

Prenzel, Manfred [Hrsg.]; Doll, Jörg [Hrsg.]

Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen

Weinheim u.a. : Beltz 2002, 389 S. - (Zeitschrift für Pädagogik, Beiheft; 45)



Quellenangabe/ Reference:

Prenzel, Manfred [Hrsg.]; Doll, Jörg [Hrsg.]: Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. Weinheim u.a. : Beltz 2002, 389 S. - (Zeitschrift für Pädagogik, Beiheft; 45) - URN: urn:nbn:de:0111-opus-78625 - DOI: 10.25656/01:7862

<https://nbn-resolving.org/urn:nbn:de:0111-opus-78625>

<https://doi.org/10.25656/01:7862>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der


Leibniz-Gemeinschaft

Zeitschrift für Pädagogik · 45. Beiheft

Zeitschrift für Pädagogik · 45. Beiheft

Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen

Herausgegeben von Manfred Prenzel und Jörg Doll

Beltz Verlag · Weinheim und Basel

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genützte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, von der die einzelnen Zahlungsmodalitäten zu erfragen sind.

© 2002 Beltz Verlag • Weinheim und Basel
Herstellung: Klaus Kaltenberg
Druck: Druckhaus »Thomas Müntzer«, Bad Langensalza
Printed in Germany
ISSN 0514-2717

Bestell-Nr. 41146

Inhaltsverzeichnis

<i>Jörg Doll/Manfred Prenzel</i> Einleitung in das Beiheft	9
Teil I:	
Unterrichtsforschung in Mathematik	
Förderung des mathematischen Verständnisses, Problemlösens und der Herausbildung zutreffender mathematischer Weltbilder von Schülerinnen und Schülern	31
<i>Kristina Reiss</i> Einleitung	32
<i>Christoph Wassner/Laura Martignon/Peter Sedlmeier</i> Die Bedeutung der Darbietungsform für das alltagsorientierte Lehren von Stochastik	35
<i>Kristina Reiss/Frank Hellmich/Joachim Thomas</i> Individuelle und schulische Bedingungsfaktoren für Argumentationen und Beweise im Mathematikunterricht	51
<i>Ingmar Hosenfeld/Andreas Helmke/Friedrich-Wilhelm Schrader</i> Diagnostische Kompetenz: Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Unterrichtsstudie SALVE	65
<i>Rudolf vom Hofe/Reinhard Pekrun/Michael Kleine/Thomas Götz</i> Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA). Konstruktion des Regensburger Mathematikleistungstests für 5.–10. Klassen	83

Teil II:

Lehrerexpertise und Unterrichtsmuster in Mathematik und Physik

Videografie von Unterrichtssequenzen in Mathematik und Physik: Diagnose, Analyse und Training erfolgreicher Unterrichtsskripts 101

Eckhard Klieme

Einleitung 102

Martina Diedrich/Claudia Thußbas/Eckhard Klieme

Professionelles Lehrerwissen und selbstberichtete Unterrichtspraxis im Fach Mathematik 107

Hans E. Fischer/Thomas Reyer/Tina Wirz/Wilfried Bos/Nicole Höllrich

Unterrichtsgestaltung und Lernerfolg im Physikunterricht 124

*Manfred Prenzel/Tina Seidel/Manfred Lehrke/Rolf Rimmele/Reinders Duit/
Manfred Euler/Helmut Geiser/Lore Hoffmann/Christoph Müller/Ari Widodo*

Lehr-Lernprozesse im Physikunterricht – eine Videostudie 139

Helmut Fischler/Hans-Joachim Schröder/Cornelia Tönhäuser/Peter Zedler

Unterrichtsskripts und Lehrerexpertise: Bedingungen ihrer Modifikation 157

Teil III:

Entwicklung und Evaluation von Unterrichtsmodulen und Trainingsprogrammen

Schulische Lehr-Lernumgebungen und außerschulische Trainings zur Förderung fächerübergreifender Kompetenzen von Schülerinnen und Schülern 173

Bernhard Schmitz

Einleitung 174

Kornelia Möller/Angela Jonen/Ilonca Hardy/Elsbeth Stern

Die Förderung von naturwissenschaftlichem Verständnis bei Grundschulkindern durch Strukturierung der Lernumgebung 176

Beate Sodian/Claudia Thoermer/Ernst Kircher/Patricia Grygier/Johannes Günther

Vermittlung von Wissenschaftsverständnis in der Grundschule 192

<i>Elke Sumfleth/Elke Wild/Stefan Rumann/Josef Exeler</i> Wege zur Förderung der naturwissenschaftlichen Grundbildung im Chemie- unterricht: kooperatives Problemlösen im schulischen und familialen Kontext zum Themenbereich Säure-Base	207
<i>Tina Gürtler/Franziska Perels/Bernhard Schmitz/Regina Bruder</i> Training zur Förderung selbstregulativer Fähigkeiten in Kombination mit Problemlösen in Mathematik	222
<i>Claudia Leopold/Detlev Leutner</i> Der Einsatz von Lernstrategien in einer konkreten Lernsituation bei Schülern unterschiedlicher Jahrgangsstufen	240
<i>Alexander Renkl/Silke Schworm</i> Lernen, mit Lösungsbeispielen zu lehren	259
Teil IV:	
Diagnose und Förderung von Interessen und Lernmotivation	
Förderung des Interesses und der Motivation von Schülerinnen und Schülern für mathematisch-naturwissenschaftliche Fächer: Zum Einfluss schulischer und familiärer Lehr-Lernumgebungen	271
<i>Elke Wild</i> Einleitung	272
<i>Elke Wild/Katharina Remy</i> Quantität und Qualität der elterlichen Hausaufgabenbetreuung von Drittklässlern in Mathematik	276
<i>Annette Upmeier zu Belzen/Helmut Vogt/Barbara Wieder/Franka Christen</i> Schulische und außerschulische Einflüsse auf die Entwicklungen von naturwissenschaftlichen Interessen bei Grundschulkindern	291
<i>Falko Rheinberg/Mirko Wendland</i> Veränderung der Lernmotivation in Mathematik: eine Komponentenanalyse auf der Sekundarstufe I	308

**Teil V:
Einstellungen und Werte als förderliche oder hinderliche Bedingungen
schulischer Leistungsfähigkeit**

Mathematisch-naturwissenschaftliche Fächer als Einstellungsobjekte: Einflüsse
von Makro- und Mesoebene auf die Einstellungsbildung 321

Bettina Hannover

Einleitung 322

Anna-Katharina Pelkner/Ralph Günther/Klaus Boehnke

Die Angst vor sozialer Ausgrenzung als leistungshemmender Faktor?

Zum Stellenwert guter mathematischer Schulleistungen unter Gleichaltrigen 326

Bettina Hannover/Ursula Kessels

Challenge the science stereotype! Der Einfluss von Technik-Freizeitkursen auf das

Naturwissenschaften-Stereotyp von Schülerinnen und Schülern 341

Juliane Strecker/Peter Noack

Wichtigkeit und Nützlichkeit von Mathematik aus Schülersicht 359

**Teil VI:
Schulforschung**

Evaluation und Feedback auf Klassen- und Schulebene 373

Hartmut Ditton/Bettina Arnoldt/Eva Bornemann

Entwicklung und Implementation eines extern unterstützenden Systems der

Qualitätssicherung an Schulen – QuaSSu 374

Jörg Doll /Manfred Prenzel

Einleitung in das Beiheft

Der Senat der Deutschen Forschungsgemeinschaft (DFG) hat im Mai 1999 die Einrichtung eines Schwerpunktprogramms beschlossen mit dem Thema „Die Bildungsqualität von Schule“. Dieses Forschungsprogramm untersucht das fachliche und fächerübergreifende Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten. Die Konzeption für ein entsprechendes Forschungsprogramm war von einer interdisziplinären Arbeitsgruppe vorbereitet worden (Prenzel/Merkens/Noack 1999). Die Ausschreibung des Schwerpunktprogramms fand starke Resonanz in zahlreichen Projektanträgen. Nach der Begutachtung konnte das Programm im Sommer 2000 die Arbeit aufnehmen. Nach Abschluss des ersten zweijährigen Förderungszeitraums in dem auf insgesamt sechs Jahre angelegten Schwerpunktprogramm präsentiert der vorliegende Band einen umfassenden Überblick über die Projekte und ihre Ergebnisse.

Schwerpunktprogramme der DFG bieten besondere Möglichkeiten für die Abstimmung und Bündelung von Forschungsvorhaben mit einer längeren Zeitperspektive. Es handelt sich also nicht um einen lockeren Verbund von thematisch verwandten Einzelprojekten, sondern um ein Forschungsunternehmen, das auf Zusammenarbeit angelegt und angewiesen ist. Den Rahmen für die Fragestellungen der einzelnen Projekte beschreibt die Konzeption des Schwerpunktprogramms (Prenzel/Merkens/Noack 1999), die der Ausschreibung zugrunde lag und die bestimmte Anforderungen an die Projekte definierte. Diese übergeordneten Fragestellungen und ihre Begründung werden im vorliegenden Einleitungskapitel skizziert. Sie betreffen die Forschungsziele, den Gegenstandsbereich sowie methodische Aspekte. Trotz dieser Vorgaben bleibt den einzelnen Projekten dennoch viel Raum, um eigene Fragestellungen, theoretische Perspektiven und Untersuchungszugänge zu verfolgen.

Die thematische Ausrichtung des Programms auf den Unterricht (Mathematik, Naturwissenschaften), auf schulische und außerschulische Kontexte eröffnet vielfältige Möglichkeiten, das Bedingungsgefüge für den Aufbau von Kompetenz bzw. für die Bildungsqualität von Schule zu analysieren. Um die Einordnung der Fragestellungen und die Abstimmung zwischen den Projekten zu erleichtern, soll die Strukturierung des Forschungsfeldes nachfolgend weiter ausdifferenziert werden, die in der Konzeption des Schwerpunktprogramms vorgezeichnet wurde. Daher wird der Orientierungsrahmen für das Schwerpunktprogramm unter Rückgriff auf vorliegende Modelle zur Schulleffektivität und Bildungsqualität weiterentwickelt, um Beziehungen zwischen dem Unterricht und schulischen wie außerschulischen Kontexten aufzuschlüsseln und die Relationen zwischen den einzelnen Projekten klären zu können. Im zweiten Teil des Einlei-

tungskapitels wird dieses Rahmenmodell vorgestellt, das zugleich eine Einordnung der Beiträge in diesem Band gestattet.

1. Befunde zur Qualität der mathematisch-naturwissenschaftlichen Bildung in Deutschland

Die Diskussion über die Qualität von Schule, die hierzulande seit geraumer Zeit geführt wird (z.B. Aurin 1990; Fend 1998; Helmke/Hornstein/Terhart 2000; Steffen/Bargel 1993; Terhart 2000; Tillmann 1989) hat durch die TIMS-Studie starke Impulse erhalten. Diese auf den mathematisch-naturwissenschaftlichen Bereich konzentrierte internationale Vergleichsuntersuchung beschreibt ein insgesamt ernüchterndes Bild des Leistungsstands deutscher Schülerinnen und Schüler in den Sekundarstufen. Ihre Leistungen liegen im unteren Mittelfeld beim Ländervergleich (Beaton u.a. 1996a, b). Wie die detaillierten Analysen der nationalen und internationalen Ergebnisse zeigen (Baumert u.a. 1997; Baumert/Bos/Lehmann 2000), bereiten insbesondere komplexere Aufgaben, die konzeptuelles Verständnis voraussetzen oder eine flexible Anwendung des Wissens verlangen, Schwierigkeiten. Im Bereich des naturwissenschaftlichen Denkens und Argumentierens werden ausgeprägte Schwächen offensichtlich. Es lassen sich weder in der Mathematik noch in den Naturwissenschaften ausgeprägte Spitzengruppen identifizieren. Relativ große Schülergruppen (ca. 25 Prozent) weisen fundamentale Wissensdefizite auf. Insgesamt sind verhältnismäßig geringe Leistungszuwächse über die Schuljahre festzustellen; stattdessen finden sich Hinweise auf ein abnehmendes Interesse an Mathematik, Physik und Chemie, besonders stark ausgeprägt bei den Mädchen.

Die aktuellen Ergebnisse der ersten Erhebungsrunde im „Programme for International Student Assessment“ (PISA) der OECD (1999, 2001) unterstreichen die Aussagen von TIMSS und liefern durch ein erweitertes und theoretisch wie methodisch fortgeschrittenes Erhebungskonzept zusätzliche Evidenz. Die Leistungen der deutschen Schülerinnen und Schüler liegen im Vergleich mit den OECD-Staaten unter dem internationalen Durchschnitt und dies gleichermaßen in den drei untersuchten Domänen der Lesekompetenz, der mathematischen und der naturwissenschaftlichen Grundbildung (Baumert u.a. 2001). Zusammen mit den ergänzenden nationalen Erhebungen liefert PISA, u.a. durch Analysen von Aufgaben und Kompetenzstufen, ein differenziertes Bild der Schwächen deutscher Schülerinnen und Schüler im Bereich der mathematischen und naturwissenschaftlichen Grundbildung (Klieme/Neubrand/Lüdtke 2001; Prenzel u.a. 2001). PISA zeigt aber durch die Untersuchung der Lesekompetenz (Artelt u.a. 2001) oder so genannter fächerübergreifender Kompetenzen (Cross-curricular Competencies), wie etwa der Fähigkeit, das eigene Lernen zu steuern (Artelt/Demmrich/Baumert 2001), dass die Schwächen deutscher Schülerinnen und Schülern nicht nur auf den mathematisch-naturwissenschaftlichen Bereich begrenzt sind. Vielmehr scheint die Qualität der Bildungsergebnisse in Deutschland generell problematisch zu sein. Der internationale Vergleich macht deutlich, dass Jugendliche ein (im Vergleich zu Deutschland) sehr viel höheres Kompetenzniveau erreichen können und dies auch bei einer ge-

ringeren Streuung der Leistungen. Die PISA-Befunde weisen darauf hin, dass es in anderen Ländern offenbar besser gelingt, Schülerinnen und Schüler mit unterschiedlichen Lernvoraussetzungen und Erfahrungen in ihrer Kompetenzentwicklung zu fördern.

