

## Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik

### Ansätze der TIMSS 1999 Video Studie und ihrer schweizerischen Erweiterung<sup>1</sup>

Dominik Petko, Monika Waldis, Christine Pauli und Kurt Reusser, Zürich (Switzerland)

**Abstract:** In the last decade video based classroom research was increasingly applied and further developed in the field of research on mathematics instruction. The main advantages and the methodological challenges of this approach are elaborated from a Swiss perspective on the research process of the TIMSS 1999 video study. Based on a systemic concept of teaching quality the crucial steps of data collection, data processing and analysis of the Swiss sample are briefly described and discussed.

**Kurzreferat:** Der Ansatz der videobasierten Unterrichtsforschung fand insbesondere im Fach Mathematik in den letzten Jahren intensiv Verwendung. Dabei kristallisierten sich sowohl die möglichen Vorteile dieses Zugangs als auch eine Reihe von methodologischen Herausforderungen heraus, die hier am Beispiel der TIMSS 1999 Videostudie aus schweizerischer Sicht beschrieben werden. Ausgehend von einem systemischen Konzept von Unterrichtsqualität werden bedeutsame Entscheidungen bei der Datenerhebung, Datenaufbereitung und Datenanalyse anhand der schweizerischen Stichprobe kurz dargestellt und diskutiert.

**ZDM-Classification:** D20

#### Weshalb videobasierte Unterrichtsforschung?

Videotechnologie ist mehr als vierzig Jahre nach ihrem ersten Auftreten in der Unterrichtsforschung wieder ein viel beachtetes Thema (Wild 2003b, Ulewicz & Beatty 2001, Aufschnaiter & Welzel 2001). Dieser Umstand kann auf technische und auf wissenschaftliche Entwicklungen zurückgeführt werden. Technische Entwicklungen führten zu einem Qualitätssprung der Videoaufzeichnung bei gleichzeitig sinkenden Kosten. Während in den 1970-er und 1980-er Jahren Schulklassen in hochtechnisierten Mitschulanlagen aufgezeichnet wurden, können heute mit vergleichsweise geringem finanziellem und technischem Aufwand digitale Videogeräte und Funkmikrophone in Klassenzimmern installiert werden. Die modernen Geräte liefern zudem eine technisch höherwertige Aufzeichnung als frühere Anlagen. Tonaufnahmen, die in Klassenzimmern ein besonderes Problem darstellen, lassen sich beispielsweise schon mit Standardgeräten

ohne störendes Rauschen realisieren. Computerunterstützte Videoformate erlauben zudem eine verbesserte Archivierung, Distribuierung, Bearbeitung und Analyse auch großer Datensätze mit Hilfe qualitativer Datenanalysesoftware (z.B. vPrism™, Videograph, Catmovie, ATLAS.ti).

Die erweiterten technischen Möglichkeiten wurden von einer Reihe von Autoren in ihrem Wert für die Unterrichtsforschung erkannt und in wissenschaftliche Entwicklungen überführt (Stigler 1998, Stigler, Gallimore & Hiebert 2000, Reusser, Pauli & Zollinger 1998, Stigler, Gonzales, Kawanaka, Knoll & Serrano 1999). Die Vorteile von Videoanalysen zeigen sich nach diesen Autoren vor allem angesichts der Grenzen von Fragebögen, Interviews und Unterrichtsbeobachtungen. Die auf diesen Wegen erhobenen Daten werden durch das jeweilige Erhebungsinstrument (Items und Skalen, Leitfaden, Protokollraster) und die dahinter liegenden theoretischen Annahmen und Vorüberlegungen in hohem Maße vorstrukturiert. Videodaten haben demgegenüber eine weniger subjekt- und theoriegebundene Qualität, da die analytischen Fragestellungen und Kategorien nicht bereits vor der Erhebung festgelegt werden müssen. Anhand von Videos sind grundsätzlich wiederholte Analysen mit unterschiedlichen Kriterien und Perspektiven, mit qualitativen oder quantitativen Verfahren, im Vergleich zu anderen Datenebenen und in Kooperation internationaler oder transdisziplinärer Teams möglich. Videoanalysen können, insbesondere in Kombination mit anderen Verfahren, die Komplexität von Unterrichtsprozessen, so die Hoffnung, am ehesten abbilden. Aufgrund ihrer hohen Anschaulichkeit ist es möglich, Ergebnisse von Analysen eindrücklicher an Beispielen zu illustrieren, als das mit Fragebögen oder Beobachtungsrastern möglich wäre. Videos erleichtern die fachsprachliche Verständigung und überbrücken möglicherweise die Kluft zwischen Theorie und Praxis. Videos, die im wissenschaftlichen Kontext erhoben wurden, können unter Umständen als Anschauungsmaterial für die Lehrerinnen- und Lehrerbildung dienen. Solchem Anschauungsmaterial muss gerade bei Kultur vergleichenden Studien ein hoher Wert beigemessen werden. Die eigene Praxis kann im Lichte des fremden Unterrichts neu hinterfragt und diskutiert werden. Dabei muss eine gemeinsame Sprache für die videografierten Phänomene gefunden werden, und zwar nicht nur im internationalen Forschungskontext, sondern auch in der Zusammenarbeit von Lehrpersonen, Lehrerbildnern, Fachdidaktikern, Allgemeindidaktikern und Erziehungswissenschaftlern innerhalb eines einzelnen Landes. Da digitale Videodaten langlebig sind, ist in zukünftigen Zeiten die Kontrastierung der eigenen mit der fremden Unterrichtspraxis auch in der historischen Rückschau möglich.

#### TIMSS 1999 Mathematik-Videostudie

Mit den internationalen Mathematikleistungsstudien, die Anfang der 70er Jahre mit FIMS begannen und nun mit PISA einen vorläufigen Höhepunkt erreichten, stellte sich auch verstärkt die Frage nach den *Hintergründen* der Schülerleistungen. Neben sozialen Faktoren stellt auch *Unterricht*, so die Hypothese, einen wichtigen Faktor dar.

<sup>1</sup> Das im vorliegenden Text im Zentrum stehende Projekt wurde vom Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung (Gesuch-Nr. 4033-054871) sowie von weiteren nationalen und internationalen Geldgebern unterstützt (Schweizerische Projektleitung: Kurt Reusser und Christine Pauli).

In der mathematikdidaktischen Forschung versuchten vor allem die groß angelegten internationalen Videostudien TIMSS 1995 Video (Stigler et al. 1999) und die noch erweiterte Nachfolgestudie TIMSS 1999 Video (Hiebert, Gallimore, Garnier, Bogard Givvin, Hollingsworth, Jacobs, Chui, Wearne, Smith, Kersting, Manaster, Tseng, Etterbeek, Manaster, Gonzales & Stigler 2003) Unterricht auf nationaler und internationaler Ebene zu beschreiben. Diese Projekte erhoben Videos nicht mehr nur zur Dokumentation von Einzelfällen bzw. ausgewählten Versuchs- und Kontrollklassen, sondern filmten im Sinne von „video surveys“ Hunderte von Videos als repräsentative Stichproben der jeweiligen Länder (Stigler et al. 2000).

Bei der TIMSS 1995 Videostudie wurden 230 Videos von Mathematiklektionen der 8. Klasse in Deutschland, Japan und den USA erhoben und analysiert. Dabei war Japan das einzige Hochleistungsland, Deutschland bewegte sich im Mittelfeld und die USA im unteren Drittel des Mathematikleistungsvergleichs. Die Ergebnisse der Videostudie zeigten, dass in Japan üblicherweise ein stärker explorierender Unterrichtsstil Anwendung fand, bei dem pro Lektion nur verhältnismäßig wenige Aufgaben mit hohem Niveau gelöst wurden. In Deutschland und noch stärker in den USA fand sich in Stillarbeitsphasen vor allem eine Beschäftigung mit prozeduralen Übungsaufgaben mit niedriger Komplexität. Diese kulturspezifischen Gestaltungsmuster von Unterricht schienen in Beziehung zu den Testleistungen der Schülerinnen und Schüler zu stehen.

Angeregt von den Ergebnissen wurde in der TIMSS 1999 Videostudie noch einmal ein repräsentatives Sample von Schulstunden der achten Jahrgangsstufe aus den USA und aus fünf weiteren Ländern gefilmt, die in der TIMSS 1995 Studie besonders gute Leistungen aufzuweisen hatten. Vertieft sollte dabei der Frage nach den Gemeinsamkeiten und Unterschieden in der Unterrichtsgestaltung dieser Länder nachgegangen werden. Neu wurden Australien, Hong Kong, Tschechien, die Niederlande und die Schweiz in die Studie einbezogen. Die Aufnahmen der japanischen Schulstunden wurden im Rahmen der TIMSS 1999 Videostudien noch einmal reanalysiert.

Fragestellung und Ziel der Studie war es, einen repräsentativen Querschnitt durch die Unterrichtswirklichkeit im Fach Mathematik auf der achten Jahrgangsstufe jeder teilnehmenden Nation zu gewinnen. Dieses Ziel ist zunächst ein rein deskriptives. In Bezug auf die TIMSS Leistungsdaten ergeben sich jedoch weiter gehende Schlussfolgerungen hinsichtlich des mit diesem Unterricht erreichten nationalen Leistungsniveaus. Aus US-amerikanischer Sicht handelt es sich bei der Studie um ein Benchmarking des amerikanischen Unterrichts mit der Unterrichtspraxis in sechs anderen, in Leistungstests dieser Jahrgangsstufe weltweit führenden Ländern. Der Vergleich untereinander ist jedoch auch für die anderen teilnehmenden Länder von Interesse. Über diese Ziele hinaus verfolgte die Studie auch das methodologische Ziel der Entwicklung von geeigneten Analysestrategien, mittels derer eine Videostudie dieses Umfangs bewältigt werden kann.

Insgesamt umfasst der Datenkorpus der TIMSS 1999 Studie 638 zufällig ausgewählte Schulstunden, d.h. zwi-

schen 50 und 100 pro Land. Die Schweiz war, aufgrund des Bestrebens, einen Vergleich der drei großen Sprachregionen (Deutschschweiz, französischsprachige und italienischsprachige Schweiz) zu erreichen, mit 140 Videoaufnahmen vertreten. Es wurde eine Mathematiklektion pro Klasse auf Video aufgezeichnet. Die Aufnahmen streuten über das gesamte Schuljahr. Im Rahmen der internationalen Studie wurden neben den Videoaufnahmen auch ein Lehrerfragebogen und ein kurzer Schülerfragebogen eingesetzt. Diese Instrumente erhoben vor allem Hintergrundinformationen zur gefilmten Lektion.

Die Größe und Internationalität dieser Studie stellt ein Novum dar. Die Auswertung der umfangreichen Videodaten, deren Basisauswertung allein für das schweizerische Sample annähernd 3000 Transkriptions- und 10000 Kodierstunden erforderte, war nur unter Ausreizung der technischen Entwicklungen, insbesondere der computerunterstützten Datenanalyse möglich. Für die Auswertungen wurde in den USA eigens das Software-Tool vPrism™ entwickelt, das Video und Transkript zeitsynchron wiederzugeben in der Lage ist und das eine zeitgenaue Kodierung sowie den Export der Kodierungen in ein statistisch auswertbares Tabellenformat erlaubt. Die Herausforderungen waren jedoch nicht nur technischer, sondern auch inhaltlicher Art. Die internationale Anlage verunmöglichte eine einfache Kategorisierung der Videodaten aus der Sicht einer einzelnen Kultur. Statt dessen mussten Kulturkreis übergreifende Kriterien der Unterrichtsbeobachtung entwickelt und eine gemeinsame Sprache gefunden werden. Die Beschäftigung mit Videodaten führte so zu einer Entselbstverständlichung der eigenen Sichtweise auf Unterricht. Dabei hat sich auch das Bewusstsein um die Komplexität der Zusammenhänge zwischen systemischen Einflüssen, Lernvoraussetzungen, Unterrichts- bzw. Lernprozessen und den resultierenden Bildungswirkungen erhöht.

Die Ergebnisse dieser Untersuchung (Hiebert et al. 2003) werfen ein neues Licht auf die Ergebnisse der TIMSS 1995 Vorgängerstudie (Stigler et al. 1999). Es ergaben sich eine Reihe von zentralen Ähnlichkeiten zwischen den Ländern, mit gleichzeitigen Abweichungen in Einzelaspekten. Alle untersuchten Länder teilen das Merkmal, dass 80% der Unterrichtszeit mit Aufgabenlösen verbracht wird, in einem Wechsel von Klassen- und Einzelarbeit. Die überwiegende Mehrheit der Lektionen besteht aus einer je eigenen Abfolge von Repetition, Einführung neuer Inhalte, und vertiefender Verarbeitung, Übung oder Anwendung. In über 90% der Lektionen wird mit einem Lehrmittel gearbeitet. Lehrpersonen reden in allen Ländern, auch dies eine Gemeinsamkeit, ungefähr acht mal mehr als alle Schüler zusammen. Unterschiede zwischen den Ländern ergaben sich unter anderem bezüglich der Beschaffenheit mathematischer Aufgaben, vor allem hinsichtlich ihrer Komplexität, Kohärenz, fachdidaktischen Qualität und ihres Alltagsbezugs.

Der Ertrag war jedoch nicht nur inhaltlicher, sondern auch methodischer und methodologischer Art. In der Schweiz wurden, einem systemischen Verständnis von Unterrichtsqualität folgend, zusätzliche nationale Erhebungen durchgeführt (Reusser & Pauli 2003). Hinzu kamen ein umfangreicher Schülerfragebogen und ein erweiterter Lehrerfragebogen, welche Aspekte des Ma-

thematikunterrichts im Allgemeinen thematisierten. Außerdem wurde zu zwei Zeitpunkten, im Abstand von einem Jahr, der TIMSS-Leistungstest durchgeführt und ein kognitiver Fähigkeitstest bei den Schülerinnen- und Schülern eingesetzt. Zusätzlich zu den 140 repräsentativen Klassen wurden noch 16 Klassen gefilmt, deren Lehrpersonen nach eigener Auskunft regelmäßig erweiterte Lehr- und Lernformen einsetzten, die den Schülerinnen und Schülern verstärkte Möglichkeiten eigenverantworteten Lernens ermöglichen. Das systemische Verständnis von Unterrichtsqualität, das zu diesen Erweiterungen führte, wird nachfolgend kurz erläutert.

