziele von Unterrichtsphasen bzw. der Lektion werden erkennbar.
- Die Ziele der Lektion sind ersichtlich. - Die Ziele werden erreicht. - Die einzelnen Unterrichtsphasen sind zielorientiert.
Die Anweisungen und Aufträge der LP sind klar formuliert.
- Meint nur die inhaltliche Verständlichkeit der Aufträge. - Die Arbeitsaufträge sind prägnant, im Sinne von angemessener Länge und inhaltlicher Verständlichkeit. - Sowohl mündlich als auch schriftlich möglich. - Sowohl zu einzelnen als auch zur ganzen Klasse möglich. - Festzumachen auch an S-Reaktionen, z.B. Nachfragen.
Der Zusammenhang zwischen einzelnen Unterrichtsphasen ist klar erkennbar.
- Die Abfolge der Unterrichtsphasen erscheint kohärent, so dass eine sinnvolle Verbindung zwischen den Phasen erkennbar ist. - Zusammenhänge werden nicht erzwungen oder weit hergeholt. - Die LP verwendet Beispiele, die den Sachverhalt anschaulich machen und erklären. - In der Übungsphase wird das geübt, welches vorher besprochen wurde. - Unterrichtsphasen sind in ihrer Funktion voneinander unterscheidbar. - Die LP gibt am Anfang der Stunde einen Überblick und hält sich auch daran („roter Faden“).

Eidesstattliche Versicherung des Oliver Küster zur Dissertation „Praktika und ihre Lernpotentiale in der Lehrerbildung – Eine längsschnittliche Videostudie zur Untersuchung der Entwicklung unterrichtlicher Handlungskompetenzen in verlängerten Praxisphasen“

Belehrt, dass die Abgabe einer falschen Versicherung an Eides Statt strafbar sein kann, erkläre ich hiermit an Eides Statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich / unentgeltlich geholfen:

1. ……...
2. ……...
3. ……...

……..

Weitere Personen waren am Verfassen der vorliegenden Arbeit nicht beteiligt.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

(Ort, Datum)

……...

(Unterschrift)

……...

(Ort, Datum und Unterschrift des die Versicherung an Eides Staat aufnehmenden Beamten)