Die Ergebnisse aus TIMSS und PISA zeigen einen deutlichen Abstand zwischen dem Leistungsniveau in Deutschland und anderen Industrienationen. Aus einer kriterienorientierten Perspektive betrachtet belegen die Befunde, dass Bildungsziele des mathematischen und naturwissenschaftlichen Unterrichts, die national wie international als bedeutsam gelten, hierzulande nicht erreicht werden. Dies gilt für Zielvorstellungen einer hinreichend breit gefächerten Allgemeinbildung (z.B. Klafki 1994, 1995), einer mathematischen bzw. naturwissenschaftlichen Grundbildung im Sinne von „Literacy“ (American Association for the Advancement of Science 1993; Freudenthal 1991; Gräber/Bolte 1997; National Council of Teachers of Mathematics 1989, 1991, 1995; National Research Council 1995); letztlich aber auch für die Ziele, die in deutschen Lehrplänen deklariert werden (Riquarts/Wadewitz 1999).

Anlass für Vorbereitung und Einrichtung eines Schwerpunktprogramms waren die alarmierenden TIMSS-Befunde und die daran anschließenden Fragen nach Ursachen für das Abschneiden deutscher Schülerinnen und Schüler sowie nach Handlungsmöglichkeiten. Die aktuellen internationalen Vergleichsstudien liefern methodisch solides und aussagekräftiges Datenmaterial über Kompetenzstand und Merkmale von schulischen wie außerschulischen Lernumgebungen oder von Bildungssystemen. Sie lassen aufgrund ihres Designs jedoch keine angemessene Prüfung von Erklärungsmodellen zu. Die Studien führen damit zu Fragen, die auch auf der Basis der vorliegenden, empirisch gesicherten Befunde der Unterrichts-, Schul- oder Sozialisationsforschung oft nicht zufriedenstellend beantwortet werden können.

2. Fachspezifisches Lernen und fächerübergreifende Kompetenzen

Bereits in der TIMS-Studie (Baumert u.a. 2000) waren anspruchsvolle Kompetenzen und ein Reihe weiterer Schülermerkmale untersucht worden, die für das Lehren und Lernen im Mathematik- oder Naturwissenschaftsunterricht relevant sind. Bei PISA wurden weitere umfassende Kompetenzbereiche und lernrelevante Konstrukte erhoben. Die PISA-Ergebnisse lassen überraschend starke Zusammenhänge zwischen unterschiedlichen Kompetenzbereichen erkennen (Baumert u.a. 2001). So erweist sich etwa die Lesekompetenz als Schlüsselqualifikation für den Aufbau von Wissen verschiedener Domänen. Die Befunde regen dazu an, die Beziehungen zwischen dem fachbezogenen Lernen und dem Aufbau fächerübergreifender Kompetenzen zu überdenken. Im Hinblick auf die Qualität von Bildungsprozessen und -ergebnissen gewinnen die Klärung und Untersuchung entsprechender Beziehungen an Bedeutung.

Wie didaktische Konzeptionen oder auch die Lehrpläne zeigen, zielt der Unterricht in Mathematik oder in den Naturwissenschaften auf Kompetenzen, die über das Fach hinausweisen: Einen besonderen Stellenwert erhalten (mathematische/naturwissenschaftliche) Kompetenzen, die grundlegend und relativ breit anwendbar sind bzw. ein

hohes Transferpotenzial zu besitzen scheinen. Der Unterricht zielt daneben auch auf die Entwicklung von Einstellungen oder Orientierungen, die mehr oder weniger eng mit dem Fach verbunden sind. Diese gewinnen aus schulischer Sicht u.a. deshalb Bedeutung, weil sie das weitere Lernen und Entscheiden beeinflussen.

Die Forschung zum Lehren und Lernen in der Mathematik und in den Naturwissenschaften hat in den letzten Jahren zum Beispiel auf fachbezogene Überzeugungen (Beliefs) aufmerksam gemacht (Köller/Baumert/Neubrand 2000; Schommer u.a. 1997; Törner/Grigutsch 1994). Sie wirken sich aus auf die Wahrnehmung des Unterrichtsstoffs oder von Fachinhalten und auf die Zuweisung von Bedeutung, und sie beeinflussen die kognitiven und motivationalen Lernzugänge.

Viele Studien zeigen weiterhin, dass der Unterricht in Mathematik und in den Naturwissenschaften für die Entwicklung des Selbstkonzepts einen herausgehobenen Stellenwert hat. Bezogen auf die besonderen Anforderungen dieser Schulfächer erleben sich die Schülerinnen und Schüler als mehr oder weniger befähigt und kompetent – in Relation zu den Fortschritten anderer, aber auch im Vergleich mit anderen Inhalten und Fächern. Durch entsprechende Vergleichsprozesse entwickeln Schülerinnen und Schüler (mit beträchtlicher Differenz zwischen den Geschlechtern) Vorstellungen über ihre eigenen Fähigkeiten und Präferenzen. In engem Zusammenhang dazu erfolgt die Entwicklung von fach- und gegenstandsbezogenen Wertorientierungen und Interessen (Hofmann/Häußler/Lehrke 1998; Krapp 1998, 2001; Krapp/Prenzel 1992; Schiefele/Wild 2000). Das fachbezogene Interesse wiederum beeinflusst nicht nur das Lernen und die (auch emotionale) Beteiligung im Unterricht, sondern vor allem die Entscheidung bei der Wahl von Kursen, Studienfächern oder Ausbildungsgängen.

Die Forschung zum Lernen und Lehren in Mathematik und in den Naturwissenschaften hat sich bisher intensiv mit Unterschieden in Überzeugungen, Interessen, motivationalen Orientierungen oder im Selbstkonzept befasst und deren Bedingungen oder Wirkungen analysiert. Doch leuchten viele Studien oft nur Teilbereiche aus; selten untersucht wurden bisher das Wechselspiel von kognitiven, motivationalen und emotionalen Prozessen oder die Zusammenhänge zwischen fachlichem Lernen und der Entwicklung von fachbezogenen Einstellungen. Unberücksichtigt bleibt meist, dass entsprechende Lern- oder Entwicklungsprozesse im Mathematik- oder Naturwissenschaftsunterricht nicht losgelöst von der Wahrnehmung anderer Fächer analysiert werden können. Aus der Schülerperspektive liefert der Fächervergleich wichtige Anhaltspunkte für die Beurteilung, wie relevant, schwierig oder interessant der Unterricht in der Mathematik und in den Naturwissenschaften mit den jeweiligen Stoffgebieten und Zugangsweisen ist. Diese Überlegungen sprechen dafür, das Schwerpunktprogramm relativ umfassend auf die Entwicklung fachbezogener Kompetenzen und damit verbundener Überzeugungen, Orientierungen oder Einstellungen auszurichten, die für die Entwicklung der Persönlichkeit bedeutsam sind (Prenzel/Merkens/Noack 1999).

Ein weiterer Grund, den Einfluss anderer Fächer zu analysieren, resultiert aus aktuellen Forderungen und aus dem von der Schule selbst vertretenen Anspruch, gezielt auch fächerübergreifende Kompetenzen zu fördern. Bei diesen übergeordneten Kompetenzen, zum Teil wird von „Schlüsselqualifikationen“ oder von „cross curricular compe-

tencies“ (OECD 1999) gesprochen, handelt es sich meist um Syndrome von kognitiven, motivationalen und metakognitiven Fähigkeiten. Sie betreffen z.B. die Steuerung des eigenen Lernens oder das Verhalten in sozialen Lern- oder Arbeitssituationen.

Der Ort für die Entwicklung solcher übergeordneter Kompetenzen ist der Fachunterricht. Gleichgültig, ob von der Lehrkraft intendiert oder nicht, beeinflusst der mathematisch-naturwissenschaftliche Unterricht den Aufbau von metakognitiven Kompetenzen, von Lernstrategien, von Problemlösekompetenzen, von motivationalen Orientierungen, von Selbstwirksamkeits- und Kontrollüberzeugungen, von sozialen Kompetenzen einschließlich Kooperationsfähigkeit bis hin zu grundlegenden Handlungs- und Wertorientierungen (Baumert 1993; Lambert/McCombs 1997; Mandl/Friedrich 1992; Weinert 1998).

Prinzipiell bietet der Unterricht in Mathematik und in den Naturwissenschaften günstige Gelegenheiten zur Untersuchung der Frage, inwieweit und auf welche Weise im Fachzusammenhang generalisierbare Kompetenzen aufgebaut werden, die auch über die Schule hinaus als Basis für lebenslanges Lernen dienen können (Achtenhagen/Lempert 1999). Damit verbunden ist die Frage, wie die Entwicklung solcher fächerübergreifender Kompetenzen gezielt gefördert werden kann. So wird oft dem Mathematikunterricht die Funktion zugesprochen, Heuristiken des Problemlösens zu entwickeln (Polya 1965; Schoenfeld 1992). In ähnlicher Weise wird eine breite Förderwirkung naturwissenschaftlicher Denk- und Arbeitsweisen auf das Problemlösen vermutet. Bisher ist jedoch nicht hinreichend geklärt, inwieweit ein Transfer bzw. Anpassen von Heuristiken oder von Lernstrategien auf andere Inhaltsgebiete und Anforderungssituationen gelingt bzw. gezielt unterstützt werden kann.

Das Schwerpunktprogramm konzentriert sich so auf das Lehren und Lernen im Mathematik- und Naturwissenschaftsunterricht, um Fragen zur Bildungsqualität von Schule an einem eingrenzbaaren Gegenstandsbereich beispielhaft klären zu können. Die Fokussierung auf den Mathematik- und Naturwissenschaftsunterricht verlangt aber auch, fächerübergreifende Aspekte des Lehrens und Lernens in diesen Fächern nicht aus dem Blick zu verlieren. Von den Projekten im Schwerpunktprogramm wird erwartet, dass sie sich mit dem Aufbau von Kompetenz im Mathematik- oder Naturwissenschaftsunterricht befassen: mit fachlicher Kompetenz in einem engen Sinn und mit fächerübergreifenden Kompetenzen bzw. mit fachbezogenen Persönlichkeitsmerkmalen. Das heißt auch, dass die einzelnen Projekte theoriegeleitete Relationen zwischen diesen Konstrukten untersuchen.

3. Zur Bedeutung von Unterricht und schulischen wie außerschulischen Lernumwelten in Schuleffektivitätsmodellen

In diesem Abschnitt wird kurz auf die Bedeutung des mathematisch-naturwissenschaftlichen Unterrichts für die Bildungsqualität von Schule eingegangen. Daran anschließend werden exemplarisch vier Modelle der Schuleffektivität und Bildungsqualität vorgestellt und hinsichtlich ihres Potenzials kritisch diskutiert, theoretische Bausteine

zur Entwicklung eines Rahmenmodells für das Schwerpunktprogramm zu liefern, das am Ende dieses Abschnitts vorgestellt wird.

Unterricht bietet besondere Gelegenheiten, metakognitive und selbstregulatorische Kompetenzen aufzubauen, aber auch Arbeits- und Lerntechniken zu erwerben. Koooperative Arbeitsformen können das für mathematisch-naturwissenschaftliche Fächer notwendige verständnisvolle Lernen in besonderer Weise unterstützen, den Aufbau sozialer Kompetenzen und Fertigkeiten fördern (Brown/Campione 1990) und die Schülerinnen und Schüler als Mitglieder einer Lerngemeinschaft, Fach- oder Expertenkultur ansprechen und gewinnen (Brown 1997; Cobb 1994). Mathematik und Naturwissenschaften sind deshalb bevorzugte Beispielanwendungen für sogenannte „powerful teaching-learning environments“ (DeCorte u.a. 1996; Mandl/Prenzel 1991; Sherwood u.a. 1998), die auf die Entwicklung fachbezogener wie fächerübergreifender Kompetenzen abzielen. Die in der internationalen Forschung entwickelten Konzepte und Modelle zu den Wirkungen mathematisch-naturwissenschaftlichen Unterrichts (Duit/Häußler 1997; Fraser/Tobin 1998; Treagust/Duit/Fraser 1996) werden im Schwerpunktprogramm empirisch geprüft und weiterentwickelt.

Qualitätskriterien für Unterricht und Schule sind multidimensional und auf unterschiedlichen Ebenen angesiedelt. Ausgehend von dem Mehrebenencharakter des Schulwesens hat Fend (2001) ein Klassifikationssystem für Qualitätskriterien entwickelt (s. Tab. 1), in dem er den vier Ebenen des Schulwesens, dem Schulsystem, der Schulebene, der Klassenebene und der Personenebene (Lehrpersonen, Schülerinnen und Schüler), ebenenspezifische Qualitätskriterien zuweist. Tabelle 1 verdeutlicht die Multikriterialität von Bildungsqualität innerhalb und zwischen diesen Ebenen. Es fällt auf, dass von Fend auf Schul-, Klassen- und Personenebene zusätzlich zu den Qualitätskriterien charakteristische *Problemkriterien* (z.B. Vandalismus, Alkoholkonsum, Anomie) benannt werden, die damit in ihrer Wichtigkeit besonders akzentuiert werden. Darüber hinaus wird auch der *Qualität der sozialen Beziehungen* und der Interaktionen zwischen einzelnen Schülerinnen und Schülern und einzelnen Lehrpersonen (auf Personenebene), zwischen Klassen und einzelnen Lehrpersonen (auf Klassenebene) oder zwischen Lehrpersonen eines Kollegiums (auf Schulebene) zentrale Bedeutung zugewiesen.