**Unterrichtsforschung – Vom Prozess-Produkt-Paradigma zum systemischen Ansatz der Unterrichtsforschung**

Zur Erläuterung des Forschungsverständnisses der schweizerischen Videostudie soll zunächst kurz auf die Entwicklung der Forschungsparadigmen in der Unterrichtsforschung eingegangen werden.

In den 1970-er und 1980-er Jahren hat sich eine Vielzahl empirischer Untersuchungen zu Unterrichtsqualität und Bildungswirkungen an einem einfachen Prozess-Produkt-Modell orientiert. Diese Designs waren darauf angelegt, *direkte Beziehungen* zwischen dem alltäglichen Lehrerhandeln im Klassenzimmer – traditionellerweise erfasst mittels Beobachtungsdaten – und den Schülerleistungen, welche mittels Leistungstests erhoben wurden, herzustellen. Die unzähligen Studien brachten zunächst wenig überzeugende und z.T. widersprüchliche Ergebnisse hervor, dies nicht zuletzt darum, weil nach relativ globalen Persönlichkeitseigenschaften von Lehrpersonen

(vgl. Terhart 1997) beziehungsweise in wenig theoriegeleiteter Weise nach isolierten didaktischen Qualitätsmerkmalen des Unterrichts gesucht wurde. Die ab den 1980-er Jahren in Metaanalysen zusammengestellten Ergebnisse sind etwas ergiebiger: In Bezug auf (hohe) Fachleistungen verweisen sie auf die Bedeutung eines störungsfreien Verlaufs von Unterrichtsstunden, einer intensiv genutzten Unterrichtszeit, klaren und gut strukturierten Lehrerklärungen und individuellen Hilfestellungen im Sinne der direkten Instruktion (Wang, Haertel & Walberg 1993). Einher gehend mit der Etablierung kognitiver Lerntheorien in der pädagogischen Psychologie, wodurch die Bedeutung von Mediatorvariablen hervorgehoben wurde (Shuell 1996), und später mit dem Einbezug systemtheoretischer Überlegungen in die Arbeiten der empirischen Schulforschung (z.B. Fend 1998) wurde klar, dass Bildungswirkungen nicht mit einfachen Input-Output-Modellen erfasst werden können. Bildungswirkungen, wie sie beispielsweise in internationalen Schulleistungsvergleichen wie TIMSS oder PISA erfasst werden, müssen als Ergebnis des Zusammenwirkens unterschiedlicher Faktoren in einem Bildungssystem verstanden werden. Die Aufgabe der Unterrichtsforschung ist es demnach, empirisch gehaltvolle Modelle der Wechselwirkung von Faktoren innerhalb und zwischen verschiedenen Ebenen eines Bildungssystems zu finden (vgl. auch Helmke 2003). Das der schweizerischen Studie zugrunde liegende Modell (Reusser & Pauli 1999, 2003) illustriert das Zusammenspiel multipler Faktoren und bettet das Unterrichtsgeschehen in den erweiterten Kontext des Bildungssystems ein (vgl. Abb.1).

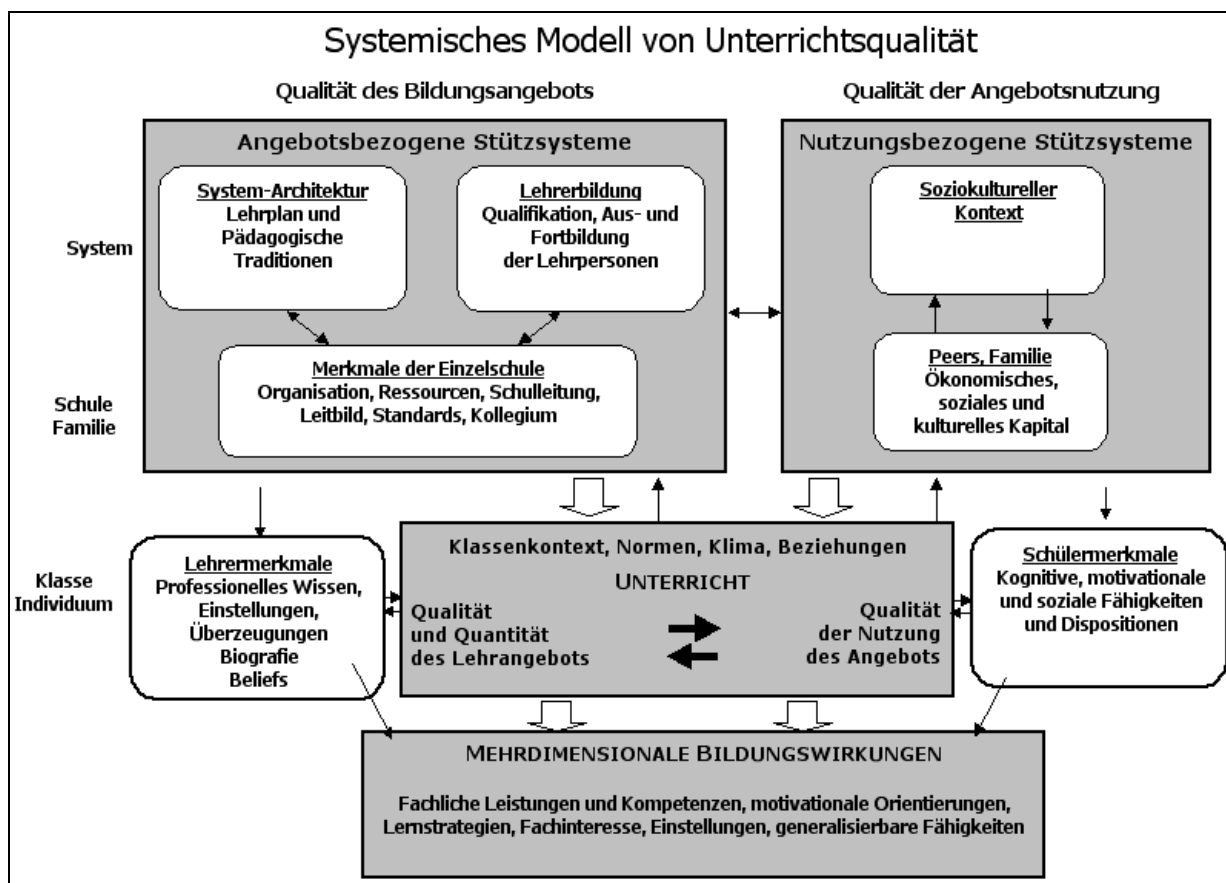


Abbildung 1: Systemisches Modell von Unterrichtsqualität (Reusser & Pauli 1999, 2003)

Im Zentrum dieses Modells steht der Gedanke des Zusammenspiels von Angebots- und Nutzungsseite (Fend, 1998) und der damit verbundenen Mediationsprozesse. Beide Seiten werden wiederum von verschiedenen systemimmanenten Voraussetzungen – den so genannten Stützsystemen – beeinflusst bzw. geprägt.

Idealerweise ergibt sich aus diesem Modell ein Forschungsdesign, in welchem folgende Punkte berücksichtigt werden (Reusser 2001):

- der Einbezug von Daten, welche Ausprägungen auf verschiedenen Ebenen des Bildungssystems erfassen: System, Schule, Klassenzimmer und Individuum
- die Analyse multipler Perspektiven auf Unterricht: die Innensicht von Lehrpersonen, von Schülerinnen und Schülern sowie die Außensicht von externen Experten
- die Erfassung von verschiedenen Typen des Lehrerwissens: fachliches, curriculares, fachspezifisch-pädagogisches und pädagogisch-didaktisches Wissen zu Unterrichtsgestaltung und Lernprozessen (Bromme 1997), z.T. repräsentiert in subjektiven Theorien, welche sich wiederum im unterrichtlichen Handeln niederschlagen (Dann 1997, Wahl 1991),
- die Verwendung verschiedener Methoden bei der Datenerhebung und Datenanalyse (Videodaten, schriftliche Befragungen und Tests; hoch-inferenter und niedrig-inferenter Zugang zu den Videodaten, etc.)
- die Untersuchung verschiedener schulischer Entwicklungskriterien, wie z.B. der Einbezug von Fachleistungen, Fach- und Sachinteresse, fachlichem Selbstkonzept und Leistungsmotivation
- die Bearbeitung von fachdidaktischen, allgemeindidaktischen, lernstoffbezogenen und pädagogischen Fragestellungen und Forschungshypothesen

Der hier formulierte Anspruch ist außerordentlich anspruchsvoll und kann schwerlich in einem einzelnen Forschungsprojekt eingelöst werden. Außerdem bedingt die Realisierung dieses Anspruchs transdisziplinäre Forschungsteams und Kooperationen zwischen Spezialisten aus verschiedenen Teilgebieten.

Der Fokus der *internationalen* TIMSS 1999 Video Studie richtete sich auf die Erforschung der beobachtbaren Unterrichtsaktivitäten in einer ökologisch validen Erhebungssituation. Nur am Rande wurden einige zusätzliche Informationen zu den dahinter liegenden Stützsystemen erfasst. Der Erkenntnisgewinn resultiert hier insbesondere aus dem internationalen Vergleich von Daten zur konkreten Unterrichtsgestaltung. Die *schweizerische Erweiterung* des Projekts zielte deshalb darauf ab, wesentlich erweiterte Daten vor allem von den Akteuren des Unterrichtsgeschehens (Lehrpersonen, Schülerinnen und Schüler) zu erfassen und diese zu den Videodaten in Beziehung zu setzen. In Bezug auf die Datenvielfalt, den Einsatz verschiedener Methoden, die erhobenen Perspektiven und die erfassten Bildungswirkungen konnte einiges im schweizerischen Projekt umgesetzt werden. In einem gegenwärtig laufenden Folgeprojekt, einer schweizerisch-deutschen Mathematik-Videostudie, wird dieser umfassende Ansatz weiter ausgebaut (Klieme 1999, Reusser & Pauli 2000, Klieme & Reusser 2003).

### Methodologische Herausforderungen

Im Rahmen der TIMSS 1999 Video Studie und ihrer schweizerischen Erweiterung stellten sich eine Reihe von methodologischen Fragen, auf die Antworten gefunden werden mussten. Die nachfolgenden Ausführungen skizzieren das Spektrum der wichtigsten methodologischen Probleme und präsentieren die im Projekt, auch unter Berücksichtigung forschungsökonomischer Gesichtspunkte, gewählten Lösungen aus internationaler (Jacobs, Garnier, Gallimore, Hollingsworth, Bogard Givvin, Rust, Kawanaka, Smith, Wearne, Manaster, Etterbeek, Hiebert, Stigler & Gonzales 2003) und schweizerischer Perspektive (Reusser & Pauli 2003). Die Darstellung gliedert sich in drei Teile: Datenerhebung, Datenaufbereitung und Datenanalyse.

#### Datenerhebung

Die folgenden Ausführungen zu methodologischen Fragen der Datenerhebung beziehen sich vor allem auf die Erhebung von Videodaten. Auf eine Darstellung der in der schweizerischen Studie ebenfalls sehr wichtigen Erhebungen von Befragungs- und Testdaten wird weitgehend verzichtet.

#### Repräsentativität und Validität

Im Gegensatz zu anderen Unterrichtsdokumentationen, bei denen die Videodaten entweder unsystematisch oder nach qualitativen Samplingmethoden (Kelle & Kluge 1999) erhoben werden, wurde bei TIMSS Video 1995 und später auch bei der TIMSS 1999 Videostudie mit einer repräsentativen Landesstichprobe gearbeitet. Für das ausgewählte Land wurde eine stratifizierte Zufallsstichprobe von Schulen gezogen, welche die Verteilung der Schultypen und der Schulgrößen in der Grundgesamtheit systematisch repräsentiert. Im Anschluss an die Stichprobenziehung wurden im ersten Schritt die Schulen, im zweiten Schritt eine wiederum zufällig ermittelte Lehrperson der jeweiligen Schule um Teilnahme an der Studie angefragt. Pro Schule kam nur eine zufällig ausgewählte Lehrperson zur Teilnahme an der Studie in Frage. Im Falle einer Absage wurde die zum Voraus bestimmte Ersatzschule kontaktiert. So war es der einzelnen Schule verunmöglicht, die Lehrperson ihrer Wahl für die Studienteilnahme vorzuschlagen. Da die Lehrperson eine Teilnahme an der TIMSS 1999 Studie auch ablehnen konnte, ist allerdings ein gewisser Selbstselektionsprozess bei den angefragten Lehrpersonen nicht auszuschließen, braucht es doch etwas Mut und eine gewisse Sicherheit im Umgang mit den Schülerinnen und Schülern, um sich vor die Kamera zu stellen. Ob man sich tatsächlich frei fühlte, auch abzusagen, ist jedoch wiederum von der jeweiligen (Schul-)Kultur eines Landes abhängig. So variiert denn auch der Anteil der Ersatzschulen in der Schweiz von Region zu Region<sup>2</sup>.

Die Frage nach der Validität von Videodaten kann letztlich jedoch nur unter Berücksichtigung des Ziels der

<sup>2</sup> In der deutschsprachigen Schweiz waren beispielsweise annähernd zwei Drittel der angefragten Schulen bereit, an der Studie teilzunehmen, in der französischsprachigen und italienischsprachigen Schweiz war die Bereitschaft zur Teilnahme deutlich höher (Jacobs et al. 2003, 40f.).