Die nachfolgend skizzierten Modelle der Unterrichts- und Schulqualität gehen über diese Klassifikation von Qualitätskriterien hinaus und formulieren modellhafte Variablenzusammenhänge mit dem Ziel, Erklärungen für eine hohe oder niedrige Bildungsqualität zu liefern. Die Modelle unterscheiden sich u.a. darin,

- dass sie *unterschiedlich große Ausschnitte* aus dem Geflecht der möglichen Bedingungskonstellationen von Bildungsqualität betrachten,
- ob sie eine *unikriteriale* oder *multikriteriale Wirksamkeit* (Helmke/Schrader 2001) von Bildung im Sinne des Verfolgens vielfältiger Ziele im kognitiven, motivationalen, volitionalen und sozialen Bereich berücksichtigen,
- welche Bedeutung sie *distalen Bedingungen* (z.B. Faktoren auf einer höheren Ebene) in Mehrebenenmodellen schulischen Lernens (Ditton 1998; Saldern 2001) im Unterschied zu proximalen Bedingungen (z.B. Unterrichtsbedingungen) zuweisen,

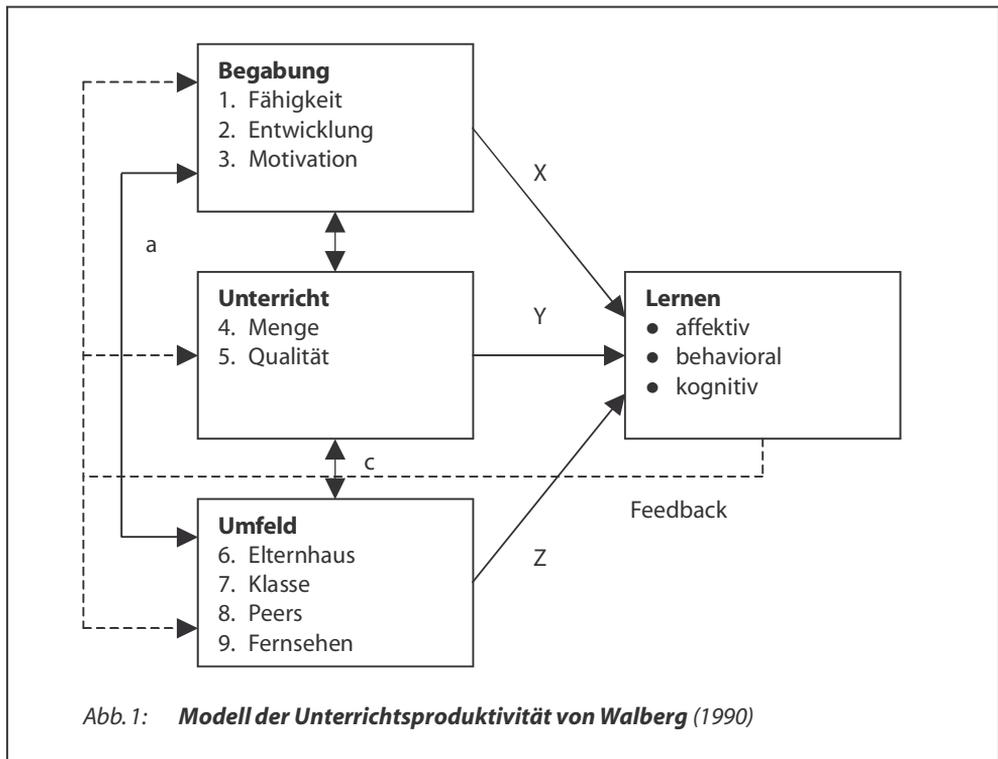
Tab. 1: **Ebenenspezifische Qualitätskriterien nach Fend (2001)**

<p>Schulsystem</p> <ol style="list-style-type: none"> 1. Effizienz und Leistungsfähigkeit 2. Sozialpolitische Verantwortbarkeit Chancengleichheit 3. Flexibilität der Schullaufbahngestaltung <ul style="list-style-type: none"> - Lehrgangcharakter - Kanon, Orientierungsbereiche - Übersicht und Planungsvorgaben (kein Abschluss ohne Anschluss) 4. Versorgungsdichte <ul style="list-style-type: none"> - Angebotsqualität und Ausstattung 5. Humanität <ul style="list-style-type: none"> - Regelungen von Freiheiten, Beteiligungen und Verpflichtungen - Überfachliche Wirkungen 	<p>Schulebene</p> <ol style="list-style-type: none"> 1. Schulpolitik <ul style="list-style-type: none"> - Außenpolitik: lokale Einbettung - Innenpolitik: Verwaltungseffizienz und soziale Integration - Kinder- bzw. jugendspezifische Gestaltung des Schullebens - Leistungsniveau 2. Kollegium <ul style="list-style-type: none"> - Pädagogisches Ethos - Konfliktlösungsfähigkeit - Arbeitsethos - Qualitätsbewusstsein in Bezug auf humanes Zusammenleben - Kooperationsintensität - Gestaltungswillen 3. Schülerschaft <ul style="list-style-type: none"> - Vandalismus und Aggression oder prosozialer Verhaltensstil und moralische Atmosphäre - Schuldistanz oder Vertrauen - Verschüchterungsgrad oder Offenheit
<p>Klassenebene</p> <ol style="list-style-type: none"> 1. Allgemein <ul style="list-style-type: none"> - Leistungsniveau - Motivation und Leistungsbereitschaft - Problembelastung: Disziplinprobleme, Rauchen, Alkoholkonsum, Delinquenz 2. Qualität sozialer Beziehungen zu den Lehrern <ul style="list-style-type: none"> - Gesprächsfähigkeit mit den Lehrern - Distanz und Gleichgültigkeit in den Schüler-Lehrer-Beziehungen - Konformitätsdruck und Verschüch- terung durch die Lehrerschaft - Verstrickungen mit einzelnen Lehrern - Ablehnung und Distanz 3. Qualität der sozialen Beziehungen der Schüler untereinander <ul style="list-style-type: none"> - Konformitätszwang - Rivalitätsgrad und Ostrazierungs- tendenzen - Isolation oder Vercliquung - Rohheit und Grobheit als informelle Erfolgswege 	<p>Personebene</p> <ol style="list-style-type: none"> 1. Lehrer <ul style="list-style-type: none"> - Kompetenzen - Mentalitäten und Weltbilder - Beziehungsmerkmale aus der Schülerperspektive: <ul style="list-style-type: none"> - Ablehnung oder Distanz - Angst und Furcht (machtorientierte Lehrer) - Anomie (chaotische Lehrer) - Regelungsunfähigkeit (hifflöse Lehrer) 2. Schüler <ul style="list-style-type: none"> - Kompetenzen - Mentalitäten - Leistungsbereitschaften - Arbeitshaltungen - Zuverlässigkeit, Pflichtbewusstsein - Ich-Stärke - Soziale Integrationsfähigkeit - Führungsfähigkeit - Verantwortungsbereitschaft - Soziale Einsatzbereitschaft

- welche *schulischen* und *außerschulischen Lernumwelten* (Familie, Gleichaltrigen-
gruppe, Massenmedien) auf welchen Systemebenen (Mikro-, Meso- und Exoebene
sensu Bronfenbrenner 1978, 1981) berücksichtigt werden, und
- ob vor allem *Zustände* (d.h. Produkte oder Ergebnisse von Bildung) betrachtet wer-
den oder ob die *Prozesse* und *Mechanismen* spezifiziert werden, die diese Zustände
herbeiführen und die Ansatzpunkte für Interventionen bieten.

Das Modell der Unterrichtsproduktivität von Walberg (1990) in Abbildung 1 kann als Erweiterung des einflussreichen Modells schulischen Lernens von Carroll (1963) um die Variable der psychischen Umwelt von Schülerinnen und Schülern betrachtet werden. Dieses Modell der Unterrichtsproduktivität beruht auf Metaanalysen und unterscheidet neun Faktoren, die für die Produktivität von Unterricht wichtig sind. Diese neun Faktoren werden in die drei Bereiche Schülerkompetenzen, Unterrichtsvariablen und psychologisches Umfeld eingeteilt. Das hier besonders interessierende *psychologische Umfeld* bezieht sich auf Aspekte der häuslichen Umwelt (z.B. elterliche Hausaufgabenkontrolle), des Klassenklimas (z.B. Klassenkohäsion), der Gleichaltrigengruppe (z.B. Bildungsaspiration des Freundeskreises) und der Mediennutzung (z.B. Fernsehnutzungsdauer).

Das Modell der Unterrichtsproduktivität ist ein *Einebenenmodell* für multikriterialen Schulerfolg, das ausschließlich direkte Effekte erfasst zwischen Schülerkompetenzen, Unterrichtsvariablen und Variablen des psychologischen Umfelds einerseits und der Schulleistung andererseits (vgl. die Pfeile mit Großbuchstaben in Abb. 1). Das Modell lässt allerdings auch Wechselwirkungen zwischen allen Produktivitätsfaktoren zu und Rückwirkungen vom Schulerfolg auf die Produktivitätsfaktoren, d.h. dass beispielsweise das Leistungsniveau einer Klasse auf die Strukturiertheit des Unterrichts in dieser Klasse zurückwirken kann.



Aufgrund seiner Entstehung aus Metaanalysen sollten sich nach Walberg keine weiteren Produktivitätsfaktoren identifizieren lassen (z.B. keine Faktoren der Schulorganisation), die einen zusätzlichen direkten Effekt auf die Schulleistung haben. Alle sonstigen denkbaren Produktivitätsfaktoren sollten bei Gültigkeit des Vollständigkeitsprinzips nur *indirekt* die Schulleistung beeinflussen, vermittelt über die berücksichtigten neun Faktoren.

Das Modell stimmt mit dem Ziel des Schwerpunktprogramms überein, die Bedeutung unterschiedlicher Lernumwelten für die Bildungsqualität zu untersuchen. Im Konzept des psychologischen Umfeldes werden auf unterschiedlichen Ebenen angeordnete Merkmale aus unterschiedlichen Lernumwelten in ihrer Prädiktionskraft für den schulischen Lernerfolg berücksichtigt. Das Modell der Unterrichtsproduktivität ist jedoch in mehreren Aspekten nicht als Rahmenmodell des Schwerpunktprogramms geeignet:

- Es ignoriert solche Bedingungen von Bildungsqualität, die in der schulischen Lernumwelt liegen,
- es ist ein Prädiktormodell kein Prozessmodell, d.h. dass zwar Zustände jedoch nicht die Prozesse modelliert werden, die Bildungsqualität fördern oder hemmen,
- es bildet den Mehrebenencharakter des Bildungssystems nicht adäquat ab, und
- es berücksichtigt nicht die Bedeutung kultureller Hintergrundfaktoren.

Scheerens (1990) „integratives Modell der Schuleffektivität“ (Abb. 2, S. 18) ist umfassender konzipiert als das Modell der Unterrichtsproduktivität, denn es zielt darauf ab, Forschungsergebnisse zur Leistungsfähigkeit von Unterricht *und* von Schule in einem Modell zu integrieren. Es ist ein *Mehrebenenmodell* mit drei Modellebenen: den Ebenen des Kontexts, der Schule und der Klasse. Es geht von der Kernthese aus, dass Bedingungen auf einer höheren Ebene Prozesse auf einer niedrigeren Ebene unterstützen oder erleichtern können. Es betont außerdem durch die Verwendung der Konzepte Input, Prozess und Output den dynamischen Charakter (Ditton 2000) von Bildung, sodass Prozesse oder Mechanismen den Modellkern bilden, durch die Eingangsinvestitionen auf unterschiedlichen Ebenen in erzielte unikriterial konzipierte Ergebnisse (Schülerleistungen) überführt werden.

Scheerens/Bosker (1997, S. 309ff.) haben allerdings im Abstand von einigen Jahren aufgrund eigener empirischer Studien und weiterer Metaanalysen zur Schuleffektivitätsforschung die Gültigkeit ihrer ursprünglichen zentralen Hypothese, dass Faktoren der Schulorganisation die Erfolge auf der Ebene einzelner Schulklassen oder einzelner Schülerinnen und Schüler erleichtern oder verbessern können, stark in Zweifel gezogen. Sie stellen zusammenfassend fest (Scheerens/Bosker, 1997, S. 309ff.), dass es *keine* empirischen Belege für *direkte* Effekte von Variablen der Schulorganisation auf die Ebene der Schülerleistungen gibt und dass auch die wenigen Studien negativ ausgegangen seien, die indirekte Effekte von der Schulebene auf die Schülerleistung, vermittelt über die Klassenebene, getestet hätten. Effekte von Variablen der Schulorganisation ließen sich nach Scheerens/Bosker nur dann nachweisen, wenn sie simultan auf Schul- und Klassenebene eingeführt worden seien wie etwa die Implementation eines neuen Curricu-