Studie abschließend beurteilt werden. Dies betrifft Fragen des Studiendesigns wie Anzahl der gefilmten Lehrpersonen, Anzahl gefilmter Lektionen pro Lehrperson sowie Standardisierung der Inhalte. Soll beispielsweise ein valides und reliables Bild einer individuellen Lehrperson gezeichnet werden, so ist die Aufzeichnung mehrerer Unterrichtslektionen im gleichen Fach notwendig. Auf diese Weise kann z.B. das Methodenrepertoire einer Lehrperson erfasst werden. Wenn das Studienziel darauf gerichtet ist, ein Bild der Unterrichtskultur in einer bestimmten Schule oder gar auf nationaler Ebene zu gewinnen, so ist es wichtig, dass eine genügend große Anzahl von verschiedenen Lehrpersonen gefilmt wird. Nur auf diese Weise können verlässliche Durchschnittswerte gewonnen werden. Im Extremfall wird – wie dies bei den beiden TIMSS Video Studien der Fall war – eine Lektion pro Lehrperson videografiert. Die auf diese Art und Weise gesammelten Daten sind als „national sample“ zu interpretieren und auf keinen Fall als repräsentativ für den Unterricht einer einzelnen Lehrperson zu betrachten (Stigler et al. 2000).

Hinsichtlich der Repräsentativität der gefilmten (Lektions-) Inhalte bieten sich ebenfalls verschiedene Möglichkeiten an. Naheliegender wäre die Vorgabe des fachlichen Themas der gefilmten Lektion. Damit wird die Vergleichbarkeit der einzelnen Lektionen untereinander erhöht. Wird die untersuchte Jahrgangsstufe konstant gehalten, bedingt dies allerdings, dass das gewählte Thema tatsächlich mit den Jahresvorgaben (Lehrpläne bzw. Lehrmittel) der untersuchten Schultypen kompatibel ist. Im stark föderalistisch organisierten Schulsystem der Schweiz hat es sich beispielsweise als ausgesprochen schwierig herausgestellt, in vier Kantonen des Mittellandes (Bern, Aargau, Zürich und Thurgau) zwei mathematische Themen zu identifizieren, welche in allen drei Schultypen (Gymnasium, Sekundar- und Realschule) in der 8. Klasse behandelt werden<sup>3</sup>.

Dieses Problems gewahr, wurde im internationalen Projekt die Ausdehnung des Erhebungszeitpunktes über ein Schuljahr beschlossen und die Wahl der Lektionsinhalte und –themen den Lehrpersonen überlassen<sup>4</sup>. Anstelle der Standardisierung der Inhalte wurde demnach eine zufällige Auswahl von Lektionen in einem zum voraus festgelegten Zeitrahmen<sup>5</sup> gefilmt. Unter der Voraussetzung einer genügend großen Stichprobe kann mit einem solchen Vorgehen ein relativ breites Spektrum der in Lehrplänen und Lehrmitteln festgelegten Unterrichtsinhalte der ausgewählten Jahrgangsstufe erfasst werden.

Die bewusst in Kauf genommene Variation der Inhalte hat allerdings zur Folge, dass einzelne Lektionen nicht mehr direkt miteinander verglichen werden können. Bei der Analyse der Videodaten kann es in diesem Falle nicht mehr darum gehen, einzelne Unterrichtsaktivitäten, die eng an bestimmte Inhalte geknüpft sind, miteinander zu kontrastieren. Vielmehr geht es darum, ein Analyseraster

zu finden, das es ermöglicht, Mathematikunterricht lehrstoffübergreifend zu beschreiben<sup>6</sup>.

#### *Datenschutz*

Üblicherweise wird Datenschutz im Zuge der Aufbereitung der Daten geleistet, etwa durch die Anonymisierung von Namen. In einer Videostudie ist dieses Verfahren nicht hinreichend. Eine anonymisierte Transkription bildet ein komplexes Geschehen wie den Schulunterricht nur unzureichend ab. Kommunikation im Unterricht verläuft nicht nur sprachlich, sondern auch via Tafel, Hellraumprojektor, Arbeitsblätter oder Hefte. Auch Aspekte wie Gestik, Mimik und Prosodie können nur mit großem Aufwand transkribiert werden. Deshalb musste in den Analysen der TIMSS 1999 Videostudie neben den anonymisierten Transkripten immer auch mit den Primärdaten, den Videoaufzeichnungen, gearbeitet werden. Gesichter von Lehrpersonen, Schülerinnen und Schülern bleiben im Video erkennbar und damit grundsätzlich identifizierbar. Eine Anonymisierung von Videos wäre nur unter großem technischem Aufwand machbar und würde das Video in seiner Aussagekraft deutlich einschränken. Der Datenschutz kann deshalb nur bis zu einem gewissen Grad mit der Datenaufbereitung geleistet werden. Personenbezogene Daten müssen vor allem durch verantwortliche Regelungen des Zugangs zu den Daten geschützt werden.

Der Schutz personenbezogener Daten ist bei videogestützter Forschung nicht nur eine ethische, sondern auch eine rechtliche Frage. Die Rechtslage ist wenig explizit und schwer überschaubar. Das Filmen zu wissenschaftlichen Zwecken kann grundsätzlich als Eingriff in die Privatsphäre von Menschen gewertet werden, die durch die Europäische Menschenrechtskonvention (Art. 8) geschützt ist. Bei videogestützten Unterrichtsstudien ist deshalb eine ausdrückliche Absicherung des Datenschutzes auf mehreren Ebenen anzuraten. Bildungsverantwortliche, d.h. Schulbehörden, Lehrpersonen und Erziehungsberechtigte müssen der Aufzeichnung und der zweckgebundenen Handhabung der Daten ausdrücklich, d.h. schriftlich zustimmen. In der schweizerischen Studie wurden die teilnehmenden Lehrpersonen und Eltern in mehreren Stufen um ihre Zustimmung zur Verwendung

<sup>3</sup> Es wurden Lehrplan- und Lehrmittelanalysen durchgeführt.

<sup>4</sup> Die Lehrpersonen wurden darauf aufmerksam gemacht, dass bei ihnen eine alltägliche, der persönlichen Unterrichtsplanung folgende Mathematiklektion gefilmt werden sollte.

<sup>5</sup> In der TIMSS 1999 Videostudie mussten mit der Datenerhebung mindestens 8 Monate abgedeckt werden.

<sup>6</sup> Eine andere Strategie wurde in einem binationalen Forschungsprojekt, das zur Zeit an der Universität Zürich und der Universität Frankfurt durchgeführt wird, gewählt (Projektleitung: Prof. Dr. K. Reusser, Dr. C. Pauli, Prof. Dr. E. Klieme). In jeweils 20 Klassen aus Deutschland und der Schweiz werden über ein Schuljahr Lehr-Lernprozesse unter anderem mit Videoaufzeichnungen untersucht. Im Rahmen eines quasi-experimentellen Designs wurden über zwei Lektionseinheiten in der Dauer von je drei Unterrichtsstunden die Themen („Satzgruppe des Pythagoras“ sowie „Algebraische Textaufgaben“) und die didaktisch-methodischen Unterrichtsmerkmale (Verwendung von Beweisen, Arbeit in Gruppen, Unterricht im Klassenverband) variiert. Der längsschnittliche Charakter der Studie erlaubt mikrogenetische Untersuchungen zur Entwicklung des mathematischen Verständnisses innerhalb eines spezifischen Teilgebiets der Mathematik. Die Festlegung der Inhalte ermöglicht die vergleichende Analyse inhaltsspezifischer Prozesse auf Klassenebene (vgl. Klieme & Reusser 2003).

ihrer Daten angefragt.

Lehrpersonen und Erziehungsberechtigten wurde schriftlich versichert, dass die Aufnahmen nur von wissenschaftlichem Personal betrachtet werden und dass keine personen-, klassen- oder schulidentifizierenden Daten veröffentlicht werden. Im Gegenzug unterschrieben sowohl Lehrpersonen wie auch die gesetzlichen Vertreter der Schülerinnen und Schüler (in der Regel die Eltern) eine diesbezügliche Einverständniserklärung. Das wissenschaftliche Personal unterzeichnete ebenfalls eine Erklärung, in der es sich verpflichtete, keine personenbezogenen Informationen aus den Daten weiterzugeben.

Eine Auswahl von videografierten Lektionen bzw. von Teilen besonders interessanter Lektionen sollen der Lehrerbildung und weiteren Kreisen für Dokumentations-, Informations-, Aus- und Weiterbildungszwecke zugänglich gemacht werden. Hierfür war eine erweiterte Datenschutzerklärung nötig. Für eine vollständig öffentliche Verwendung wurde schließlich ein gesondertes Set von Lektionen gefilmt, das lediglich als Anschauungsmaterial Verwendung findet. Hier wurden Lehrpersonen und Eltern ausdrücklich um die Erlaubnis jeglicher Verwendung, z.B. auch einer Publikation im Internet, angefragt. Es wurde jedoch ebenfalls darauf geachtet, dass sämtliche für den Verwendungszweck nicht unbedingt nötigen Informationen nicht öffentlich werden.

Zentrales Element sämtlicher Stufen des Datenschutzes war die schriftliche Vereinbarung mit den Verantwortlichen und deren strikte Einhaltung.

#### *Invasivität und Kameraeffekte*

Bei den Videoaufnahmen der TIMSS 1999 Videostudie handelte es sich um offene Beobachtungen. Zwar haben diese, nachdem die Frage des Datenschutzes geklärt wurde, nicht die ethischen Probleme einer verdeckten Beobachtung (Ellgring 1995), jedoch stellt sich die Frage, ob der Umstand, dass eine Videoaufnahme durchgeführt wird, den gefilmten Unterricht beeinflusst. Da die TIMSS 1999 Videostudie den Anspruch hatte, mit Hilfe ihrer Videoaufnahmen den *alltäglichen* Unterricht in den Schulklassen der teilnehmenden Länder repräsentativ zu dokumentieren, muss über mögliche Effekte der Erhebungssituation nachgedacht und nach Strategien zur Vermeidung von Invasivität gesucht werden.

Im Vorfeld wurden die Lehrpersonen in einem Schreiben ausdrücklich über das Interesse der Studie informiert, „eine alltägliche, wie üblich vorbereitete Mathematiklektion“ aufzuzeichnen. Die Studie begegnete mit diesem expliziten Hinweis bereits im Vorfeld einer möglichen Verzerrung.

Weitere Invasivität vermindern Maßnahmen boten sich im Moment des Filmens an. Durch zurückhaltendes Verhalten der Kamerapersonen konnte die Störung des Unterrichts auf ein Minimum reduziert werden. Die Filmer waren instruiert, sich möglichst unaufdringlich zu verhalten, sich nicht auf Interaktionen mit der Klasse einzulassen und sich vordergründig auf die Handhabung des Equipments zu konzentrieren. Auf diese Weise sollte das Interesse der Lernenden an der Kameraperson in Grenzen gehalten werden. Die Videoaufnahmen fanden im normalen Klassenzimmer der Schülerinnen und Schüler statt. Im Gegensatz zu früheren Formen der

Unterrichtsdokumentation mussten die Klassen nicht in ein ungewohntes, technisch ausgestattetes Klassenzimmer umziehen. Auch dies war ein Beitrag zur Verminderung möglicher Kameraeffekte.

Angesichts der hohen Zahl der teilnehmenden Klassen war ein Vorgehen, das bei kleinen Studien mit Erfolg angewendet wurde, nämlich die Klassen zuerst an die Situation der Videoaufnahme zu gewöhnen, bevor nach einigen Lektionen die eigentliche Erhebungssituation stattfindet, nicht realisierbar.

Schließlich wurden die Lehrpersonen mittels eines Fragebogens zur Typikalität der gefilmten Lektion befragt. Die Einschätzungen bezogen sich auf die Üblichkeit und das Anspruchsniveau des Unterrichtsstoffes, auf die gewählten Lehrformen und das Verhalten der Schülerinnen und Schüler (Hiebert et al. 2003). In der Schweiz berichteten 88% der videografierten Lehrpersonen, dass sie die gezeigten Lehrformen häufig oder fast immer verwendeten. Das Schülerverhalten schätzten 74% als normal ein, 21% als besser und 6% als schlechter als üblich. 88% der gefilmten Lehrpersonen gaben zudem an, das Anspruchsniveau der Inhalte sei nicht wesentlich anders als üblicherweise, 10% meinten, es sei in der gefilmten Lektion höher und nur 2% siedelten es tiefer an. In einer umfassenden Einschätzung gaben 86% der schweizerischen Lehrpersonen an, die Lektion sei wie üblich verlaufen, 7% schätzten den Verlauf als besser und 6% als schlechter als normal ein. Die Vorbereitungszeit war durchschnittlich mit 36 Minuten gegenüber den üblichen 32 Minuten leicht erhöht. Verglichen mit den Typikalitätswerten der anderen teilnehmenden Länder erreicht die Schweiz damit gute Werte, d.h. dass schweizerische Lehrpersonen die gefilmten Lektionen in ihrer überwiegenden Mehrheit als typisch für ihren eigenen Unterricht charakterisieren.

Insgesamt scheint das Problem des „Kameraeffekts“ weniger gewichtig zu sein als vermutet werden könnte. Zunächst einmal besteht kaum ein Grund zur Annahme, dass Lehrpersonen ihren Unterrichtsstil ohne weiteres von einer Stunde zur nächsten ändern können, haben doch einschlägige empirische Forschungsarbeiten gezeigt, wie schwierig Änderungen im Handeln von Lehrpersonen herbeizuführen sind (z.B. Wahl 1991). Mit der Verwendung von Lehrmethoden, die im üblichen Unterricht keine Verwendung finden, geht die Lehrperson zudem ein gewisses Risiko ein. Schülerinnen und Schüler könnten nicht verstehen, was von ihnen verlangt wird und könnten die ungewohnte Situation sogar kommentieren, was im Video sichtbar wäre. Außerdem ist aufgrund der Indizien nicht davon auszugehen, dass durch die bloße Anwesenheit einer Kamera eine signifikante Verbesserung oder Verschlechterung des Unterrichts bewirkt wurde. Ein spontaner Umlernprozess bei Lehrpersonen aufgrund der Filmsituation ist nicht zu erwarten. Allerdings muss angenommen werden, dass die Lehrpersonen in der Tendenz versuchten, eine aus persönlicher Sicht „gute“ bzw. „attraktive“ Lektion zu halten. Aus diesem Grund bieten die Videos nach Einschätzung von Stigler (1998, 141) eine „...somewhat idealized version of what the teacher normally does in the classroom“.

*Selektivität und Standardisierung der Aufzeichnung*

Es ist ein verbreitetes Missverständnis, dass Videodaten in der Lage seien, sämtliche Prozesse, die im Klassenraum ablaufen, zu dokumentieren. Entgegen dieser theoretischen Hoffnung prägen in der Praxis häufig die Limitiertheit des Bildausschnitts und die oft mangelhafte Tonqualität die spätere Auswertbarkeit. Für videogestützte Unterrichtsdokumentation gilt, ebenso wie für andere beobachtende Verfahren, die von Evertson & Green (1986) formulierte Aussage:

„The selectivity issue operates beyond the level of perspective or approach. Selectivity is an inevitable characteristic of any tool, representational system, or program of research. It is not possible to record all aspects of reality with any given system or tool or in any single research project“ (p. 165).

Kameras können im Klassenraum immer nur einen bestimmten Bildausschnitt festhalten. Gleiches gilt, sogar noch in verschärftem Maße, für den Ton. Es wäre nur mit sehr großem technischem Aufwand möglich, sämtliche Gespräche innerhalb eines Klassenzimmers aufzuzeichnen. Die Bestückung eines normalen Klassenzimmers mit dem dafür nötigen technischen Equipment würde das latente Problem der Invasivität zudem in hohem Maße verstärken. Video- und Tonaufzeichnungen müssen sich deshalb auf bestimmte Aspekte des Geschehens im Klassenzimmer ausrichten. Hall (2000) argumentiert, dass die notwendige Selektivität dazu führt, dass auch die scheinbar neutralen Daten letztendlich theoriegeladen sind. Je nach Wahl des Bildausschnitts geraten bestimmte Aspekte in den Blick und andere werden ausgeblendet.

In der TIMSS 1999 Videostudie stand das Problem der Selektivität zudem im Zusammenhang mit der Frage der Vergleichbarkeit der Aufnahmen aus unterschiedlichen Ländern. Die Aufnahmen wurden jeweils von einheimischen Kamerapersonen durchgeführt. Mit der Standardisierung der Kameraprozeduren durch ein Kameraskript wurde versucht, die Entscheidungen beim Zoomen und Schwenken der Kamera in bestimmten Momenten des Unterrichts zu vereinheitlichen (Jacobs et al. 2003, Appendix D).

Das zum Einsatz gekommene *Kameraskript* kann in groben Zügen wie folgt umrissen werden: Der Unterricht wurde von zwei Kameras gefilmt. Eine Kamera dokumentierte das Klassengeschehen statisch im Überblick aus einer vorderen Ecke des Klassenraums, üblicherweise neben der Tafel. Die andere Kamera, welche von einer Kameraperson dynamisch geführt wurde, stand idealerweise auf halber Höhe der Fensterseite. Diese Kamera hatte primär die Aufgabe, die Lehrperson in ihrer Interaktion mit den Schülerinnen und Schülern zu dokumentieren. Leitgedanke bei diesen Aufnahmen war das Prinzip „Folge immer der Lehrperson“ und die sogenannte „Zone der Interaktion“, nach der die Lehrperson mit den Schülerinnen und Schülern, denen sie sich situativ zuwendet, im Bildausschnitt zu sehen sein sollte. Darüber hinaus wurde das Kameraverhalten in einer Reihe von Ausnahme- und Sondersituationen bestimmt, ohne dabei eine restlose Standardisierung anzustreben. Das Skript enthält demzufolge auch Spielräume, da keineswegs zu erwarten ist, dass sämtliche Situationen mittels Vorgaben abgedeckt werden können.

Das Skript stellt einen Balanceakt zwischen Standardi-

sierung und Offenheit dar. Standardisierung führt zu größerer Vergleichbarkeit, beinhaltet aber auch die Gefahr, dass interessante Aspekte des Unterrichts systematisch ausgeblendet werden. Bei geringerer Standardisierung leidet die Auswertbarkeit und insbesondere die Vergleichbarkeit bestimmter Aspekte.

*Datenaufbereitung*

Die Prozesse der Datenaufbereitung umfassen die sorgsame Dokumentation jedes Datenelements und jeder Datenebene, die Kontrolle des Rücklaufes und die Archivierung. Dabei stellen sich vor allem methodische Probleme, die jedoch grundsätzlich mit Hilfe einer Datenbank lösbar sind. Methodologische Probleme bei der Aufbereitung von Videodaten entstehen vor allem bei der Digitalisierung und Transkription und der Einbettung in eine Analysesoftware.

*Digitalisierung und Transkription*

Das sehr speicheraufwendige Digital-Video-Format (DV), mit dem die Videos aufgezeichnet wurden, musste zur computerunterstützten Transkription und Auswertung zunächst in ein stark komprimiertes, ebenfalls computerlesbares Format umgewandelt werden. Ziel der Komprimierung war einerseits eine Reduktion der zu speichernden Datenmenge und andererseits die Umwandlung in ein Format, das von der Auswertungssoftware unterstützt wurde. Die Datenmenge einer Unterrichtslektion (ca. 9 Gigabyte) wurde auf die Speicherkapazität eines handelsüblichen CD-Rohlings (ca. 650 Megabyte) komprimiert. Die Komprimierung erfolgte im MPEG1-Format. Die dabei entstandenen Verluste an Bild- und Tonqualität erlauben nach wie vor eine komfortable Betrachtung und Analyse des Datenmaterials. Die Entscheidung für dieses Kompressionsformat richtete sich auch nach dem Aspekt seiner Integrierbarkeit in die Auswertungssoftware vPrism™, innerhalb derer sich die Daten auch transkribieren lassen.

Die Aufbereitung von audiovisuell aufgezeichneten Daten erfolgt üblicherweise durch Transkription, d.h. durch die detaillierte Verschriftlichung der für die Auswertung zentralen Informationen. Dies sind in der Regel vor allem die aufgezeichneten Redebeiträge, die in unterschiedlicher Genauigkeit wörtlich festgehalten werden, aber auch weitere Geschehnisse, die zusätzlich protokolliert werden. Transkription bietet die Möglichkeit für eine analytische Verlangsamung des Geschehens bei gleichzeitiger Beibehaltung der Linearität der Ereignisse. Durch die Verschriftlichung wird den Ereignissen ihre Flüchtigkeit genommen. Dennoch sind Transkriptionen immer selektiv (Brinker & Sager 2001, Dittmar 2002). Neben die Identifikation von Sprechern und gesprochenen Wörtern, können beispielsweise auch Pausen, Überlappungen, Phonetik, Prosodie, nonverbale Signale und Kontextinformationen treten. Es ist praktisch unmöglich, sämtliche auf Video aufgezeichneten Informationen in einem Transkriptionssystem abzubilden. Die Reduzierung der Information muss deshalb in Zusammenhang mit der späteren Analyse gesehen werden (Flick 1995). Einerseits sollten unter forschungswirtschaftlichen Gesichtspunkten möglichst nur jene Informationen transkribiert werden, die für die Analyse wertvoll sind, andererseits muss



vermieden werden, dass eine sich im weiteren Analyseverfahren als unerlässlich herausstellende Information nicht bereits im Transkriptionsprozess systematisch ausgeblendet wird.

Dieses Dilemma wurde in der TIMSS 1999 Videostudie durch die Entwicklung der Software vPrism™ entschärft (Stigler 1998, Knoll & Stigler 1999). Diese Software macht es neben anderen Funktionen möglich, Videodaten mit Abschnitten des Transkriptes so zu verknüpfen, dass beim Abspielen des Videobandes auf dem Bildschirm zugleich die entsprechende Stelle des Transkriptes angezeigt wird. Dieselbe Funktion ermöglicht eine sekunden-schnelle Navigation in den Videodaten, indem über das Anklicken einer bestimmten Stelle im Transkript das Video an genau diesem Punkt gestartet wird. Da der Bezug zur primären Datenebene stets erhalten blieb, war es möglich, eine verhältnismäßig einfache Transkriptionsweise zu wählen. Die in der TIMSS 1999 Video Studie transkribierten Informationen umfassen den Zeitpunkt des Beginns jedes Redezeuges, die Identifikation des jeweiligen Sprechers und eine orthographisch korrigierte, wörtliche Wiedergabe des betreffenden Redebeitrages (Jacobs et al. 2003, Appendix A). Zusätzlich wurden auch weitere Kommunikationsmerkmale vermerkt, die zum Verständnis unerlässlich sind, beispielsweise Überlappungen oder Pausen. Eine besondere Herausforderung stellte die Tatsache dar, dass Unterrichtsvideos in acht verschiedenen Unterrichtssprachen vorlagen. Mit Ausnahme der schweizerischen Lektionen wurden die Lektionen in Los Angeles in englischer Sprache transkribiert. Zugleich wurden Kodierpersonen verpflichtet, die mit der Originalsprache der von ihnen kodierten Videos muttersprachlich vertraut waren. Der Bezug von englischem Transkript und Originalton wurde durch die Software vPrism™ gewährleistet. Die schweizerischen Videos wurden in der Schweiz bearbeitet. Die Transkriptionen erfolgten in der jeweiligen Unterrichtssprache (deutsch, italienisch, französisch). Auch hier wurden die Videos von muttersprachlichem Transkribier- und Kodierpersonal bearbeitet. Die Bedeutung von vPrism™ lag hier in der Ergänzung nicht-transkribierter Information.

#### *Computerunterstützte Datenanalyse*

Neben vPrism™ existieren heute eine Reihe weiterer Softwarepakete, die Video in Verbindung mit Transkripten zu verarbeiten in der Lage sind. Dies sind vor allem Catmovie (Wild 2003a), Videograph (Rimmele 2003) und, mit Einschränkung, auch ATLAS.ti (Muhr 1994). Während vPrism™ eine Software für Apple Macintosh ist, laufen die übrigen Programme auf Microsoft Windows. Die Grundfunktionalität sämtlicher Programme ist ähnlich. Sie erlauben es, Videodaten abzuspielen und sie mit einem Transkript zu koppeln. Bei Catmovie, Videograph und vPrism™ ist diese Koppelung sehr komfortabel gelöst, ATLAS.ti hingegen erlaubt dies nur mit größerem Aufwand. Die eigentliche Auswertungsfunktion (zur Datenkodierung) liegt in der Möglichkeit der Abgrenzung von Sequenzen innerhalb des Videomaterials und der Zuordnung dieser Sequenzen zu einer beliebigen Menge von Kategorien. Die auf diese Weise getätigten Analysen werden im Hintergrund in einer Datenbank gespeichert

und können zu einer quantitativen Auswertung in ein Tabellenformat exportiert werden. Es bestehen beträchtliche Unterschiede zwischen den Programmen in der Handhabung der Analyseeinheiten und der Kategorien. Während Catmovie die Videodaten gleichmäßig in beliebig lange Auswertungseinheiten unterteilt und auch Videograph ein solches time-sampling in besonderer Weise unterstützt, können bzw. müssen in vPrism™ und ATLAS.ti die Auswertungseinheiten für jede Analyseeinheit neu spezifiziert werden. Die Handhabung der Kategorien unterscheidet sich ebenfalls. Für Catmovie ist es von Vorteil, die Auswertungskategorien vor Beginn der Analysen festzulegen, da entsprechende Softwareprogrammierungen vor Auswertungsbeginn vorgenommen werden müssen. Bei Videograph und vPrism™ können während der Analysen weitere Kategorien hinzugefügt werden. Die flexibelste Handhabung erlaubt ATLAS.ti, wo Kategorien während der Analyse geordnet oder zusammengefasst werden können. ATLAS.ti erlaubt auch ein Arbeiten mit Texten, Audiodateien und Bildern. Grundsätzlich lässt sich mit diesen Softwarelösungen nahezu jede beliebige Auswertungsweise von Videodaten realisieren. Es ist jedoch zu empfehlen, die Software zu wählen, die das eigene Vorgehen am besten unterstützt. Gerade in der Arbeit mit großen Datenmengen summiert sich jeder überflüssige Mausklick zu einer großen Anzahl von Arbeitsstunden. Anzuraten ist auch ein frühzeitiges, probeweises Exportieren der Daten in ein Tabellenformat und erste statistische Analysen derselben, da auf diese Weise Fehler bei der Kodierung im Hinblick auf das Auswertungsinteresse frühzeitig vermieden werden können.

#### **Datenanalyse**

Die Datenanalyse von Videoaufzeichnungen mit dem Ziel der Quantifizierung von deskriptiven Aussagen umfasst zunächst die Schritte der Bestimmung der Analyseeinheit, der Bildung eines Kodiersystems und dessen reliabler Anwendung. Hierbei können Kodierungen bzw. Ratings unterschiedlicher Reichweite zum Tragen kommen. Besondere Herausforderungen liegen weiterhin in der Aggregation der Analysen zu sinnvollen Aussagen und in der Verbindung von Videoanalysen mit anderen Datenquellen.