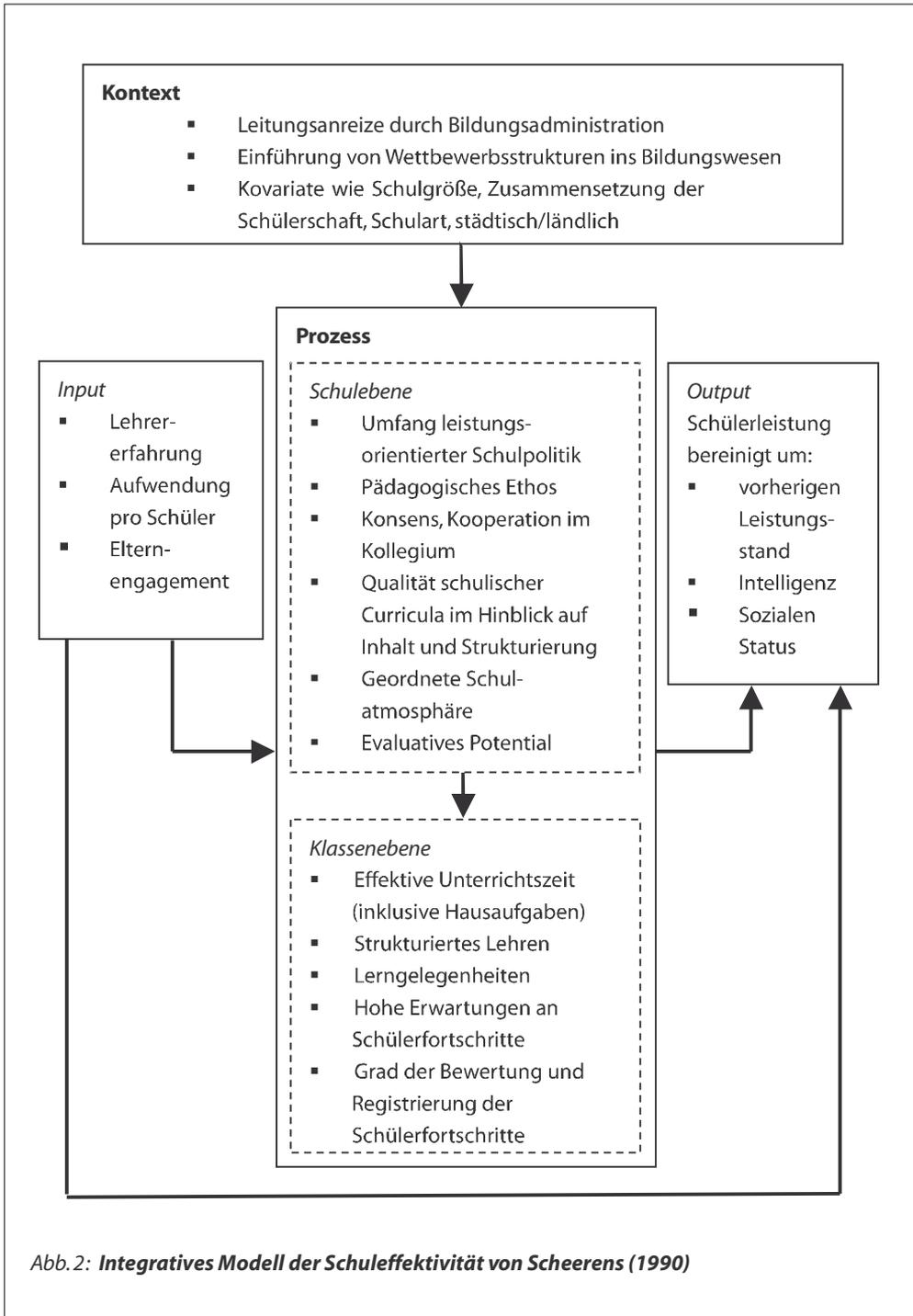


Abb.2: *Integratives Modell der Schuleffektivität von Scheerens (1990)*

lums. Scheerens/Bosker versuchen die genannten negativen Ergebnisse zur Unterstützungshypothese ins Positive zu wenden, indem sie sich dem Subsidiaritätsprinzip anschließen und den Lehrpersonen in dem hierarchisch gegliederten Schulsystem eine möglichst große Autonomie von den Vorschriften höherer Systemebenen zugestehen und die Möglichkeit zu selbstorganisierter Unterrichtsgestaltung hervorheben.

Das „integrative Modell der Schuleffektivität“ umreißt einen theoretischen Rahmen, der insbesondere durch seinen Mehrebenencharakter und durch die Modellierung von Prozessen auf Schul- und Klassenebene, die Ansatzpunkte für Interventionen bieten können, theoretische Bausteine enthält, die gut zu dem Vorhaben des Schwerpunktprogrammes passen. Allerdings sind mindestens zwei zentrale Ziele des Schwerpunktprogrammes, die multikriteriale Erfassung von Bildungsqualität und die Modellierung des Einflusses von außerschulischen Lernumwelten, auch mit diesem Rahmenmodell nicht umsetzbar.

Theoretisch noch umfassender ist der Anspruch des TIMSS-Rahmenmodells „schulischer Lehr-Lern-Gelegenheiten“ von Schmidt/Cogan (1996) mit der Unterscheidung von vier Systemebenen (Schulsystem, Schule, Klasse und Schülerin/Schüler), das als Bezugsrahmen zur Konstruktion des Kontextfragebogens der TIMS-Studie entwickelt wurde. Im Zentrum dieses Modells (Abb. 3) steht die Transformation eines von der Bildungsadministration *intendierten Curriculums*, das nationale und regionale Bildungsziele umsetzt, zuerst in ein auf der Ebene von Schule und Unterricht *implementiertes Curriculum*, das durch die Ziele einzelner Schulen und Lehrpersonen ergänzt und konkretisiert wird, und schließlich in ein *erreichtes Curriculum* auf der Ebene der Leistungsergebnisse der Schülerinnen und Schüler mündet, das seinerseits durch die Lern- und Lebensziele der Schülerinnen und Schüler mitbestimmt wird. Dieses TIMSS-Mehrebenenmodell verdeutlicht insbesondere zwei theoretisch zentrale Aspekte: Es betont die Intentionalität und Zielgerichtetheit von Lehr-Lernprozessen auf allen Ebenen des Schulwesens, und es erfasst die Vernetzungen zwischen den vier Ebenen des Bildungssystems durch das Konzept der *Curriculum-Transformationen* beim Übergang von einer Ebene auf eine andere (auch Ditton 2000). Aus Sicht des Schwerpunktprogramms stellt das TIMSS-Modell einen Rahmen zur Verfügung, der durch das Schwerpunktprogramm umfassender ausgefüllt werden kann als durch die internationale Schulvergleichsstudie TIMSS selbst. Dies ergibt sich daraus, dass die TIMS-Studie

- als Fragebogenstudie mit begrenztem Fragebogenumfang wichtige Variablen nur durch wenige Indikatoren erfassen konnte,
- auf Grund ihres querschnittlichen Designs keine kausalen Zusammenhänge überprüfen konnte und
- das implementierte Curriculum nicht hinreichend erfasst hat.
- Schließlich finden auch hier außerschulische Lernumwelten zu wenig Berücksichtigung, sodass etwa Eigenschaften der familiären Lernumwelt nur verkürzt als Schülereigenschaften berücksichtigt werden.

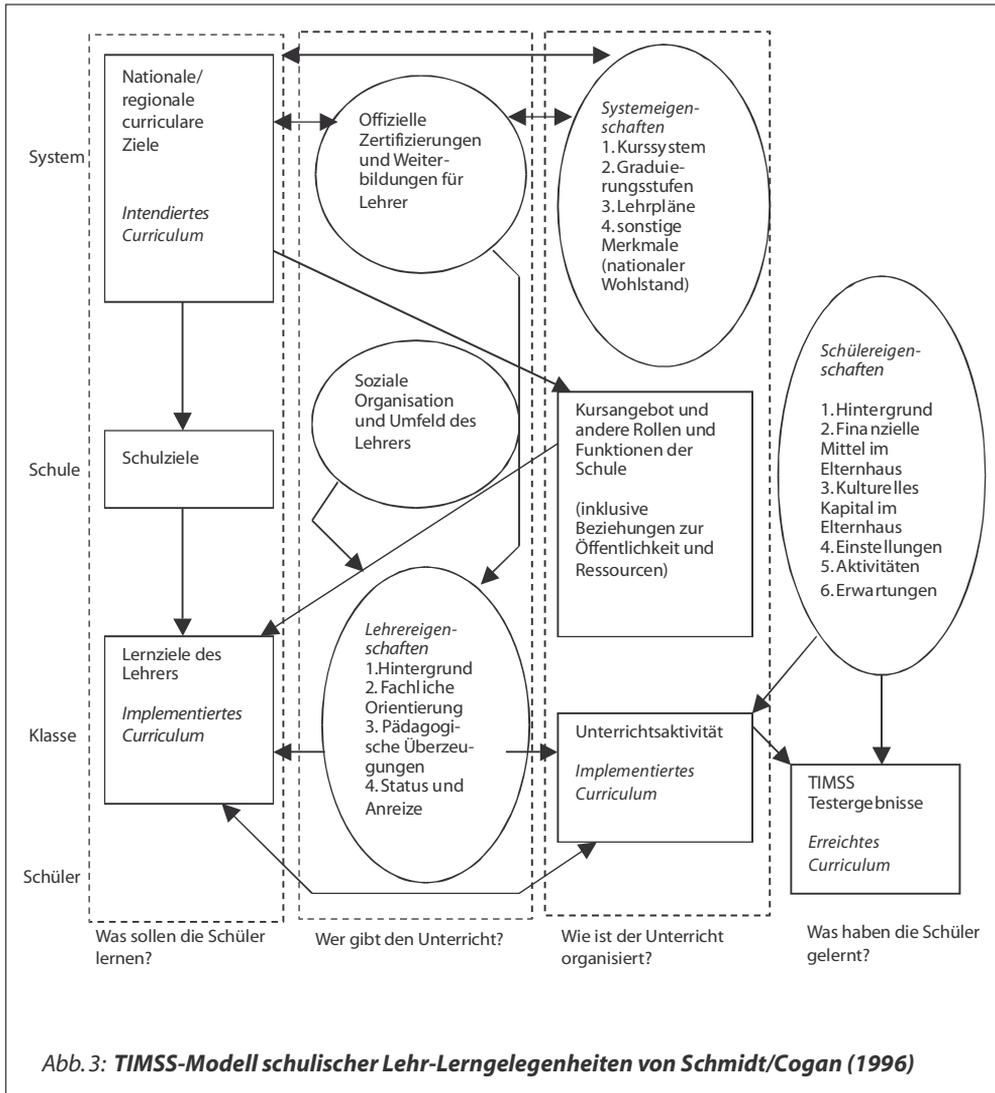


Abb. 3: TIMSS-Modell schulischer Lehr-Lerngelegenheiten von Schmidt/Cogan (1996)

Das vierte „Modell der Entwicklungsumwelten“ entstammt dem Ansatz der *ökologischen Sozialisationsforschung*, der miteinander systemisch vernetzte Entwicklungsumwelten auf Mikro-, Meso-, Exo- und Makroebene unterscheidet (Bronfenbrenner 1978, 1981). Dieser Ansatz betrachtet eine Person als dynamische Einheit, die einerseits ihre Umwelt im Laufe der Entwicklung fortschreitend in Besitz nimmt und umformt, die sich aber andererseits auch den Umwelteinflüssen anpasst. Die Umwelt einer Person umfasst mehrere Lebensbereiche und die Verbindungen zwischen diesen Bereichen. Dabei bezieht sich das *Mikrosystem* auf die unmittelbaren Lebensbereiche, in denen das sich entwickelnde Individuum lebt (z.B. Familie, Schulklasse, Gleichaltrigengruppe). Das *Mesosystem* besteht aus zwei oder mehreren Lebensbereichen, an denen das sich

entwickelnde Individuum aktiv beteiligt ist (für ein Kind etwa die Beziehungen zwischen Elternhaus, Schule und Spielkameraden). Die Wechselwirkungen zwischen diesen Lebensbereichen (z.B. wechselseitige Behinderungen oder Unterstützungen) bilden ein eigenes entwicklungsrelevantes *System von Mikrosystemen*. Unter dem *Exosystem* wird ein Lebensbereich oder mehrere Lebensbereiche verstanden, an denen die sich entwickelnde Person nicht selbst beteiligt ist, in denen aber Ereignisse stattfinden, die ihren Lebensbereich beeinflussen oder von ihrem Lebensbereich beeinflusst werden (für ein Kind bilden der Arbeitsplatz der Eltern oder verbindliche Lehr- und Stundenpläne Teile des Exosystems). Das *Makrosystem* schließlich bezieht sich auf Übereinstimmungen in den genannten Systemen niedrigerer Ordnung (Mikro-, Meso- und Exosystemen), die auf der Ebene der Gesamtkultur oder von Subkulturen bestehen und denen geteilte Überzeugungssysteme oder Ideologien zugrunde liegen. Die moderne Arbeitsstruktur, die Ideologie persönlicher Autonomie und Kontrolle und der Fortschrittsglaube sind nach Oerter (1998) exemplarische Bestandteile des Makrosystems unseres westlichen Kulturkreises.

Bronfenbrenners Systematisierung von kindlichen Entwicklungsumwelten als hierarchisch strukturierten und vernetzten Systemen hat die Sozialisationsforschung stark beeinflusst. Im Schwerpunktprogramm wird dieser Ansatz mit Ansätzen der Unterrichtsforschung zur Gestaltung optimaler Lehr-Lernumgebungen verbunden. In diesem Zusammenhang ist auf das methodische Problem hinzuweisen, dass bei der Überprüfung von Mehrebenenmodellen sensu Bronfenbrenner hierarchisch verschachtelte Datenstrukturen entstehen, für deren statistische Auswertung erst seit einigen Jahren adäquate Programme zur Verfügung stehen (Ditton 1998; Saldern 2001).

Diese Überlegungen zu systemisch vernetzten Entwicklungsumwelten führen zum „BIQUA-Rahmenmodell schulischer und außerschulischer Lernumwelten“ (vgl. Abb. 4). Das Modell greift auf die soeben diskutierten Ansätze zurück und orientiert sich an der Rahmenkonzeption von PISA (Baumert/Stanat/Demmrich 2001, S.33). Das systemische BIQUA-Modell unterscheidet mit der Schule, der Schulklasse, der Familie, der Gleichaltrigengruppe und den alten und neuen Medien fünf miteinander vernetzte Lernumwelten. Bedingungen und Prozesse in jeder dieser Lernumwelten oder in der Wechselwirkung zwischen zwei oder mehreren Lernumwelten (Mesosystemebene), können direkt oder indirekt sowohl die individuellen *Lernvoraussetzungen* der Schülerinnen und Schüler im kognitiven, motivationalen, volitionalen, affektiven und sozialen Bereich mitbestimmen als auch das *Unterrichtsgeschehen* beeinflussen. Individuelle Lernvoraussetzungen und Unterrichtsprozesse beeinflussen dann individuelle Verarbeitungsprozesse, die zu kurz-, mittel- und langfristigen Veränderungen im Sinne von fachlichen und fächerübergreifenden Wissensbeständen und Kompetenzen führen.