#### *Die Wahl der Analyseeinheit*

Videodaten können unter unterschiedlichen Fragestellungen analysiert werden, die jeweils eine entsprechende Wahl der Analyseeinheit nach sich ziehen. Wenn die Wahl der Analyseeinheit für eine Fragestellung zu klein oder zu groß ausfällt, lassen sich aus den Kategorisierungen nur schwerlich sinnvolle Aussagen gewinnen. Bei der Auswertung von Videodaten durch ein analytisches Raster, d.h. durch Zuordnung von Kategorien zu Ereignissen, kommt der Bestimmung der Analyseeinheit deshalb eine Schlüsselrolle zu. Die Spanne der Herangehensweisen liegt zwischen den Extremen der kleinschrittigen Zergliederung in einzelne Sprecherwechsel und der Gliederung in längere Sequenzen, bis zu ganzen Lektionen. Ein Beispiel für kleinschrittige Zergliederung bieten klassische Interaktionsanalysen, in denen jeder Redebeitrag in der Unterrichtsbeobachtung bzw. -aufzeichnung



einzelnen kodiert wird (vgl. Hanke, Mandl & Prell 1973). Ein derartiges Vorgehen wurde auch in der TIMSS 1995 Videostudie von verschiedenen Forschergruppen angewendet (Schümer 1997), wobei insbesondere für fachdidaktische Fragestellungen die Schwäche dieses Ansatzes deutlich hervortrat. Der Gruppe um Stigler gelang es zwar, bei der Analyse der Klassenzimmerinteraktion jede einzelne Sprechaktivität (turn) der Lehrperson und der Schülerinnen und Schüler klar zu kategorisieren und die durchschnittlichen Auftretenshäufigkeiten der einzelnen Kategorien zu bestimmen, es konnten aber keinerlei Aussagen über die Gesprächsabfolge oder die Verbindung zum inhaltlichen Kontext gemacht werden (Stigler et al. 2000).

Im Gegensatz zu einer derart kleinen Analyseeinheit bieten sich grundsätzlich auch ganze Lektionen als naheliegende Einheit an, allerdings sind damit auch gewisse Unschärfen verbunden. So kann zwar durch eine nominale Kategorie (0/1) ausgedrückt werden, ob ein bestimmtes Element innerhalb einer Lektion auftritt oder nicht, es können jedoch keine Aussagen über die Häufigkeit und den Umfang des Auftretens gemacht werden. Auch Aussagen über zeitliche Erstreckungen oder über die Eingebundenheit einer Unterrichtshandlung in einen bestimmten Kontext sind nur begrenzt möglich.

Die Sequenzierung der Lektion in kleinere Analyseeinheiten, die mehrere Sprecheraktivitäten (turns) umfassen, hat sich in Bezug auf allgemein- und fachdidaktische Fragestellungen als praktikabel erwiesen. Diese Analyseeinheiten sind weiter gefasst als einzelne Redebeiträge und flexibler eingrenzbar als ganze Lektionen.

Die wichtigste Analyseeinheit in der TIMSS 1999 Videostudie ist die Zeitspanne, in der eine mathematische Aufgabe („mathematical problem“) bearbeitet wird<sup>7</sup>. Um eine Lektion in Analyseeinheiten zu *segmentieren*, die zeitgenaue Aussagen über Spannen der Aufgabenbearbeitung zulassen, können grundsätzlich zwei Strategien angewendet werden, das *time-sampling* oder das *event-sampling* (vgl. Bakeman & Gottman 1994)<sup>8</sup>. Im *time-sampling* wird die gesamte Lektion in gleich lange Zeitabschnitte gegliedert, beispielsweise in 10-Sekunden-Schritte. Auf jede dieser 10 Sekunden langen Einheiten werden sämtliche Kodierkategorien in ihrer jeweiligen Ausprägung angewendet. Das *event-sampling*, die Lösung der TIMSS 1999 Videostudie, unterteilt die Lektion in eine Reihe unterschiedlich großer Zeitabschnitte mit jeweils eigenem, präzise bestimmtem Anfangs- und Endpunkt für jedes Auftreten des Kriteriums. Im Gegensatz zum *time-sampling* wird auf diese Weise die Abgrenzung von Kodiereinheiten derselben Kategorie (z.B. einzelner Aufgaben) gegeneinander präzise erfasst.

Komplexe Interaktions- und Sachstrukturen erschweren die klare Abgrenzung von gegenstandsbezogenen Analyseeinheiten. Das methodologische Problem der gegenstandsbezogenen Abgrenzung von Sequenzen zeigt sich

<sup>7</sup> Der Problembegriff wurde in der internationalen Studie sehr weit gefasst und findet seine adäquate Übersetzung im deutschen Ausdruck „mathematische Aufgabe“.

<sup>8</sup> Diese Herangehensweisen kennzeichnen üblicherweise verschiedene Intervalle der Datenerhebung bei Beobachtungsstudien. Sie erhalten eine etwas andere Bedeutung, wenn sie auf die Datenanalyse von Videodaten angewendet werden.

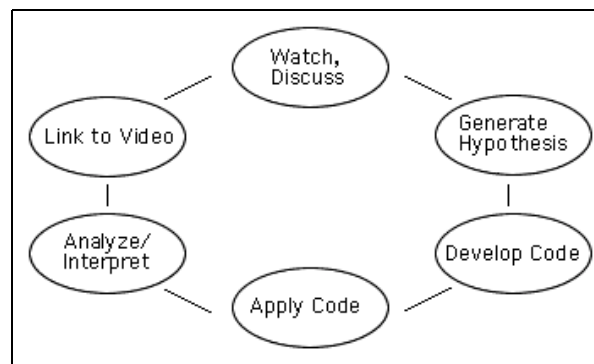
beispielsweise, wenn eine Aufgabenbearbeitung durch eine thematisch verbundene, jedoch eigenständige Aufgabe unterbrochen wird und die Lösung der Rahmenaufgabe erst nach Bearbeitung der eingeschobenen Aufgabe fortgesetzt wird.

Die Lösung solcher Probleme lag in der TIMSS 1999 Videostudie im sukzessiven Aufbau des Kategoriensystems, das bei klar abgrenzbaren Phänomenen beginnt und auf diesen Kodierentscheidungen komplexere Kodierungen aufbaut.

#### *Bildung eines Kodiersystems*

Durch die internationale Anlage der TIMSS 1999 Videostudie ergab sich das Problem der Entwicklung eines Kodiersystems, das einen länder- und kulturübergreifenden Vergleich von Mathematiklektionen erlaubt. Unterricht, so die zugrunde liegende Überzeugung, ist immer in eine bestimmte Kultur eingebunden und die beobachtbaren Lehr- und Lernaktivitäten können nur auf dem Hintergrund dieser Kultur adäquat interpretiert werden. Der Vergleich verschiedener Unterrichtskulturen konnte deshalb nicht aus der einseitigen Sichtweise je nationaler Forschungsteams erfolgen, sondern es mussten kulturübergreifende Kriterien der Beschreibung erarbeitet werden. Dieses kulturübergreifende „Grounding“, d.h. die Verständigung auf eine gemeinsame Sichtweise, erfolgte in einer multinationalen Kategorienentwicklergruppe in Los Angeles.

Die Kategorienentwicklung war als zyklischer Prozess konzipiert. Videosurveys profitieren von dem Vorteil, dass die vorhandenen Daten sowohl eine qualitative Beschreibung als auch, aufgrund des umfangreichen Datensatzes, eine Quantifizierung und statistische Reliabilitätsprüfung der beschriebenen Phänomene zulassen (Jacobs, Kawanaka & Stigler 1999, 719, vgl. Abb. 2).



**Abbildung 2: Kategorienentwicklungszyklus bei der Analyse von Videodaten (Jacobs et al. 1999)**

Dieser *zyklische Kategorienentwicklungsprozess* startet mit der Sichtung und Diskussion von einzelnen Videosequenzen bzw. videografierten Lektionen. Erste Vermutungen, Fragen und Forschungshypothesen können formuliert werden. Auf der Basis dieser Hypothesen werden die Entwicklung eines Kategoriensystems und die Definition einzelner Kategorien in Angriff genommen, wobei gefundene Beispielsequenzen zur Erläuterung der einzelnen Kategorien herangezogen werden können. Die Kategoriendefinitionen werden im nächsten Schritt auf ein

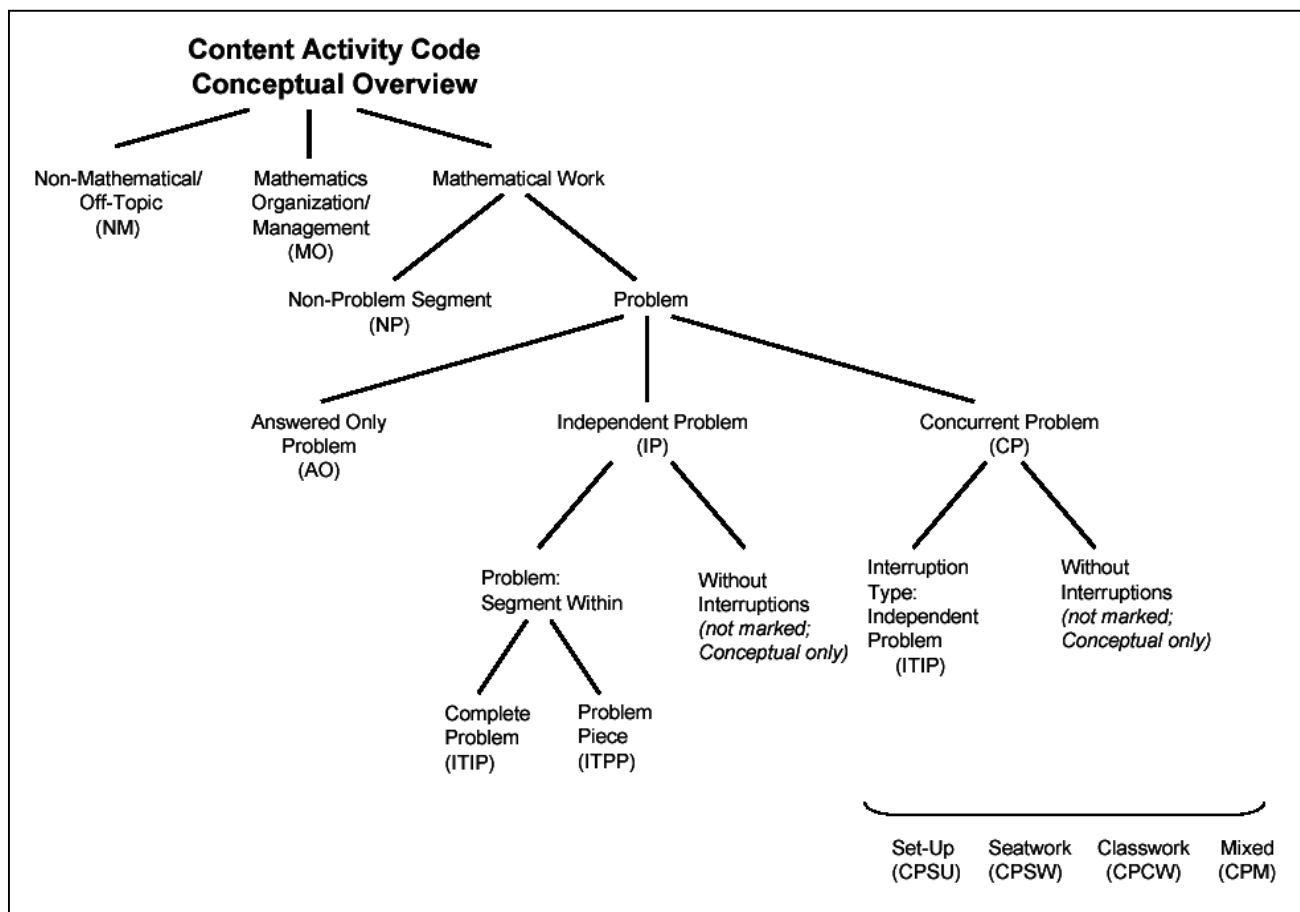
größeres Teilsample bzw. auf das ganze Sample angewandt. In dieser Phase wird getestet, ob und inwiefern eine Verallgemeinerung der eingangs formulierten Hypothesen zulässig ist. Statistische Analysen können zur Feststellung der Validität einer Kategoriendefinition bzw. eines Kategoriensystems herangezogen werden. Dabei geht es um die Prüfung der Frage, ob tatsächlich die zentralen Phänomene erfasst werden konnten oder ob mit den gefundenen Definitionen eher vereinzelt vorkommende Randphänomene erfasst wurden. Schließlich können die quantifizierten Ergebnisse wieder auf die Videodaten zurückgeführt werden. Die gefilmten Unterrichtssequenzen können von den Forschern wiederum als Hilfe bei der Interpretation der quantitativen Ergebnisse herangezogen werden. Dieser analytische Zirkel kann mehrere Male wiederholt werden, bis eine Kategoriendefinition im Detail passt oder bis eine zu Beginn eher weit gefasste Forschungshypothese eine praxisrelevante Fokussierung erfahren hat.

Dieser zyklische Prozess kann allerdings nur mit einer genügend großen Videostichprobe angewendet werden. Bei kleineren Studien können die Videodaten zwar zur qualitativen Beschreibung einzelner Aktivitäten und Prozesse herangezogen werden, es muss aber auf die quantitative Validierung anhand einer großen, repräsentativen Stichprobe verzichtet werden (ebd., 718).

Die wichtigsten Merkmale des Kategoriensystems der TIMSS 1999 Videostudie werden nachfolgend kurz dargestellt. Eine ausführliche Übersicht findet sich bei Jacobs et al. (2003, Appendix I). Das Kategoriensystem unterschied zunächst coverage-codes von occurrence-codes. Coverage-codes bilden ein Set von sich auf einer

Dimension gegenseitig ausschließenden Kategorien, von denen auf jeden Moment der Lektion nur *eine* zutrifft. Ein Beispiel ist die Unterteilung der Lektion in die disjunkten Sozialformen „öffentliche Klassenarbeit“, „Stillarbeit mit optionalem öffentlichem Lehrerbeitrag“, „Stillarbeit mit optionalem öffentlichem Schülerbeitrag“, und „Stillarbeit“. Der Beginn, das Ende und damit auch der Wechsel jeder Sozialform kann verhältnismäßig genau und reliabel bestimmt werden. Die Kodierung mit diesen Kategorien ist also „flächendeckend“. Ein weiteres Beispiel für eine flächendeckende Kategorisierung innerhalb der TIMSS 1999 Videostudie ist die Unterteilung der Lektion in Abschnitte mit verschiedenen Weisen der Aufgabenbearbeitung (vgl. Abb. 3 nach Jacobs et al. 2003, 407). Die Lektion wird zunächst vollständig in Abschnitte variabler Länge unterteilt, in denen entweder Nicht-Mathematik-bezogene, Mathematik-organisatorische oder Mathematik-bezogene Aktivitäten stattfinden. In einem nächsten Schritt werden die Mathematik-bezogenen Sequenzen dahingehend unterteilt, ob es sich um theoretische Ausführungen (Non-Problem Segment) oder um Arbeit an mathematischen Aufgaben handelt. Diese Sequenzen werden weiter gegliedert. Im Falle einer Aufgabenbearbeitung wird zwischen Aufgabenkontrolle (Answered Only Problems), im Unterricht bearbeiteten Einzelaufgaben (Independent Problems) und im Unterricht bearbeiteten Sets von Aufgaben (Concurrent Problems) unterschieden. Weitere Unterscheidungen betreffen unterbrechende Aufgaben und Einschübe.

Abbildung 3: Entscheidungsbaum eines Kategoriensystems der TIMSS 1999 Videostudie (Jacobs et al. 2003, 407)



Mit der gewählten Unterteilung kann bestimmt werden, wie viel Unterrichtszeit durchschnittlich mit Aufgabenlösen verbracht wird. Außerdem gibt sie Hinweise darauf, wie viele Aufgaben durchschnittlich im Unterricht gelöst werden und in welcher Taktung sie auftreten. Die Unterteilung bildet zudem eine Basis für weitergehende Kodierungen, etwa bezüglich des Anspruchsniveaus der Aufgaben, ihrer Verknüpftheit oder ihrem didaktischen Zweck. Für einige dieser Kodierungen war eine erhöhte fachspezifische Kompetenz auf Seiten der Kodierpersonen vonnöten. Die betreffenden Kodierungen wurden deshalb von Expertengruppen durchgeführt (vgl. Jacobs et al. 2003, 75; Hiebert et al. 2003, Appendix D).

Während zur Vergabe einer flächendeckenden Kategorie immer ein Anfangs- und Endpunkt in Abgrenzung zu anders kodierten Tätigkeitsphasen bestimmt werden muss, ist eine solche Abgrenzung für occurrence-codes nicht nötig. *Occurrence-codes* bezeichnen das bloße Auftreten bestimmter Elemente, beispielsweise das Aufgeben von Hausaufgaben oder der Verweis auf einen lebensweltlichen Bezug. Diese Kategorien schließen sich gegenseitig *nicht* aus. Die Bestimmung des Anfangszeitpunktes identifiziert den Moment des Auftretens. Die Bestimmung des Endpunktes ist auch hier zumindest theoretisch notwendig, damit eine wiederholte Kodierung möglich wird und eine Aussage über die Häufigkeit des Auftretens gemacht werden kann.

#### *Reliabilität der Kodierungen*

Mit der Reliabilitätsprüfung des Kategoriensystems wird versucht, Auswertungsfehler, die durch eine unsystematische Anwendung der Auswertungskategorien entstehen, zu vermeiden (vgl. Bakeman & Gottman 1994). Dieser Anspruch war im internationalen Kontext der Studie in besonderem Masse vonnöten. Auswertungen durch Kodierpersonal an verschiedenen Standorten (Zürich und Los Angeles) mussten nachweislich vergleichbar sein. Die Reliabilität diente auch als wichtiges Indiz für den erreichten Grad der Intersubjektivität über die entwickelten Kategorien.

Im Anschluss an das Kodiertraining sowie nach der Kodierung der Hälfte der Unterrichtsvideos wurden Reliabilitätsprüfungen zur Ermittlung der Beobachterübereinstimmung durchgeführt.

Vor dem Beginn der Kodierarbeiten wurde die Eingangsreliabilität als Übereinstimmung zwischen Entscheidungen von Kodierpersonen und der Kodiervorlage, welche von der Kodeentwicklungsgruppe erstellt wurde<sup>9</sup>, ermittelt. Nach der Bearbeitung der Hälfte der Lektionen wurden die Zwischenreliabilitäten mittels Feststellung der Übereinstimmung der Kodierentscheidungen je eines Kodiererpaares erhoben. Das Minimum der akzeptierbaren Reliabilität (gemessen durch Cohen's Kappa) wurde im Durchschnitt für alle Kodierpersonen auf .85 festgelegt. Kodierpaare mussten mindestens eine Übereinstimmung von .80 für jede einzelne Kategorie erreichen. Die Reliabilität wurde hierbei immer in einem doppelten

Sinne gemessen, einerseits in Bezug auf die übereinstimmende Bestimmung des Anfangs- und Endpunktes der Analyseeinheit (+/- 10 Sekunden) und andererseits in Bezug auf die Zuordnung der jeweiligen Kategorien.

Wegen der komplexen Kategorienstruktur wurde die Analyse in mehrere, aufeinander aufbauende *Kodierdurchgänge* aufgeteilt. Damit blieb die Aufmerksamkeit des Kodierpersonals in jedem Durchgang relativ fokussiert und eine reliable Kodierung wurde erleichtert. Dennoch führte der strenge Reliabilitätsanspruch, insbesondere im internationalen Kontext, zum Verlust bedeutender Kategorien. Während die bisher beschriebenen Kodierungen reliabel zu bestimmen waren, scheiterte beispielsweise die Differenzierung der Bearbeitungsweise von Einzelaufgaben in die vordergründig relativ eindeutigen Kategorien „Demonstrieren“, „Entwickeln“, „Antworten hervorrufen“, „Erklärungen hervorrufen“ und „Exploration erleichtern“ an diesem Kriterium. Zur Erfassung von derartigen Unterrichtsmerkmalen müsste ein anderes, hoch-inferentes Verfahren der Datenbeschreibung gewählt werden.

#### *Hoch-inferente Ratings oder niedrig-inferente Kodierungen*

Das vorhergehend beschriebene Kodiersystem der internationalen Studie kann als Beispiel für eine niedrig-inferente Kodierung gelten. Darüber hinaus wurde im schweizerisch-deutschen Videoprojekt ein hoch-inferenter Beurteilungsansatz an das gesamte schweizerische Videosample und die deutschen Videolektionen, welche aus der TIMSS 1995 Studie vorlagen, angelegt (Clausen, Reusser & Klieme 2003). In den folgenden Ausführungen sollen die beiden Verfahren kurz beschrieben werden.

*Niedrig-inferente Kodierungen* beziehen sich auf die Erfassung von *direkt beobachtbaren* Aspekten der Sicht- oder Oberflächenstruktur des Unterrichts, wie z.B. die Unterrichtsorganisation hinsichtlich der Interaktionsstruktur und der Sozialformen, die Formulierung des Lektionsziels durch die Lehrperson beim Stundeneinstieg oder die Verwendung von technischen Hilfsmitteln, wie Taschenrechner und Computer im Verlauf der Lektion, usw. Die Bestimmung von Anfangs- und Endpunkt eines bestimmten Vorkommnisses erlaubt später bei der Datenanalyse die Auszählung von Häufigkeiten wie auch die Feststellung der zeitlichen Dauer von bestimmten Unterrichtsaktivitäten. Bei der Kategorienentwicklung wird demzufolge das Ziel fokussiert, exakt formulierte Definitionen der Analysekatoren zu formulieren, welche dem Beobachtenden einen möglichst kleinen Interpretationsspielraum überlassen. Die Kodierung soll „objektiv“ geschehen und erfordert möglichst wenig schlussfolgernde Kognitionen beim Beobachter. Bevor die eigentliche Kodierarbeit beginnt, wird die Übereinstimmung der Beobachter festgestellt. Sie sollten in möglichst allen Fällen die selben Kodierentscheidungen treffen (vgl. Kapitel Reliabilität). Ist dies nicht der Fall, muss ein erneutes Training der Kodiererinnen und Kodierer durchgeführt oder eine Überarbeitung der Kategoriendefinition vorgenommen werden. Hinsichtlich der Qualifikation der Kodierpersonen ist eine profunde Kenntnis der Kodierregeln an die erste Stelle zu setzen. So wurden auch im

<sup>9</sup> Diese Vorlage wurde erstellt, indem die einzelnen Mitglieder der Kategorienentwicklungsgruppe dieselbe Lektion kodierten, dann im Team die individuellen Kodierungen besprachen und die Nichtübereinstimmungen so lange diskutierten, bis Konsens hergestellt werden konnte.

TIMSS 1999 Video Kodiermanual präzise Kategorienbeschreibungen und Ankerbeispiele schriftlich festgehalten (Jacobs et al. 2003, Appendix I). Zur Ausbildung der Kodierpersonen wurde für jeden Kodierdurchgang eigens ein Trainingscurriculum entwickelt. Mit dieser Vorgehensweise waren beträchtliche zeitliche und finanzielle Aufwendungen verbunden.

Informationen – oder besser: *Urteile* – über Instruktionseffizienz der Lektion, zur Klarheit und Strukturiertheit der Lehrerklärungen oder zum Grad der kognitiven Aktivierung der Schülerinnen und Schüler können jedoch kaum durch niedrig-inferente Kodierungen gewonnen werden. Zwar finden sich in den Videos hinsichtlich solcher abstrakteren Sachverhalte bzw. Qualitätsmerkmale durchaus Anhaltspunkte. Diese sind jedoch häufig über die gesamte Lektion verstreut und lassen sich nicht an einem spezifischen Ereignis festmachen. Zur Erfassung von globaleren Merkmalen von Unterrichtsqualität sind Beurteilungsverfahren erforderlich, welche *interpretative Prozesse* auf Seiten der Beobachter explizit einschließen. In diesem Beobachtungs- und Beurteilungsprozess spielen die *Expertise des Beobachters*, dessen Kenntnis eines breiten Spektrums von theoretischen und praktischen Grundlagen zur Gestaltung von Lehr-Lernprozessen und dessen profunde Kenntnis der fachlichen Inhalte eine wichtige Rolle. Die Beurteilung erfolgt nach der Visionierung der gesamten Lektion.

Im schweizerischen Projekt wurden in einer Kooperation mit Deutschland *hoch-inferente Ratings* entwickelt und unter Beizug von vier trainierten Beobachtern pro Lektion mit Hilfe eines Rating-Fragebogens durchgeführt. Das eingesetzte Instrument umfasste insgesamt 94 Items, wobei jedes interessierende Unterrichtsmerkmal mit jeweils drei bis vier (eine Skala bildenden) Items erfasst wurde, beispielsweise „Interaktionstempo“ mit Items wie „Der Lehrer lässt bei Fragen kaum Zeit zum Nachdenken“, oder „Der Lehrer geht gleich zum Nächsten, wenn ein Schüler nicht sofort antwortet“. Die Auswahl der erfassten Merkmale orientierte sich an von der empirischen Unterrichtsforschung als relevant identifizierten Unterrichtsqualitätsmerkmalen oder an theoretischen Postulaten der aktuellen Lehr-Lernforschung. Aufgrund einer empirischen (faktorenanalytischen) Überprüfung konnten jeweils drei bis vier Subskalen den vier übergreifenden Merkmalsbereichen *Instruktionseffizienz, Schülerorientierung, kognitive Aktivierung und Klarheit und Strukturiertheit* zugeordnet werden (Clausen et al. 2003).

Im Gegensatz zu einer niedrig-inferenten Kodierung ist die Schulung von Expertinnen und Experten für hoch-

inferente Unterrichtsbeurteilungen weniger aufwändig. Im Rahmen eines einwöchigen Trainingsworkshops wurden die Beurteilungsstandards erarbeitet und innerhalb der Beobachtergruppe in Übereinstimmung gebracht. Die fachliche Qualifikation der Beobachterinnen und Beobachter spielte bei deren Rekrutierung eine wichtige Rolle.

Hoch-inferente Beurteilungen erlauben eine Zusammenfassung von Einzelratings zu übergeordneten Qualitätsmerkmalen. Konstrukte wie „Klarheit und Strukturiertheit“ oder „Kognitive Aktivierung“ sind intuitiv zugänglich und können mit konkreten Vorstellungen unterrichtlichen Handelns in Verbindung gebracht werden. Im Gegensatz zu hoch-inferenten Ratings sind niedrig-inferente Kodierungen für sich genommen deutlich weniger aussagestark. Erst in ihrer Zusammenschau entfalten sie ein umfassenderes Bild von Unterricht.

#### *Integration von Kodierungen zu einer Gesamtsicht*

Niedrig-inferente Kodierungen generieren Aussagen mit einfachem Informationswert. Gemittelte Werte über Länge und Menge bearbeiteter Aufgaben beispielsweise sind relativ abstrakt und sagen noch nichts aus über das Anspruchsniveau dieser Aufgaben und die Art und Weise, wie diese Aufgaben im Lektionsverlauf präsentiert und bearbeitet werden. Interessant ist jedoch gerade die Frage nach der Choreographie des Unterrichts, z.B. nach der Verknüpfung einzelner Unterrichtsaktivitäten oder nach charakteristischen Verlaufsmustern des Unterrichts. Es geht dabei um die typischen Konstellationen und Anordnungen von Unterrichtsaktivitäten, Lehr- und Sozialformen im zeitlichen Verlauf von Schulstunden, welche sich eventuell auf ein gemeinsam geteiltes Unterrichtsskript zurückführen lassen (Stigler & Hiebert 1999, Pauli & Reusser 2003).