Die unmittelbar an den Fachunterricht in der Klasse angrenzende Lernumwelt bildet die jeweilige Schule. Bedeutsame Bedingungen der *schulischen Lernumwelt* betreffen etwa die Schulkultur, das Schulklima, das Profil der Schule, Aktivitäten zur Fortbildung oder Schulentwicklung im Kollegium, Ausstattung und Räumlichkeiten, Arbeitsgruppen und andere Gelegenheitsstrukturen für Schülerbeteiligungen; die „Brückenbildung“ der Schule kann erfolgen durch Kooperationen mit der Gemeinde, mit Betrieben, Ver-

einen, Forschungseinrichtungen und nicht zuletzt mit den Eltern (Fend 1996, 1998; Hofer 2000; OECD 1989, 1995; Terhart 1987).

Von den außerschulischen Lernumwelten ist an erster Stelle das *Elternhaus* zu nennen. Im Hinblick auf das schulische Lernen sind nicht nur die Anregungsbedingungen, das intellektuelle Klima und die Arbeitsgewohnheiten zu Hause bedeutsam, sondern ebenso die Erwartungen, Unterstützungen und Rückmeldungen der Eltern, die emotionale Absicherung und das Vertrauen, die Verständigung über schulische Belange und Probleme, und das Engagement in Schulfragen, das bis zur Mitwirkung an der Gestaltung von Schule reichen kann (Coleman 1998; Krumm 1996, 1998; Pekrun 1997).

Als weitere Lernumwelt ist die „dritte Hauptinstanz der Sozialisation“ (Fend/Stöckli 1997, S. 4), die *Gleichaltrigen- und Freundesgruppe*, in Betracht zu ziehen: Die Schule bietet Gelegenheitsstrukturen zur Cliquen- und Freundschaftsbildung, im Klassenzusammenhang wird, häufig abhängig von Leistungen, ein sozialer Status erworben oder zugewiesen. Auf der anderen Seite entwickeln Cliquen im Verlauf des Jugendalters Präferenz- und Normenstrukturen, die – mit Verbindlichkeit und hohem Gewicht – die Einstellungen gegenüber der Schule beeinflussen. Bezogen auf schulische Bildungsprozesse sind u.a. bedeutsam: die Identitätsbildung in Cliquen und Freundeskreisen mit bestimmten sozialen Umgangsformen, die gegenseitige Stützung und emotionale Stabilisierung, die inhaltlichen Interessen und die Art der Freizeitgestaltung, die Wertschätzung von Schulleistungen und die Einstellung zur Schule generell (Fend 1997; Mac Iver/Reuman/Main 1995; Noack 1990; Youniss/Hyanie 1992).

Es gibt weitere Lernumwelten, die für das fachliche und fächerübergreifende Lernen in der Schule und speziell im mathematisch-naturwissenschaftlichen Unterricht wichtig sind. Zu nennen sind u.a. die *Medien* (Kerres 2001; Moser 2000; Winterhoff-Spurk 1999), aber auch die Berufs- und Ausbildungssituation oder besondere Lerngelegenheiten, die sich aufgrund bestimmter Lebenssituationen (z.B. Migration) konstituieren. Das Makrosystem als übergreifender kultureller Hintergrund bildet einen noch umfassenderen Kontext, der für das Lernen in der Schule ebenfalls eine nicht zu vernachlässigende Rolle spielt. Dies belegen nicht zuletzt die aktuellen internationalen Vergleichsstudien zum mathematisch-naturwissenschaftlichen Unterricht (Schümer 1998).

Weiterführende Erkenntnisse zur Verbesserung der Bildungsqualität können von theoretisch abgeleiteten Mehrebenenmodellen erwartet werden, die das Zusammenwirken von Unterrichtsbedingungen und Lernumwelten bei der Herausbildung von fachlichen und fächerübergreifenden Kompetenzen theoretisch spezifizieren. Schulische wie außerschulische Lernumwelten sind einerseits *Anwendungs- und Erprobungsfelder* für fachliche und fächerübergreifende Kompetenzen, die im mathematisch-naturwissenschaftlichen Unterricht aufgebaut wurden. Sie sind andererseits *Lern- und Entwicklungsfelder* für Kompetenzen und Persönlichkeitsmerkmale, die als personspezifische Lernvoraussetzungen die Lehr- und Lernerfolge im mathematisch-naturwissenschaftlichen Unterricht bedingen. Insgesamt betrachtet werden die Ähnlichkeiten in den Anliegen wie Strukturen und die Verknüpfungen bzw. Abstimmungen zwischen Unterricht und schulischen wie außerschulischen Lernumwelten zu ausschlaggebenden Bedingungen für die schulische Bildungsqualität.

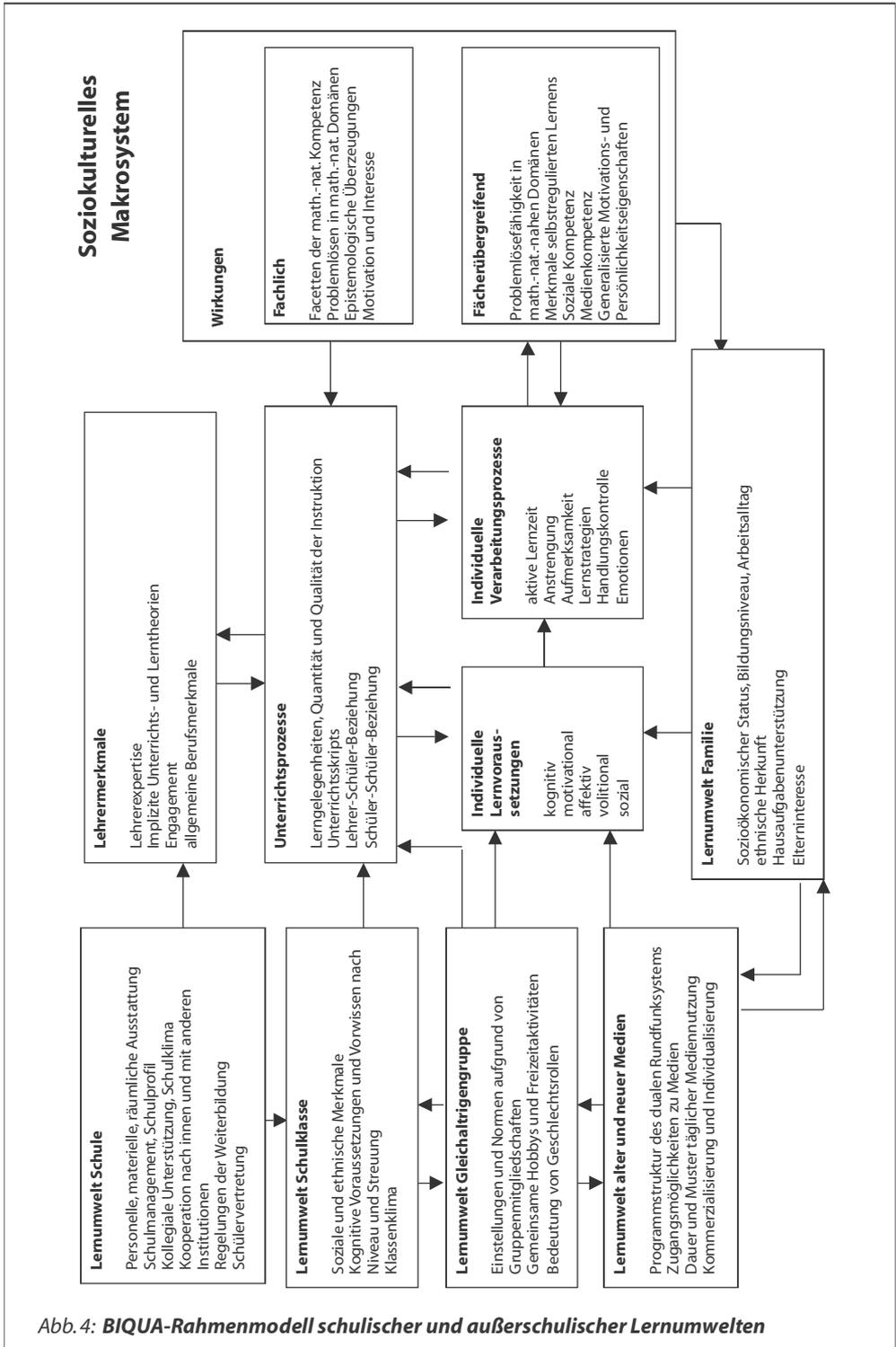


Abb. 4: BIQUA-Rahmenmodell schulischer und außerschulischer Lernumwelten

4. Zur Struktur des Schwerpunktprogramms BIQUA und zur Gliederung des Beiheftes

Am Programm beteiligen sich Projekte aus ganz Deutschland, die Fragestellungen aus der Mathematik, Physik, Chemie, Biologie und dem Sachunterricht bearbeiten. Sie beziehen in die Analyse des Zusammenwirkens von schulischen und außerschulischen Lernumwelten mit dem Unterricht das gesamte Schulspektrum mit ein: Ein Projekt gilt dem Lernen von Kindergartenkindern, acht Projekte sind in der Grundschule angesiedelt, 20 Projekte verfolgen Fragestellungen bei Schülerinnen und Schülern der Sekundarstufe I und vier bei solchen der Sekundarstufe II.

Um gemeinsame Schwerpunktsetzungen zu verdeutlichen, haben sich innerhalb des Programms sechs Arbeitsgruppen nach thematischer Nähe formiert, die auch die sechs Abschnitte dieses Beihefts bilden. Die unterschiedlich starken Besetzungen der einzelnen Abschnitte spiegeln Schwerpunktsetzungen in der deutschen Bildungsforschung wider. Zu Beginn jedes Abschnitts des vorliegenden Bandes finden die Leserin und der Leser eine kurze Einführung, in der die Gemeinsamkeiten der jeweiligen Beiträge kurz zusammenfassend präsentiert werden.

An dieser Stelle wird daher nur ein kurzer summarischer Überblick gegeben. Die vier Beiträge des ersten Abschnitts zur „Unterrichtsforschung in Mathematik“ beschäftigen sich aus unterschiedlichen Perspektiven mit dem Mathematikunterricht: Es geht u.a. um das Lehren von Stochastik (Wassner/Martignon/Sedlmeier), um mathematisches Begründen und Beweisen (Reiss/Hellmich/Thomas), um die Messung der Leistungsentwicklung in Mathematik anhand des Regensburger Mathematikleistungstests (vom Hofe/Pekrun/Kleine/Götz) und um die Einschätzung unterrichtsrelevanter Schülermerkmale durch Lehrpersonen (Hosenfeld/Helmke/Schrader). Die vier Beiträge des zweiten Abschnitts „Lehrerexpertise und Unterrichtsmuster in Mathematik und Physik“ haben gemeinsam, dass sie das Unterrichtsgeschehen, vor allem den Physikunterricht, ins Zentrum der Analyse rücken. Dies geschieht entweder durch Selbsteinschätzungen der Lehrpersonen zu ihrer Lehrexpertise (Diedrich/Thußbas/Klieme) oder per Videographie des unmittelbaren Unterrichtsgeschehens mit dem Ziel, typische Unterrichtsmuster aufzuzeichnen, deren Homogenität bzw. Heterogenität zu bestimmen und deren Potenzial für die Herstellung erfolgreicher Lehr-Lernumgebungen zu untersuchen (Fischer/Reyer/Wirz/Bos/Höllrich, Prenzel/Seidel/Lehrke/Rimmele/Duit, Euler/Geiser/Hoffmann/Müller/Widodo, Fischler/Zedler, Schröder/Tonhäuser). Die sechs Beiträge des dritten Abschnitts „Entwicklung und Evaluation von Unterrichtsmodulen und Trainingsprogrammen“ haben die Gemeinsamkeit, dass in den durchgeführten Trainings- oder Quasi-Experimenten solche abhängigen Variablen untersucht werden, die sich auf fächerübergreifende Kompetenzen beziehen. Dabei betreffen die drei Beiträge, in denen Unterrichtsmodule konzipiert und evaluiert werden, die Vermittlung eines basalen Wissenschaftsverständnisses (Sodian/Thoermer/Kircher/Grygier/Günther) und eines tieferen naturwissenschaftlichen Verständnisses (Möller/Jonen/Hardy/Stern) bereits in der Grundschule und die Vermittlung selbstorganisierten Lernens und selbstbestimmter Motivation in der Sekundarstufe I (Sumfleth/Wild/Rumann/Exeler). Die drei Beiträge,

in denen Trainingsprogramme für Schülerinnen und Schüler oder Lehrpersonen entwickelt und evaluiert werden, betreffen generelle Fähigkeiten des selbstregulierten Lernens (Gürtler/Perels/Schmitz/Bruder), spezifische Lernstrategien zum erfolgreichen Lernen aus Lehrtexten (Leopold/Leutner) und das Lehren mithilfe von Lösungsbeispielen (Renkl/Schworm). Die drei Beiträge des vierten Abschnitts „Diagnose und Förderung von Interessen und Lernmotivation“ untersuchen auf der Grundlage von Leistungsmotivationstheorien (Rheinberg/Wendland) sowie von Interessen- und Selbstbestimmungstheorien (Wild/Remy, Upmeyer zu Belzen/Vogt/Wieder/Christen) förderliche und hemmende Bedingungen für die Entwicklung von fachlichen Interessen und Lernmotivationen in den Lernumwelten der Schule und der Familie und sind längsschnittlich angelegt. Sie zielen insbesondere darauf ab, motivfördernde und -hemmende Erziehungs- und Instruktionsstrategien von Eltern und Lehrpersonen und deren Veränderung im Lauf der Schulzeit zu identifizieren. Die drei Beiträge des fünften Abschnitts „Einstellungen und Werte als förderliche oder hinderliche Bedingungen schulischer Leistungsfähigkeit“ untersuchen, inwieweit Bedingungen der sozialen Umwelt von Schülerinnen und Schülern zur Herausbildung von Einstellungen, Stereotypen, Werten oder Facetten des Selbstkonzepts beitragen, die die schulische Interessen- und Leistungsentwicklung fördern oder hemmen. U.a. wird die Bedeutung einer guten Schulleistung diskriminierenden Strebernorm in der Gleichaltrigengruppe untersucht (Pelkner/Günther/Boehnke), einer geschlechtsspezifischen Typisierung mathematisch-naturwissenschaftlicher Schulfächer (Hannover/Kessels) und die Bedeutung der Vermittlung von Wichtigkeitszuschreibungen zu einzelnen Schulfächern durch das Elternhaus (Streckler/Noack). Der sechste und letzte Abschnitt zur „Schulforschung“ besteht aus einem Beitrag (Ditton/Arnoldt/Bornemann), in dem ein standardisiertes Erhebungs- und Berichtssystem zur Qualitätskontrolle an Schulen vorgestellt wird. Dieses Berichtssystem setzt sich zusammen aus theoretisch abgeleiteten Indikatoren zur Schul- und Unterrichtsqualität. Es will den Schulen und Lehrpersonen auf der Grundlage einer Referenzstichprobe der jeweiligen Schulart Vergleichsdaten zur Verfügung stellen, die ihnen die Einordnung der eigenen schul- und unterrichtsspezifischen Ergebnisse erlauben.

Literatur

- Achtenhagen, F./Lempert, W. (1999): Entwurf zu einem Modellversuchsprogramm „Lebenslanges Lernen“. Bonn: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung.
- American Association for the Advancement of Science (AAAS) (1993): Benchmarks for science literacy. Project 2061. New York: Oxford University Press.
- Ardelt, C./Stanat, P./Schneider, W./Schiefele, U. (2001): Lesekompetenz: Testkonzeption und Ergebnisse. In: J. Baumert/E. Klieme/M. Neubrand/M. Prenzel/U. Schiefele/W. Schneider/P. Stanat/K.-J. Tillmann/M. Weiß (Hrsg.) PISA 2000. Opladen: Leske + Budrich, S. 69–140.
- Artelt, C./Demmrich, A./Baumert, J. (2001): Selbstreguliertes Lernen. In: J. Baumert/E. Klieme/M. Neubrand/M. Prenzel/U. Schiefele/W. Schneider/P. Stanat/K.-J. Tillmann/M. Weiß (Hrsg.) PISA 2000. Opladen: Leske + Budrich; S. 271–299.
- Aurin, K. (Hrsg.) (1990): Gute Schulen – worauf beruht ihre Wirksamkeit? Bad Heilbrunn: Reinhardt.

- Baumert, J. (1993): Lernstrategien, motivationale Orientierung und Selbstwirksamkeitsüberzeugungen im Kontext schulischen Lernens. In: *Unterrichtswissenschaft* 21, S. 327–354.
- Baumert, J./Bos, W./Watermann, R. (1998): TIMSS/III. Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J./Stanat, P./Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: J. Baumert/E. Klieme/M. Neubrand/M. Prenzel/U. Schiefele/W. Schneider/P. Stanat/K.-J. Tillmann/M. Weiß (Hrsg.) PISA 2000. Opladen: Leske + Budrich, S. 15–68.
- Baumert, J./Fend, H./O’Neil, H.F./Peschar, J.L. (1998): Prepared for life-long learning. Frame of reference for the measurement of self-regulated learning as a cross curricular competence (CCC) in the PISA project. Paris: OECD (unpublished manuscript).
- Baumert, J./Lehmann, R./Lehrke, M./u.a. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiß, M. (2001) (Hrsg.) PISA 2000. Opladen: Leske + Budrich.
- Beaton, A.E./Mullis, I.V.S./Martin, M.O./Gonzalez, E.J./Kelly, D.L./Smith, T.A. (1996a): Mathematics achievement in the middle school years: IEA’s third international mathematics and science study. Chestnut Hill, MA: Boston College.
- Beaton, A.E./Martin, M.O./Mullis, I.V.S./Gonzalez, E.J./Smith, T.A./Kelly, D.L. (1996b): Science achievement in the middle school years: IEA’s third international mathematics and science study. Chestnut Hill, MA: Boston College.
- Berger, P.L./Luckmann, T. (1995): Modernität, Pluralität, Sinnkrise. Gütersloh: Bertelsmann.
- Bronfenbrenner, U. (1978): Ansätze zu einer experimentellen Ökologie menschlicher Entwicklung. In: Oerter, R. (Hrsg.): *Entwicklung als lebenslanger Prozess*. Hamburg: Hoffmann & Campe, S. 33–65.
- Bronfenbrenner, U. (1981): *Die Ökologie der menschlichen Entwicklung*. Stuttgart: Klett-Cotta.
- Brown, A.L. (1997): Transforming schools into communities of thinking and learning about serious matters. In: *American Psychologist* 52, S. 399–413.
- Brown, A.L./Campione, C. (1990): Communities of learning and thinking, or, a context by any other name. In: *Contributions to Human Development* 21, S. 108–126.
- Carroll, J.B. (1963): A model of school learning. In: *Teachers College Record* 64, S. 723–733.
- Cobb, P. (1994): Where is the mind? Constructivist and sociocultural perspectives on mathematical development. In: *Educational Researcher* 23 (7), S. 13–20.
- Coleman, P. (1998): *Parent, Student and Teacher Collaboration. The Power of Three*. Thousands Oaks: Corwin.
- De Corte, E./Greer, B./Verschaffel, L. (1996): Mathematics teaching and learning. In: Berliner, D.C./Calfee, R.C. (Hrsg.): *Handbook of educational psychology*. New York: Macmillan, S. 491–547.
- Ditton, H. (1998). *Mehrebenenanalyse. Grundlage und Anwendungen des Hierarchisch Linearen Modells*. Weinheim: Juventa.
- Ditton, H. (2000): Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In: *Zeitschrift für Pädagogik* 41, Beiheft, S. 73–92.
- Duit, R./Häußler, P. (1997). Physik und andere naturwissenschaftliche Lernbereiche. In: F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie D/I/3. Psychologie des Unterrichts und der Schule* (S. 427–460). Göttingen: Hogrefe.
- Fend, H. (1996): *Schulkultur und Schulqualität*. In: *Zeitschrift für Pädagogik* 34, Beiheft, S. 85–97.
- Fend, H. (1997): *Der Umgang mit der Schule in der Adoleszenz*. Bern: Huber.
- Fend, H. (1998): *Qualität im Bildungswesen*. Weinheim: Juventa.
- Fend, H. (2001): *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung*. Weinheim: Juventa.

- Fend, H./Stöckli, G. (1997): Der Einfluss des Bildungssystems auf die Humanentwicklung: Entwicklungspsychologie der Schulzeit. In: Weinert, F.E. (Hrsg.): Enzyklopädie der Psychologie D/I/3. Psychologie des Unterrichts und der Schule. Göttingen: Hogrefe, S. 1–35.
- Fraser, B./Tobin, K. (Hrsg.) *International Handbook of Science Education* (pp. 3–25). Dordrecht, Niederlande: Kluwer.
- Freudenthal, H. (1991): *Revisiting mathematics education*. Dordrecht, The Netherlands: Kluwer.
- Gräber, W./Bolte, Gräber, W. (1992): Untersuchung zum Schülerinteresse an Chemie und am Chemieunterricht. In: *Chemie in der Schule* 39, S. 270–273.
- Helmke, A./Hornstein, W./Terhart, E. (2000): Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule. In: *Zeitschrift für Pädagogik*, 41. Beiheft. Weinheim: Beltz.
- Helmke, A./Schrader, F.-W. (2001): Determinanten der Schulleistung. In: Rost, D.H. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim: PVU Beltz, S. 81–91.
- Hofer, M. (2000): Schule: Vom Lernort zur „intermediären“ Institution. In: *Unterrichtswissenschaft* 28, S. 10–15.
- Hoffmann, L./Häußler, P./Lehrke, M. (1998): *Die IPN-Interessenstudie Physik*. Kiel: IPN.
- Kerres, M. (2001). *Multimediale und telemediale Lernumgebungen*. München: Oldenbourg.
- Klafki, W. (1994): Schlüsselprobleme als inhaltlicher Kern internationaler Erziehung. In: Seibert N./Serve, H.J. (Hrsg.): *Bildung und Erziehung an der Schwelle zum dritten Jahrtausend*. München: PimS Verlag, S. 135–161.
- Klafki, W. (1995): Zum Problem der Inhalte des Lehrens und Lernens an der Schule aus der Sicht kritisch-konstruktiver Didaktik. In: *Zeitschrift für Pädagogik* 33, Beiheft, S. 91–102.
- Klieme, E./Neubrand, M./Lüdtke, O. (2001). *Mathematische Grundbildung: Testkonzeption und Ergebnisse*. In: J. Baumert/E. Klieme/M. Neubrand/M. Prenzel/U. Schiefele/W. Schneider/P. Stanat/K.-J. Tillmann/M. Weiß (Hrsg.): *PISA 2000*. Opladen: Leske + Budrich, S. 141–191.
- Kölller, O./Baumert, J./Neubrand, J. (2000): Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): *TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie. Mathematisch und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe*. Opladen: Leske + Budrich, S. 229–270.
- Krapp, A. (1998): Entwicklung und Förderung von Interessen im Unterricht. In: *Psychologie für Erziehung und Unterricht* 44, S. 185–201.
- Krapp, A. (2001). Interesse. In D.H. Rost (Hrsg.) *Handwörterbuch Pädagogische Psychologie*, S. 286–294. Weinheim: Beltz PVU.
- Krapp, A./Prenzel, M. (Hrsg.) (1992): *Interesse, Lernen, Leistung*. Münster: Aschendorff.
- Krumm, V. (1996): Über die Vernachlässigung der Eltern durch Lehrer und Erziehungswissenschaft – Plädoyer für eine veränderte Rolle der Lehrer bei der Erziehung der Kinder. In: *Zeitschrift für Pädagogik* 34, Beiheft, S. 119–140.
- Krumm, V. (1998): Elternhaus und Schule. In: Rost, D. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz, S. 81–86.
- Lambert, N. M./McCombs, B. L. (Hrsg.) (1997): *How students learn*. Washington, DC: American Psychological Association.
- Mac Iver, D. J./Reumann, D. A./Main, S. R. (1995): Social structuring of the school: studying what is, illuminating what could be. In: *Annual Review of Psychology* 46, S. 375–400.
- Mandl, H./Friedrich, H. F. (Hrsg.) (1992): *Lern- und Denkstrategien. Analyse und Intervention*. Göttingen: Hogrefe.
- Mandl, H./Prenzel, M. (1991): Designing powerful learning environments: A constructivist perspective. In: Lowyck, J./DePotter, P/Elen, J. (Hrsg.): *Instructional design: Implementation issues*. Proceedings of the I.B.M./K.U. Leuven Conference, Dec 17–19, 1991. La Hulpe: IBM, S. 69–90.
- Max-Planck-Institut für Bildungsforschung (1996): *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). 2. Bericht für die Schulen*. Berlin: Max-Planck-Institut für Bildungsforschung.

- Moser, H. (2000). Einführung in die Medienpädagogik. Aufwachsen im Medienzeitalter. Opladen: Leske + Budrich.
- National Council of Teachers of Mathematics (NCTM) (1989): Curriculum and evaluation standards for school mathematics. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM) (1991): Professional standards for teaching mathematics. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM) (1995): Assessment standards for school mathematics. Reston, VA: NCTM.
- National Research Council (NRC). (1995): National science education standards. Washington, DC: National Academy Press.
- Noack, P. (1990): Jugendentwicklung im Kontext. Zum aktiven Umgang mit sozialen Entwicklungsaufgaben in der Freizeit. München: PVU.
- OECD (1989): Schools and Quality. An International Report. Paris: OECD.
- OECD (1995): Measuring the quality of schools. Paris: OECD.
- Oerter, R. (1998): Kultur, Ökologie und Entwicklung. In: Oerter, R./Montada, L.: Entwicklungspsychologie. Ein Lehrbuch. Weinheim: Beltz PVU, S. 84–121.
- Pekrun, R. (1997): Kooperation zwischen Elternhaus und Schule. In: Vascovics, L.A./Lipinski, H. (Hrsg.): Familiäre Lebenswelten und Bildungsarbeit (Bd. 2). Opladen: Leske & Budrich, S. 51–79.
- Pekrun, R. (1998): Schüleremotionen und ihre Förderung: Ein blinder Fleck in der Unterrichtsforschung. In: Psychologie in Erziehung und Unterricht 44, S. 230–248.
- Pekrun, R./Fend, H. (Hrsg.) (1991): Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung. Stuttgart: Enke.
- Pekrun, R./Schiefele, U. (1996): Emotions- und motivationspsychologische Bedingungen von Lernleistung. In: Weinert, F.E. (Hrsg.): Enzyklopädie der Psychologie D/II/2. Psychologie des Lernens und der Instruktion. Göttingen: Hogrefe, S. 153–180.
- Polya, G. (1965): Mathematical discovery (Vol. 2). New York: Wiley.
- Prenzel, M./Merkens, H./Noack, P. u.a. (1999): Die Bildungsqualität von Schule: Fachliches und fächerübergreifendes Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten. Antrag an den Senat der DFG zur Einrichtung eines Schwerpunktprogramms.
- Prenzel, M./Rost, J./Senkbeil, M./Häußler, P./Klopp, A. (2001): Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In: J. Baumert/E. Klieme/M. Neubrand/M. Prenzel/U. Schiefele/W. Schneider/P. Stanat/K.-J. Tillmann/M. Weiß (Hrsg.) PISA 2000. Opladen: Leske + Budrich, S. 192–250.
- Riquarts, K./Wadewitz, C. (1999): Framework for Science Education in Germany. Kiel: IPN.
- Saldern, M.v. (2001). Mehrebenenanalyse. In D.H. Rost (Hrsg.) Handwörterbuch Pädagogische Psychologie. Weinheim: Beltz PVU., S. 457–461.
- Scheerens, J. (1990): School effectiveness and the development of process indicators of school functioning. In: School Effectiveness and School Improvement 2, S. 61–80.
- Scheerens/Bosker (1997): The foundations of educational effectiveness. Oxford: Pergamon Press.
- Schiefele, U./Wild, K.-P. (Hrsg.) (2000). Interesse und Lernmotivation. Münster: Waxmann.
- Schmidt, H.W./Cogan, L.S. (1996): Development of the TIMSS context questionnaires. In: Martin, M.O./Kelly, D.L. (Hrsg.). Third International Mathematics and Science Study (TIMSS). Technical Report. Volume I. Chestnut Hill, M.A.: Boston College.
- Schoenfeld, A.H. (1992): Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In: Grouws, D.A. (Hrsg.): Handbook of research on mathematics, teaching and learning (NCTM). New York: Macmillan, S. 334–370.
- Schommer, M./Calvert, C./Gariglietti, G./Bajaj, A. (1997): The development of epistemological beliefs among secondary students: A longitudinal study. In: Journal of Educational Psychology 89, S. 37–40.