Zur Beantwortung solcher Fragen wurden in der internationalen TIMSS 1999 Video Studie zwei verschiedene Strategien zur Verknüpfung der Einzelbefunde verfolgt. Einerseits wurden Variablenkombinationen miteinander verrechnet und dargestellt, andererseits wurde zur Aufdeckung charakteristischer Verlaufsmuster des Unterrichts für jedes Teilnehmerland eine sogenannte „*lesson signature*“ erstellt. Dabei wurden die *über alle Lektionen summierten Vorkommnisse der Einzelkategorien im Lektionsverlauf prozentual abgebildet, mit dem Ziel, ländertypische Verlaufprofile* herauszuarbeiten (Hiebert et al. 2003). Zur Illustration ist an dieser Stelle das Lektionsprofil der Schweiz abgebildet (Abb. 4).

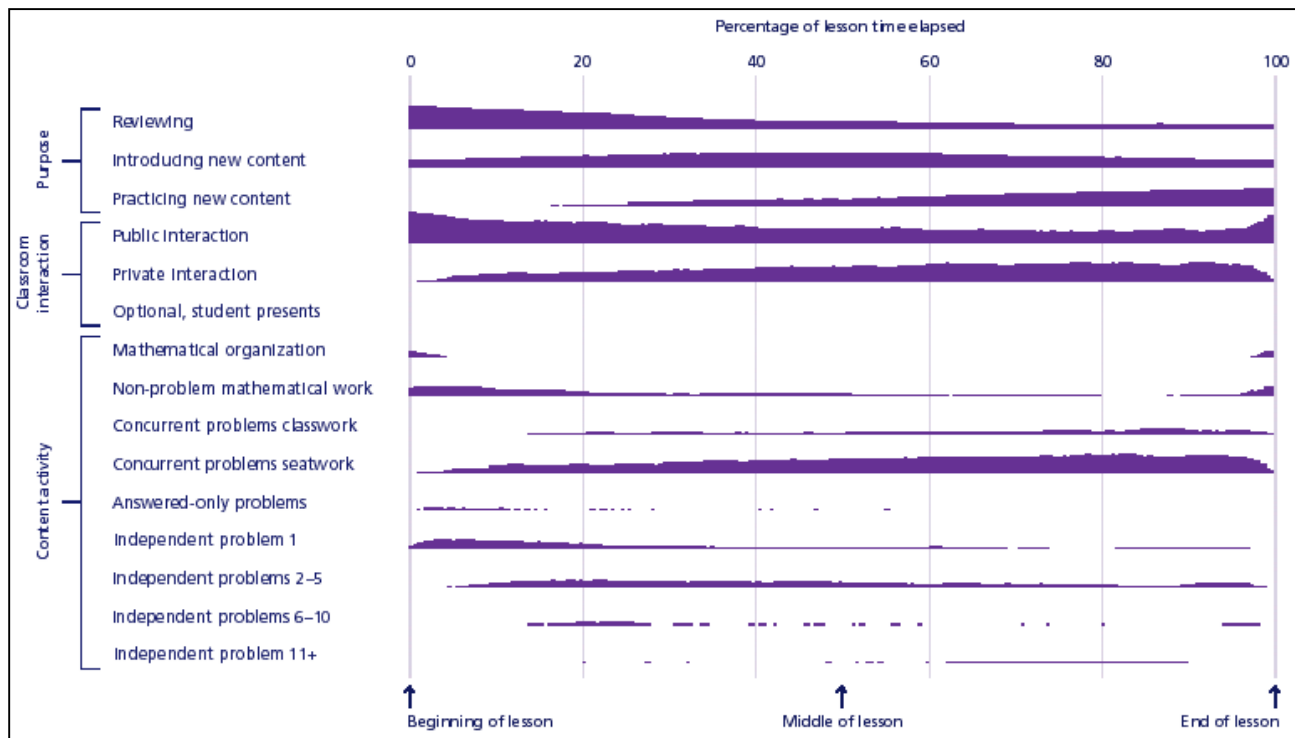


Abbildung 4: Lektionsprofil der schweizerischen Stichprobe (Hiebert et al. 2003, 143)

Tatsächlich lässt sich mit Hilfe dieser Profile die Positionierung der Einzelaktivitäten im Lektionsverlauf darstellen und es wird klar, welche Aktivitäten in der Schweiz mehrheitlich am Lektionsanfang stattfinden und welche Aktivitäten eher in der Mitte oder am Ende der schweizerischen Lektionen angesiedelt sind. Darüber hinaus kann über die Feststellung des gemeinsamen Auftretens von verschiedenen Aktivitäten auch die Funktion von gewissen Einzelercheinungen klarer interpretiert werden; z.B. deutet das gemeinsame Auftreten einer Wiederholungsphase (reviewing), öffentlicher Interaktion (public interaction) und Nicht-Aufgabenbearbeitungs-Segmenten (non problem segments) auf eine vom Lehrer geführte theoretische Repetitionsphase im Klassenverband hin, bei der mathematische Regeln etc. noch einmal öffentlich wiederholt werden.

Weiter enthalten die „lesson signatures“ auch Informationen zur *Variabilität* des Auftretens einer Kategorie in den Lektionsverläufen eines einzelnen Landes. Gleichmäßige Balken mittleren Ausmaßes ohne eigentliche Spitze deuten (über alle Lektionen gesehen) auf eine Variabilität im zeitlichen Auftreten einer Kategorie hin; Spitzenwerte mit anschließend deutlicher Verflachung hingegen auf das eher zeitgleiche Auftreten einer Einzelkategorie; d.h. dass diese Unterrichtsaktivität über die verschiedenen Lektionen betrachtet immer ungefähr an einer ähnlichen Stelle im Unterrichtsverlauf auftritt.

Im Vergleich zu den Profilen anderer Teilnehmerländer weist das Lektionsprofil der Schweiz auf eine stärkere Variationsbreite des Einsatzes bestimmter Unterrichtsaktivitäten im Lektionsverlauf hin, was aus den eher langgezogenen mitteldicken Balken ohne eigentliche Spitzen in der Grafik deutlich wird (vgl. Abb. 4). Als mögliche Ursache können hier *unterschiedliche Unterrichtskulturen* innerhalb der Schweiz in Erwägung gezogen werden.

Bereits vorgenommene Analysen deuten denn auch darauf hin, dass zumindest in der deutschsprachigen Schweiz zwischen einer eher traditionellen Unterrichtskultur und einer an neuen Lehr- und Lernformen orientierten Kultur unterschieden werden muss (Pauli, Reusser, Waldis & Grob 2003, Stebler & Reusser 2000). Für eine genauere Analyse dieser Hypothesen ist allerdings die Generierung von weiteren, der jeweiligen Fragestellung angepassten Unterrichtsprofilen notwendig. Der bei Voranalysen vorgenommene Split des gesamtschweizerischen Lektionsprofils in Subprofile der drei Sprachregionen der Schweiz weist beispielsweise auf – allerdings nicht auf einfache Weise fassbare – sprachregionale Differenzen in der Unterrichtsgestaltung hin.

Insgesamt stellen die Lektionsprofile eine wertvolle Ergänzung zu den im internationalen Bericht (Hiebert et al. 2003) präsentierten tabellarischen Auswertungen dar. Insbesondere die Positionierung der Kategorien auf der Zeitachse kann als wertvolle Zusatzinformation gewertet werden. Die vorsichtig formulierten Interpretationen der grafischen Abbildungen gehen über die statistisch quantifizierbaren Einzelauswertungen hinaus und eröffnen einen Einblick in übergreifende Zusammenhänge, die rein rechnerisch nur schwer zu erfassen wären.

#### *Kausalität von Variablenbeziehungen*

Ein querschnittliches Studiendesign, wie es z.B. in den vorliegenden Studien zur Anwendung kam, erlaubt streng genommen keine Aussagen zur Richtung der Variablenbeziehungen. Unterricht ist zugleich abhängige und unabhängige Variable des Bildungssystems. In Bezug auf Bildungswirkungen wie Fachleistungen oder motivationale Dispositionen der Schülerinnen und Schüler ist deshalb zunächst nicht zu entscheiden, in welchem Masse der Unterricht diese bestimmt bzw. in welchem Masse die

Unterrichtsgestaltung vom Leistungsstand und den dispositionalen Merkmalen der Schülerinnen und Schüler beeinflusst wird.

Wird von einem Land ein genügend großes und repräsentatives Zufallsample gezogen, wie im Fall der TIMSS 1999 Videostudie, so kann allerdings davon ausgegangen werden, dass damit nicht nur ein repräsentatives Bild des alltäglichen Unterrichts rekonstruiert werden kann, sondern sich auf dieser Grundlage auch Vermutungen über den Zusammenhang von Unterrichtsmerkmalen und Wirkungskriterien, wie sie in TIMSS oder PISA erhoben wurden, anstellen lassen.

Um allerdings gesicherte Kausalaussagen machen zu können, sind Forschungsprojekte mit längsschnittlichem Charakter erforderlich, in denen vorzugsweise lektionsübergreifende Unterrichtseinheiten gefilmt werden, und diese Dokumentation der Unterrichtsprozesse mit der inhaltspezifischen Erfassung von Veränderungsmaßen im Bereich der Wirkungskriterien (z.B. Leistungs- und Interesseentwicklung im spezifisch behandelten Teilgebiet) ergänzt wird (z.B. Reusser & Pauli 2000, Klieme & Reusser 2003). In diesem Falle können die erfassten Unterrichtsprozesse in einen kausalem Zusammenhang mit den beobachteten Entwicklungsprozessen gestellt werden. Als große Herausforderung erweist sich allerdings bei solchen Studien die Rekrutierung und Betreuung von Schulklassen und Lehrpersonen. Von der Bildung eines repräsentativen Samples muss hier abgesehen werden, da die Teilnahme an einer solchen Studie ein großes Engagement der Lehrpersonen voraussetzt, welches in der Regel nur auf freiwilliger Basis eingefordert werden kann.

#### *Methoden- und perspektivenvergleichende Analysen*

In der schweizerischen Teilstudie – wie auch in weiteren nationalen Ergänzungsprojekten zu den beiden TIMSS Videostudien, z.B. in Deutschland oder Australien – wurden neben Videodaten umfangreiche Zusatzerhebungen durchgeführt. Damit wird die Verknüpfung unterschiedlicher Datentypen möglich, z.B. qualitativer und quantitativer Art aus unterschiedlichen Datenquellen (Videoaufzeichnungen, Fragebögen, Interviews und Tests). Auch verschiedene Perspektiven auf Unterricht können miteinander verglichen werden. Solche Verfahren werden üblicherweise unter dem Begriff der Triangulation gefasst (Kelle & Erzberger 1999, Flick 2002). Triangulation wird mit zwei Zielsetzungen angewendet. Einerseits dient Triangulation der kumulativen Validierung von Forschungsergebnissen, andererseits kann Triangulation zur Erfassung von komplementären, d.h. sich ergänzenden Perspektiven auf einen Erkenntnisgegenstand eingesetzt werden. Während kumulative Triangulation auf die Feststellung von Divergenz und Konvergenz abzielt, geht es bei komplementärer Triangulation um die Generierung einer möglichst umfassenden Sichtweise, indem verschiedene Facetten eines Gegenstandsbereichs beschrieben werden. Die Abgrenzungen zwischen diesen Ansätzen sind bisher nicht vollständig geklärt. Ob es sich bei einer Triangulation um den kumulativen oder komplementären Ansatz handelt, hängt letztlich von der zugrunde liegenden Fragestellung ab (vgl. Klieme & Bos 2000).

Ein erweitertes Forschungsdesign, wie es in der schweizerischen Videostudie angelegt ist, erlaubt den Vergleich verschiedener Perspektiven auf Unterricht. Sie kann einen Vergleich von Innen- und Außensicht, von Lehrer- und Schülerperspektive, von Experten- und Laienwahrnehmung umfassen. Mit multiperspektivischen Daten ist es möglich, die jeweils besonderen Sichtweisen von Lehrpersonen, Lernenden und externen Beobachtern auf den gefilmten Unterricht herauszuarbeiten. Diese Fragestellung zielt auf eine komplementäre Beschreibung von Unterrichtsprozessen. Es kann aber auch der Frage nachgegangen werden, ob die Wahrnehmungen der Unterrichtsprozesse durch die unterschiedlichen Personengruppen eher übereinstimmen (kumulative Validierung) oder ob perspektivenspezifische Wahrnehmungen vorherrschen. Bereits vorliegende Forschungsergebnisse deuten darauf hin, dass eine konvergente Unterrichtswahrnehmung bei unterschiedlichen Personengruppen eher die Ausnahme als die Regel darstellt (vgl. Clausen 2002).

#### **Perspektiven videobasierter Unterrichtsforschung**

Auf Video aufgezeichneter Unterricht macht die Ebene konkreten Lehrens und Lernens sichtbar, ein Aspekt, der in internationalen Schulleistungsvergleichen bisher vorwiegend ausgeblendet wurde. Die Erfahrungen aus der TIMSS 1999 Videostudie ermutigen zu einer umfassenden und weiter gehenden Nutzung von Videodaten im Rahmen der Unterrichts- und Schulforschung. Technische Entwicklungen und methodologische Klärungen werden den Nutzen dieses Forschungsansatzes noch erhöhen.

Besondere Potenziale liegen in der Nutzbarmachung von Videodaten in Verbindung mit anderen Datenquellen. Videodaten sollten im Blick auf künftige Studien denn auch nicht als alleinige Datenquelle, sondern als wichtiges Element in Ergänzung zu anderen Datenebenen, z.B. Befragungen und Tests im Sinne eines systemischen Modells von Unterrichtsqualität, betrachtet werden. Große Potenziale liegen insbesondere in videounterstützten Längsschnittstudien, die eine Identifizierung von Unterrichtsmustern über längere Einheiten erlauben und diese in Beziehung zu z.B. direkt überprüfbar Leistungszuwächsen im spezifischen Teilgebiet setzen. Diese Strategie wird im schweizerisch-deutschen Videoprojekt verfolgt (Klieme & Reusser 2003).