- Schümer, G. (1998): Mathematikunterricht in Japan – Ein Überblick über den Unterricht in öffentlichen Grund- und Mittelschulen und privaten Ergänzungsschulen. In: Unterrichtswissenschaft 26, S. 195–228.
- Sherwood, R.D./Petrosino, A.J./Lin, X./Cognition and Technology Group at Vanderbilt (1998): Problem-based macro contexts in science instruction: Design issues and applications. In: Fraser, B./Tobin, K.: International handbook of science education. Dordrecht, Niederlande: Kluwer, S. 349–362.
- Steffen, U./Bargel, T. (1993): Erkundungen zur Qualität von Schule. Luchterhand: Neuwied.
- Terhart, E. (1987): Kommunikation im Kollegium. In: Die deutsche Schule 79, S. 440–450.
- Terhart, E. (2000): Qualität und Qualitätssicherung im Schulsystem. Hintergründe – Konzepte – Probleme. In: Zeitschrift für Pädagogik 46, S. 809–830.
- Tillmann, K.-J. (Hrsg.) (1989): Was ist eine gute Schule?. Hamburg: Bergmann & Helbig.
- Törner, G./Grigutsch, S. (1994): „Mathematische Weltbilder bei Studienanfängern“ – Eine Erhebung. In: Journal für Mathematik-Didaktik, 15(374), S. 211–251.
- Treagust, D./Duit, R./Fraser, B. (Hrsg.). (1996): Improving Teaching and Learning in Science and Mathematics. New York: Teacher College Press.
- Walberg, H.J. (1990): A theory of educational productivity: Fundamental substance and method. In: Vedder, P. (Hrsg.): Fundamental studies in educational research. Lisse: Swets & Zeitlinger S. 19–34.
- Weinert, F. E. (1998): Neue Unterrichtskonzepte zwischen gesellschaftlichen Notwendigkeiten, pädagogischen Visionen und psychologischen Möglichkeiten. In: Bayrisches Staatsministerium für Unterricht, Kultus, Wissenschaft und Kunst: Wissen und Werte für die Welt von morgen. München: Bayerisches Staatsministerium für Unterricht, Kultus, Wissenschaft und Kunst, S. 101–125.
- Weinert, F.E./Helmke, A. (Hrsg.) (1997): Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union.
- Wild, K.-P./Krapp, A. (1996): Die Qualität subjektiven Erlebens in schulischen und betrieblichen Lernumwelten. Untersuchungen mit der Erlebens-Stichproben-Methode. In: Unterrichtswissenschaft 24, S. 195–216.
- Winterhoff-Spurk, P. (1999): Medienpsychologie. Eine Einführung. Stuttgart: Kohlhammer.
- Youniss, J./Hyane, D.L. (1992): Friendship in adolescence. In: Development and behavioral pediatrics 13(1), S. 59–66.

Anschrift der Autoren:

PD Dr. Jörg Doll, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.

Prof. Dr. Manfred Prenzel, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.

Teil I:

Unterrichtsforschung in Mathematik

Förderung des mathematischen Verständnisses,
Problemlösens und der Herausbildung
zutreffender mathematischer Weltbilder von
Schülerinnen und Schülern

Kristina Reiss

Einleitung

Die Beiträge dieses Abschnitts beschäftigen sich aus ganz unterschiedlicher Perspektive mit dem Mathematikunterricht. Dabei ist ihnen der Bezug auf die großen internationalen Vergleichsstudien TIMMS und PISA gemeinsam, die erhebliche Defizite des Mathematikunterrichts in Deutschland gezeigt haben (Baumert u.a. 1997; Deutsches PISA-Konsortium 2001). Danach liegen die Stärken deutscher Schülerinnen und Schüler eher in der Anwendung von algorithmenorientierten Verfahren als in der eigenständigen Modellierung und Lösung mathematischer Probleme. In Ländern wie Japan, Korea, Neuseeland oder auch Finnland liegen die Gesamtleistungen um ungefähr eine halbe Standardabweichung über denen der Bundesrepublik und damit in einem Kompetenzbereich, der vor allem durch Modellieren und begriffliches Verknüpfen charakterisiert werden kann (Klieme/Neubrand/Lüdtke 2001). Doch können diese Studien aufgrund ihrer Anlage nur in sehr begrenztem Maße auch Ursachen für diese Schulleistungsunterschiede oder Wirkungszusammenhänge zwischen einzelnen Ursachenkomponenten beschreiben. Hier setzen die Arbeiten in den vier Projekten an, deren erste Ergebnisse dargestellt werden.

Der Aufsatz von Wassner, Martignon und Sedlmeier beschäftigt sich mit Möglichkeiten, wie durch die Verwendung von Baumdiagrammen die Fähigkeit zur Lösung komplexer stochastischer Aufgaben trainiert werden kann. Dabei werden drei Bedingungen als bedeutsam angesehen, nämlich die Daten- und Repräsentationsorientierung, die Unterstützung aktiven Lernhandelns der Schülerinnen und Schüler und ein hoher Alltagsbezug. Insbesondere sollen fächerübergreifende Kompetenzen beim Umgang mit unsicheren Informationen gefördert werden. In einem quasiexperimentellen Design wurde eine Trainingsgruppe mit einer Kontrollgruppe verglichen. Wie erwartet zeigte die Trainingsgruppe bessere Leistungen, es bestätigte sich damit die Überlegenheit des Repräsentationsmodells mit Häufigkeitsformaten beim Lehren von Bayes'scher Inferenz und bedingter Wahrscheinlichkeit über andere Repräsentationsformate. In einem kleinen Bereich konnte damit gezeigt werden, dass in Bezug auf die dahinter liegende Modellierung geeignete Unterrichtsansätze zur Leistungssteigerung beitragen können.

Reiss, Hellmich und Thomas beschreiben eine Untersuchung zum Beweisen und Begründen mit Schülerinnen und Schülern der Klassen 7/8 des Gymnasiums. Im Vordergrund steht dabei die Frage, welches Beweisverständnis und welche argumentativen Fähigkeiten die Schüler haben und welche Bedingungsvariablen damit verbunden sind. In der Studie wurden dazu Leistungsvariablen wie geometrische Basiskompetenzen, das Beweisverständnis und das wissenschaftliche Grundverständnis, aber auch die mathematikbezogenen Beliefs von Lehrern und Schülern sowie Angaben der Lehrerinnen und

Lehrer zum präferierten Unterrichtsstil erfasst. Die Ergebnisse des Leitungstests lassen sich in den von TIMSS und PISA aufgezeigten Kontext integrieren und bestätigen das Stufenmodell mathematischer Kompetenz für den speziellen Bereich des Argumentierens und Begründens. Darüber hinaus repliziert die Untersuchung nicht nur Ergebnisse zum Umgang mit Beweisen, sondern weist die entsprechenden Kompetenzen und Defizite bereits für einen relativ frühen Zeitpunkt in Bezug auf das Beweisen im Mathematikunterricht nach. Dies gilt auch für das wissenschaftlichen Denkens der Schülerinnen und Schüler, bei dem typische Schwierigkeiten identifiziert werden konnten.

Gegenstand der Untersuchung von Hosenfeld, Helmke und Schrader sind zentrale Determinanten der schulischen Leistung. Es werden insbesondere Zusammenhänge zwischen individuellen Bedingungsmerkmalen, der Unterrichtsqualität und dem Kontext im Hinblick auf Verständnis und Lernmotivation betrachtet. Dabei geht es den Autoren einerseits um mathematisches Verständnis, andererseits ist die Lernmotivation, die durch Merkmale wie Interesse, Fähigkeitsselbstkonzept, Wertschätzung des Fachs, Einstellung zum Lernen und Lernstrategien charakterisiert wird, ein wesentlicher Aspekt. Die Darstellung beschränkt sich auf einige ausgewählte Aspekte der Untersuchung. So stehen die Frage nach dem Vorwissen (sowohl als individuelle Variable als auch in Bezug auf Unterschiede zwischen Klassen), die Freude am Fach, die Selbsteinschätzung der Schülerinnen und Schüler sowie die Einschätzung dieser Variablen durch die Lehrkräfte im Vordergrund. Die Ergebnisse zeigen Klassenunterschiede, Unterschiede zwischen Schulformen, aber auch Überlappungsbereiche etwa zwischen guten Realschulklassen und schwächeren Gymnasialklassen. Die Einschätzungen der Lehrer in Bezug auf die Lösungen zeigen fast systematisch eine deutliche Überschätzung der Schülerleistungen, allerdings korrelieren Lehrereinschätzung und tatsächliche Leistung positiv. Zwischen Freude an der Mathematik und der Vortestleistung gibt es eine positive Korrelation, die in den Hauptschulklassen am deutlichsten ist.

Vom Hofe, Pekrun, Kleine und Götz verfolgen in ihrem Projekt das Ziel, Entwicklungsverläufe mathematischer Leistung, die entsprechenden Schülervoraussetzungen und selbstreguliertes Lernen sowie Kontextbedingungen mathematischer Leistung in Unterricht, Schulklasse und Elterhaus systematisch zu untersuchen. Dies geschieht im Rahmen einer Längsschnittstudie, in die Schülerinnen und Schüler der Klassen 5 bis 10 von Hauptschulen, Realschulen und Gymnasien einbezogen werden. Im Beitrag wird zunächst die Konstruktion eines Mathematikleistungstests für 5. bis 10. Klassen dargestellt. Allgemeines Ziel der Testkonstruktion ist es, der Frage nachzugehen, inwieweit die Entwicklung einer mathematischen Grundbildung gleichmäßig verläuft oder inwieweit sich Phasen, Stufen oder Sprünge identifizieren lassen. Dazu ist es erforderlich, einen Test zu entwickeln, der die Leistungen eindimensional skaliert. Es wird über die Ergebnisse dieser Rasch-Skalierung berichtet. Erste Befunde zur Entwicklung mathematischer Kompetenzen, die anhand des entwickelten Verfahrens gewonnen wurden, schließen sich an.

Alle in diesem Kapitel zusammengefassten Untersuchungen machen zunächst deutlich, dass es sinnvoll ist, die Leistungen von Schülern zu erfassen und sie miteinander zu vergleichen. Die Qualität einer Schule muss sich daran messen lassen, ob sie Schüler zu

den geforderten Leistungen hinführen kann. Auch wenn durch Untersuchungen wie TIMSS und PISA diese Leistungen in einer großen Stichprobe betrachtet und erste Auswertungen im Sinne von Kontextvariablen ermöglicht wurden, sind beide im Wesentlichen groß angelegte Querschnittsuntersuchungen, die produktorientiert Leistungen zu einem bestimmten Zeitpunkt erheben und Kontextvariablen nur im Ansatz erfassen können. Die in diesem Kapitel dargestellten Studien sind als ein Schritt in Richtung eines tiefergehenden Verständnisses mathematischen Lernens und Lehrens zu verstehen.

Literatur

- Baumert, J./Lehmann, R./u.a. (Hrsg.) (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske+Budrich.
- Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske+Budrich.
- Klieme, E./Neubrand, M./Lüdtke, O. (2001): Mathematische Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske+Budrich, S. 139–190.

Anschrift der Autorin:

Prof. Dr. Kristina Reiss, Universität Augsburg, Lehrstuhl Didaktik der Mathematik, 85135 Augsburg.

Christoph Wassner/Laura Martignon/Peter Sedlmeier

Die Bedeutung der Darbietungsform für das alltagsorientierte Lehren von Stochastik

1. Einleitung

„Statistisches Denken wird eines Tages genauso wichtig sein für eine aufgeklärte Gesellschaft wie die Fähigkeit, zu lesen und zu schreiben“. Diese Worte des Romanautors H.G. Wells („Die Zeitmaschine“) wiesen damals bereits auf einen absehbaren Wandel unserer Gesellschaft zu einer Informationsgesellschaft und der damit verbundenen Verschiebung von Fähigkeitsanforderungen an den heranwachsenden Menschen hin. Auch der Mathematiker Hans Freudenthal empfahl für das Lehren und Lernen von Mathematik eine starke „Orientierung an der Welt“ mit dem Ziel der Ausbildung tragfähiger mentaler Modelle für mathematische Begriffe (Freudenthal 1983). In diesem Sinne ist zu diskutieren, ob die steigende Bedeutung von statistischer Information in unserer Umwelt und deren wachsende Kommunikation auch eine Neuorientierung von Zielen, Maßnahmen und Methoden der mathematischen Grundbildung erfordert. Das Ziel des hier dargestellten Projektes¹ ist die Untersuchung von Bedingungen und Verbesserungsmöglichkeiten einer entsprechenden fächerübergreifenden Kompetenz von Schülern, mit unsicheren Informationen und statistischen Daten umzugehen und Entscheidungen, die auf Daten basieren, treffen und begründen zu können. In der Literatur finden sich Bezeichnungen wie „Datenkompetenz“ oder „Entscheiden unter Unsicherheit und Risiko“. Erforderliche Fähigkeiten erstrecken sich vom richtigen und verantwortungsvollen Umgang mit statistischen Daten, dem Wissen um grundlegende Methoden zur Interpretation und Analyse bis hin zur Fähigkeit, Werkzeuge für sinnvolle Schlussfolgerungen und Entscheidungen unter Unsicherheit parat zu haben und benutzen zu können (vgl. NCTM 2000, S. 47). In unserer vom Austausch und der Analyse von Informationen geprägten Welt erfahren mathematische Teilgebiete wie Datenanalyse, Wahrscheinlichkeitsrechnung und Statistik eine neue bildungspolitische Relevanz. Alltagsorientierte und problembasierte stochastische Bildung ist eine wichtige Voraussetzung, die die Schule schaffen muss, damit Schulabgänger als kritische und politisch ihres Handelns bewusste Erwachsene in Staat, Wirtschaft und Gesellschaft agieren können. Neuere Entwürfe zur stochastischen Bildung favorisieren deshalb Konzepte, die „stochastisches Denken“ an Stelle bloßen Erlernens von Rechenverfahren fördern. Entscheidend ist, nicht in „methodischen Ritualen“ (Gigerenzer 1999), halb verstandenen Regeln und eingepaukten Schemata verhaftet zu bleiben, sondern Stochastik möglichst realitätsbezogen und problemorientiert zu unterrichten.