Die Implementation videogestützter Interventionsstudien mit dem Ziel der Unterrichtsentwicklung, Lehrerbildung und Lehrerweiterbildung ist ebenfalls naheliegend. In solchen Projekten werden Videodaten nicht nur von einem engen Kreis von interessierten Forscherinnen und Forschern analysiert, sondern es werden zusätzlich die beteiligten Lehrpersonen in die Analyse eigener sowie auch fremder Videos einbezogen. Die gemeinsame Betrachtung der videografierten Lektionen hat zum Ziel, den eigenen Unterricht zu reflektieren und das videografierte Material zur Generierung neuer Ideen für die persönliche Unterrichtsgestaltung zu nutzen. Videodaten bilden einen Kristallisationspunkt der fachdidaktischen Verständigung über Unterrichtsprozesse und bauen eine Brücke über die Kluft zwischen didaktischer Theorie und Praxis. Aus



wissenschaftlicher Sicht bietet sich in derartigen Projekten die Gelegenheit zur empirischen Untersuchung der Entwicklung professionellen Wissens und professioneller Handlungsstrategien. Ein diesbezügliches Projekt zur Weiterbildung von Lehrpersonen ist zur Zeit unter Leitung von K. Reusser, C. Pauli und E. Klieme in Vorbereitung.

Ein anderer möglicher Weg, der in Los Angeles und Zürich in dieser Richtung eingeschlagen wurde, ist der Einsatz eines internetbasierten Tools, LessonLab Viewer™, welches die Diskussion von Videodaten über das Internet erlaubt (Petko, Reusser, Noetzi, Krammer & Hugener 2003). Dieser Ansatz verbindet die komfortable Zugänglichkeit der Videodaten mit den erweiterten Möglichkeiten hypermedialer Lernarrangements und Kommunikationsformen. Mit Videos können unterschiedliche Lernaktivitäten angeregt werden. Sie illustrieren die Variationsbreite von Unterrichtsprozessen und regen, gerade im internationalen Kontext, zum Nachdenken über die eigene Unterrichtskultur an. Videos können als Feedbackinstrument dienen, unter Umständen ein „best-practice“-Modell für modernes Unterrichten bieten oder als Ausgangspunkt für fallbasiertes Lernen genutzt werden.

Videodaten können immer wieder neu aus unterschiedlichen Perspektiven betrachtet und analysiert werden. So bleibt zu hoffen, dass – vor allem auch wegen der aufwändigen Datenaufbereitung und Datenanalyse – Videodaten nicht allzu schnell im Archiv verschwinden. Die Zugänglichkeit der Daten für weitere Forscherteams, ihre Kombination mit anderen Datenebenen und die Nutzbarmachung für die Bedürfnisse der Lehrerbildung müssen deshalb als zentrale Anliegen gelten.

## Literatur

- Aufschnaiter, S. v. & Welzel, M. (2001). Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen: Eine Einführung. In S. v. Aufschnaiter & M. Welzel (Hrsg.), *Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen. Aktuelle Methoden empirischer pädagogischer Forschung* (S. 7-15). Münster: Waxmann.
- Bakeman, R. & Gottman, J. M. (1994). *Observing interaction: An introduction to sequential analysis* (2. Ed.). Cambridge <etc.>: Cambridge University Press.
- Brinker, K. & Sager, S. F. (2001). *Linguistische Gesprächsanalyse. Eine Einführung* (3., durchges. und erg. Aufl.). Berlin: Erich Schmidt.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 177-212). Göttingen: Hogrefe.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität*. Münster: Waxmann.
- Clausen, M., Reusser, K. & Klieme, E. (2003). Unterrichtsqualität auf der Basis hoch-inferenter Unterrichtsbeurteilungen: Ein Vergleich zwischen Deutschland und der deutschsprachigen Schweiz. *Unterrichtswissenschaft*, 31(2), 122-141.
- Dann, H.-D. (1997). Pädagogisches Verstehen: Subjektive Theorien und erfolgreiches Handeln von Lehrkräften. In K. Reusser & M. Reusser-Weyeneth (Hrsg.), *Verstehen. Psychologischer Prozess und didaktische Aufgabe* (S. 163-182). Bern: Huber.
- Dittmar, N. (2002). *Transkription. Ein Leitfaden mit Aufgaben für Studenten, Forscher und Laien*. Opladen: Leske und Budrich.
- Ellgring, H. (1995). Audiovisuell unterstützte Beobachtung. In U. Flick, E. Von Kardorff, H. Keupp, L. Von Rosenstiel & S. Wolff (Hrsg.), *Handbuch Qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen* (2. ed., S. 203-208). Weinheim: Beltz. Psychologie Verlags Union.
- Evertson, C. M. & Green, J. L. (1986). Observation as Inquiry and Method. In M. C. Wittrock (Ed.), *Handbook of research on teaching: a project of the American Educational Research Association* (3rd, [4th printing] ed., pp. 162-213). New York: MacMillan.
- Fend, H. (1998). *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung*. Weinheim: Juventa.
- Flick, U. (1995). Der qualitative Forschungsprozess als Abfolge von Entscheidungen. In U. Flick, E. v. Kardorff, H. Keupp, L. v. Rosenstiel & S. Wolff (Hrsg.), *Handbuch qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen* (S. 148-173). Weinheim: Psychologie-Verlags-Union.
- Flick, U. (2002). *Qualitative Sozialforschung. Eine Einführung*. Reinbek: Rowohlt Taschenbuchverlag.
- Hall, R. (2000). Videorecording as Theory. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 647-664). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hanke, B., Mandl, H. & Prell, S. (1973). *Soziale Interaktion im Unterricht. Darstellung und Anwendung des Interaktions-Systems von N. A. Flanders*. München: R. Oldenbourg Verlag.
- Helmke, A. (2003). *Unterrichtsqualität: Erfassen - bewerten - verbessern*. Seelze: Kallmeyer.
- Hiebert, J., Gallimore, R., Garnier, H., Bogard Givvin, K., Hollingsworth, H., Jacobs, J., Chui, A. M. Y., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P. & Stigler, J. (2003). *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study* (No. NCES 2003-013). Washington DC: U.S. Department of Education, National Center for Education Statistics.
- Jacobs, J., Garnier, H., Gallimore, R., Hollingsworth, H., Bogard Givvin, K., Rust, K., Kawanaka, T., Smith, M., Wearne, D., Manaster, A., Etterbeek, W., Hiebert, J., Stigler, J. & Gonzales, P. (2003). *Third International Mathematics and Science Study 1999 Video Study Technical Report. Volume 1: Mathematics* (No. NCES (2003-012)). Washington: National Center for Education Statistics, Institute of Education Statistics, U.S. Department of Education.
- Jacobs, J., Kawanaka, T. & Stigler, J. (1999). Integrating qualitative and quantitative approaches to the analysis of video data on classroom teaching. *International Journal of Educational Research*, 31, 717-724.
- Kelle, U. & Erzberger, C. (1999). Integration qualitativer und quantitativer Methoden. Methodologische Modelle und ihre Bedeutung für die Forschungspraxis. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 51(3), 509-531.
- Kelle, U. & Kluge, S. (1999). *Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung*. Opladen: Leske und Budrich.
- Klieme, E. (1999). *Unterrichtsqualität und mathematisches Verständnis in verschiedenen Unterrichtskulturen. Antrag auf Gewährung einer Sachbeihilfe im Schwerpunktprogramm "Bildungsqualität von Schule"*. Berlin: Max Planck - Institut für Bildungsforschung.
- Klieme, E. & Bos, W. (2000). Mathematikleistung und mathematischer Unterricht in Deutschland und Japan. Triangulation qualitativer und quantitativer Analysen am Beispiel der TIMS-Studie. *Zeitschrift für Erziehungswissenschaft*, 3(3), 259-380.



- Klieme, E. & Reusser, K. (2003). Unterrichtsqualität und mathematisches Verständnis im internationalen Vergleich - Ein Forschungsprojekt und erste Schritte zur Realisierung. *Unterrichtswissenschaft*, 31(3), 194-205.
- Knoll, S. & Stigler, J. (1999). Management and analysis of large-scale video surveys using the software vPrism™. *International Journal of Educational Research*, 31, 725-734.
- Muhr, T. (1994). ATLAS/ti. Ein Werkzeug für die Textinterpretation. In A. Boehm, A. Mengel & T. Muhr (Hrsg.), *Texte verstehen. Konzepte, Methoden, Werkzeuge* (S. 317-324). Konstanz: Universitätsverlag Konstanz.
- Pauli, C. & Reusser, K. (2003). Unterrichtsskripts im schweizerischen und deutschen Mathematikunterricht. *Unterrichtswissenschaft*, 31(3), 238-272.
- Pauli, C., Reusser, K., Waldis, M. & Grob, U. (2003). "Erweiterte Lehr- und Lernformen" im Mathematikunterricht der Deutschschweiz. *Unterrichtswissenschaft*, 31(4), 291-320.
- Petko, D., Reusser, K., Noetzi, C., Krammer, K. & Hugener, I. (2003). *Collaborative video based teacher training in a virtual learning environment*. Paper presented at the 10th European Conference for Research on Learning and Instruction (EARLI), Padova / Italy, August 26-30 2003.
- Reusser, K. (2001). *Bridging Instruction to Learning - Where we come from and where we need to go. A research strategy and its implementation in a national and cross-cultural study in Switzerland. Keynote address an der 9th European Conference for Research on Learning and Instruction (EARLI)*. Unveröffentlichtes Manuskript, Fribourg.
- Reusser, K. & Pauli, C. (1999). *Unterrichtsqualität: Multiterminiert und multikriterial*. Unveröffentlichtes Manuskript, Zürich.
- Reusser, K. & Pauli, C. (2000). *Unterrichtsqualität. Lernverhalten und mathematisches Verständnis. Eine schweizerisch-deutsche Videostudie. Forschungsgesuch an den SNF*. Zürich: Universität Zürich, Pädagogisches Institut.
- Reusser, K. & Pauli, C. (2003). *Mathematikunterricht in der Schweiz und in weiteren sechs Ländern. Bericht über die Ergebnisse einer internationalen und schweizerischen Video-Unterrichtsstunde. Doppel-CD-Rom (unter Mitarbeit der Video-Projektgruppe des Pädagogischen Instituts der Universität Zürich)*. Zürich: Universität Zürich, Pädagogisches Institut.
- Reusser, K., Pauli, C. & Zollinger, A. (1998). Mathematiklernen in verschiedenen Unterrichtskulturen. *Beiträge zur Lehrerbildung*, 16(3), 427-438.
- Rimmele, R. (2003). *Das Programm Videograph*. Online verfügbar unter: <http://www.ipn.uni-kiel.de/aktuell/videograph/htmStart.htm> [Stand: 12. 11. 2003].
- Schümer, G. et al. (1997). *TIMSS: Videotape classroom study discourse coding. Manual and coding procedures*. Berlin: Max Planck Institute for Human Development.
- Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 726-764). New York: Macmillan.
- Stebler, R. & Reusser, K. (2000). Progressive, classical or balanced - a look at mathematical environments in Swiss-German lower-secondary schools. *Zentralblatt für die Didaktik der Mathematik*, 32(1), 1-10.
- Stigler, J. (1998). Video Surveys: New Data for the Improvement of Classroom Instruction. In S. G. Paris & H. M. Wellman (Eds.), *Global Prospects for Education. Development, Culture and Schooling* (pp. 129-168). Washington DC: American Psychological Association.
- Stigler, J., Gallimore, R. & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist*, 35(2), 87-100.
- Stigler, J., Gonzales, P., Kawanaka, T., Knoll, S. & Serrano, A. (1999). *The TIMSS Videotape Classroom Study: Methods and Findings From an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan and the United States* (No. NCES 1999-074). Washington DC: U.S. Department of Education, National Center for Education Statistics.
- Stigler, J. & Hiebert, J. (1999). *The teaching gap. Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.
- Terhart, E. (1997). *Lehr-Lernmethoden. Eine Einführung in Probleme der methodischen Organisation von Lehren und Lernen* (2. Aufl.). Weinheim: Juventa.
- Ulewicz, M. & Beatty, A. (2001). *The Power of Video Technology in International Comparative Research in Education*. Washington D.C.: National Academy Press.
- Wahl, D. (1991). *Handeln unter Druck. Der weite Weg vom Wissen zum Handeln bei Lehrern, Hochschullehrern und Erwachsenenbildnern*. Weinheim: Deutscher Studien Verlag.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249-294.
- Wild, K. P. (2003a). *Funktionen von CatMovie 4*. Online verfügbar unter: <http://www.catmovie.de> [Stand: 12.11. 2003].
- Wild, K. P. (2003b). Videoanalysen als neue Impulsgeber für eine praxisnahe, prozessorientierte empirische Unterrichtsforschung. *Unterrichtswissenschaft*, 31(2), 98-101.

---

**Autoren:**

Dominik Petko, M. A., Pädagogisches Institut der Universität Zürich, Gloriastr. 18a, CH-8006 Zürich  
Email: [petko@paed.unizh.ch](mailto:petko@paed.unizh.ch)

Monika Waldis, lic. phil., Pädagogisches Institut der Universität Zürich, Gloriastr. 18a, CH-8006 Zürich  
Email: [waldis@paed.unizh.ch](mailto:waldis@paed.unizh.ch)

Christine Pauli, Dr. phil., Pädagogisches Institut der Universität Zürich, Gloriastr. 18a, CH-8006 Zürich  
Email: [cpauli@paed.unizh.ch](mailto:cpauli@paed.unizh.ch)

Kurt Reusser, Prof. Dr., Pädagogisches Institut der Universität Zürich, Gloriastr. 18a, CH-8006 Zürich  
Email: [reusser@paed.unizh.ch](mailto:reusser@paed.unizh.ch)