1 Das Projekt wird seit 1.1.2001 von der DFG gefördert (Geschäftszeichen: Ma 1544 /10-1).

Wir fassen im Folgenden einige Entwicklungen im Bereich der didaktischen und psychologischen Forschung zur Wahrscheinlichkeitsrechnung und Statistik und zu Möglichkeiten von computergestütztem Training zur Förderung von Lernaktivität zusammen, die als theoretischer Hintergrund der speziellen Forschungsfragen des Projektes anzusehen sind.

2. Didaktische und kognitionspsychologische Positionen und Befunde zur stochastischen Bildung

Neben einer Vielzahl von grundlegenden didaktischen Ansätzen zur Stochastik (z.B. Heitele 1976; Riemer 1985; Wickmann 1990; Borovcnik 1992; Garfield 1993) kamen von didaktischer Seite, insbesondere aus den USA, aber auch aus der kognitionspsychologischen Forschung neue Ideen, die interessant im Hinblick auf die Förderung flexiblen und realitätsbezogenen stochastischen Denkens sind. Eine zentrale Bedeutung für die derzeitige „Mathematical literacy“-Diskussion haben die „Principles and Standards for School Mathematics“ des „National Council of Teachers of Mathematics“ (NCTM 2000) aus den USA, die Empfehlungen für einen modernen mathematischen Unterricht geben. Es finden sich nicht nur Aussagen zu Inhalten, sondern auch zu Prozesselementen, wie z.B. Repräsentation und Kommunikation. Für den Themenbereich „probability & chance“ sind einige Forderungen ein hoher Anteil an Eigenaktivität der Lernenden, Akzente auf der Anwendung der Stochastik im realen Leben und auf reale Fragestellungen und Betonung von „data handling“ (authentische Datenanalyse) im allgemeinen Mathematik-Curriculum² als fächerübergreifendes Element (Biehler 1997, 2001; Engel 2001). Demgemäß soll die Ausbildung den Lernenden umfassende Möglichkeiten bieten, Daten zu erheben und aufzubereiten, grafische Darstellungen zu verwenden und schließlich Schlussfolgerungen und Entscheidungen basierend auf den analysierten Daten zu treffen.

Solche Schlussfolgerungen bereiten nicht nur Schülern erhebliche Probleme. Auch „Experten“ wie z.B. Ärzten, Gesundheitsberatern oder Richtern fallen Entscheidungen unter Unsicherheit schwer (z.B. Gigerenzer im Druck; Krauss/Hertwig 2000). Die Behauptung, dass der Mensch hierbei resistenten kognitiven Täuschungen unterliegt (z.B. Kahneman/Slovic/Tversky 1982; Piattelli-Palmarini 1994), wurde durch neuere Forschungsarbeiten im Bereich der Kognitionspsychologie widerlegt (z.B. Gigerenzer/Hoffrage 1995; Cosmides/Tooby 1996). Der entscheidende Schlüssel zum Verstehen von Wahrscheinlichkeitsproblemen scheint die Verwendung der richtigen Darbietungsform zu sein. Im Bereich der bayesschen Schlussfolgerungen³ z.B. zeigten diverse Studien (Gigerenzer/Hoffrage 1995; Sedlmeier/Gigerenzer 2001), dass insbesondere die Frage des numerischen Formates für den Verstehenserfolg entscheidend ist. Spezielle grafische

- 2 Das ist in Deutschland nur in NRW der Fall, sonst sind Elemente der beschreibenden Statistik nicht expliziter Bestandteil der Curricula.
- 3 Schlussfolgerung durch Anwendung des Bayes-Theorems.

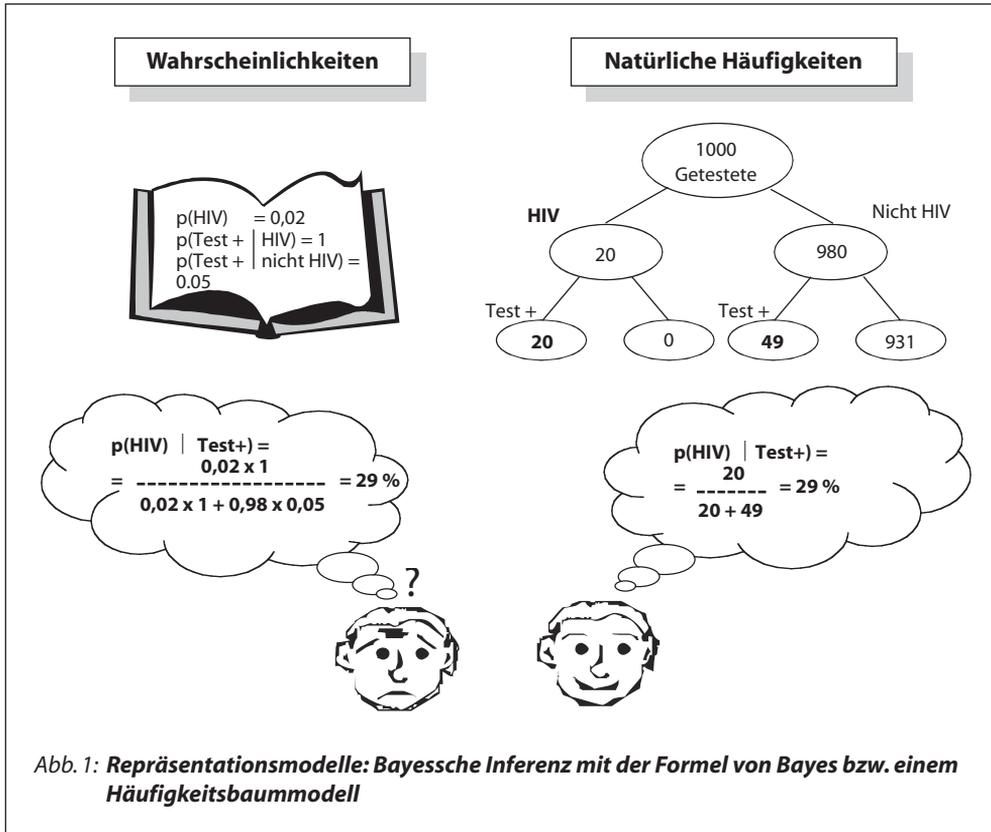
Darstellungen statistischer Information unter Verwendung von „natürlichen“, absoluten Häufigkeiten können die Performanz bei bayesschen Schlussfolgerungen erheblich verbessern. Wie kongruent Intuition und Wahrscheinlichkeitstheorie dabei sein können, zeigen Beispiele in Martignon/Wassner (2001). Der für diese Inferenz zentrale „Satz von Bayes“ ist wohl nicht an sich unintuitiv, sondern er ist nur in entsprechender Form ein Stolperstein für die Intuition (Wassner/Krauss/Martignon 2002). In weiteren Projekten des Schwerpunktprogramms wurde ebenfalls deutlich (z.B. Münster/Berlin), dass flexibles Wissen über Repräsentation von hohem Nutzen sein kann, wenn wir für Lernende eine Brücke zwischen Mathematik und Anwendung im alltäglichen Leben schlagen wollen (Stern/Aprea/Ebner im Druck).

Ausgehend von der These, dass Darbietungsformen, die im Unterricht verwendet werden, nicht optimal an vorhandene probabilistische Intuitionen der Schüler angepasst sein können (Martignon 2000), ist eine offene Forschungsfrage, inwieweit solche Schülerintuitionen durch geeignete Darbietungsformen nutzbar gemacht werden können. Erkenntnisse der kognitiven Psychologie im Bereich des stochastischen Denkens können wesentliche Anhaltspunkte liefern, wie gezielt auf vorhandene Primärintuitionen aufgebaut werden kann. Empirische Untersuchungen wurden zum kurzfristigen Erfolg verschiedener Trainingsvarianten zu bedingten Wahrscheinlichkeiten durchgeführt, denen besondere Bedeutung im Bereich stochastischen Denkens beigemessen wird. Die Ergebnisse bezogen sich zunächst auf Unterschiede im Lernerfolg durch verschiedene grafische Modelle (z.B. Bea 1995).

Sedlmeier und Gigerenzer entwickelten mehrere computergestützte Trainingsprogramme, deren Grundkonzept auf der Verwendung von Häufigkeitsformaten (z.B. „20 von 1000“) anstatt Wahrscheinlichkeitsformaten (z.B. „0,02“ oder „2%“) beruhte (z.B. in Verbindung mit einem grafischen Baummodell oder einem Flächenrastermodell, Abbildung 1, S. 38).

Trainings wurden mit Problemen zur Konjunktion von Wahrscheinlichkeiten, zu bedingten Wahrscheinlichkeiten (Sedlmeier 2000a, b) und zu bayesscher Inferenz (Sedlmeier 1997; Sedlmeier/Gigerenzer 2001) entwickelt. Es ergaben sich deutliche Unterschiede im Trainingserfolg im Vergleich zu analogen Trainings mit Wahrscheinlichkeitsformaten.

Wie kamen diese Unterschiede zustande? Es spricht einiges dafür, dass Intuitionen im Bereich des Umgangs mit Zufall und Unsicherheit als Resultat phylogenetischer Selektionsprozesse und ontogenetischer Lernprozesse in Interaktion mit der natürlichen Umgebung entstanden sind (Cosmides 1989; Cummins 1998; Sedlmeier/Wettler 1998). Dieser Auffassung folgend können wir Probleme umso leichter verarbeiten, je näher ihr Darbietungsformat dem Format ist, in dem Ereignisse und Objekte in der natürlichen Umgebung wahrgenommen werden können. Das natürliche Format ist allerdings wegen der Langsamkeit von phylogenetischen Selektionsprozessen nicht unbedingt das, welches wir gegenwärtig verwenden, sondern ein Format, mit dem unsere frühen Vorfahren zu tun hatten: Reale Häufigkeiten und nicht, wie heutzutage nach Entwicklung von Schrift und Mathematik, Prozentwerte oder Dezimalzahlen. Diese Überlegungen führten zum Konzept der „natürlichen Häufigkeiten“ und zur Annahme, dass der



Mensch sehr wohl über „adaptive Algorithmen“ verfügt, die seiner Performanz im Umgang mit Wahrscheinlichkeitsproblemen förderlich sind (Gigerenzer 1993).

Ein weiteres, von Sedlmeier (1998; 1999) entwickeltes Trainingsprogramm benutzte ein animiertes Urnenmodell, um den Einfluss der Stichprobengröße auf die Reliabilität von Schätzungen zu lehren. Eine spezielle dynamische Häufigkeitsdarstellung förderte dabei die schon bei Kindern vorhandene Intuition, dass größere Stichproben zu genaueren Schätzungen führen (Piaget/Inhelder 1951/1975) und half beim Verstehen von Stichprobenverteilungen. Zudem untersuchten wir, welchen Einfluss die Flexibilität der Darstellung und damit der Grad der Eigenaktivität des Lernenden auf den Lernerfolg hat (s.u.). Implikationen aus diesen Studien für die Gestaltung von computergestützten Statistiktutoren werden in Sedlmeier/Wettler (1998) diskutiert.

3. Entwurf einer Lehr-Lernumgebung zur Förderung stochastischer Kompetenzen

Die Ergebnisse aus der Forschung zu Urteilen unter Unsicherheit und Training solcher Urteilsprozesse dienen als Grundlage für einen Entwurf einer schulischen Lehr-Lern-

umgebung zur stochastischen Grundbildung. Das bedeutsamste Ziel ist die Förderung fächerübergreifender Kompetenzen im Sinne des Umgangs mit unsicheren Informationen und Daten. Formelwissen und innermathematische Fähigkeiten treten deshalb in diesem Ansatz zunächst in den Hintergrund. Ein vorläufige Konkretisierung des Ansatzes erfolgte in Form von Lehr-Lernmaterialien zur Wahrscheinlichkeitsrechnung mit begleitender, computergestützter Trainingssoftware (Sedlmeier/Köhlers 2001).

Wir halten drei Bedingungen ausschlaggebend für den Erfolg einer entsprechenden Lehr-Lernumgebung (Überblick über theoretischen Hintergründe und Inhalte in Sedlmeier 2001). Zentral ist die bereits dargestellte Rolle des Darbietungsformates, d.h. die besondere Bedeutung sowohl statischer (z.B. grafische Darstellungen mit Häufigkeitsformaten) als auch dynamischer Repräsentation (z.B. Simulation von Zufallsprozessen). Eine weitere entscheidende Bedingung ist die Unterstützung der Eigenaktivität der Lernenden – sie sollen durch selbstentworfenen, computergestützte Experimente Gesetzmäßigkeiten selbst erfahren können („learning by doing“). Die lernfördernde Wirkung von Eigenaktivität hat bereits Eingang in kognitionspsychologische Theorien gefunden, die als Grundlage für die Konstruktion von intelligenten Tutorssystemen benutzt werden (z.B. Anderson u.a. 1995). Demnach muss der Lernende aktiv in die Problemlösung einbezogen werden, um neues prozedurales Wissen erwerben zu können. Die dritte wichtige Bedingung ist eine Einbettung in alltagsrelevante bzw. authentische Probleme. Arbeiten im Bereich „situated cognition“ (z.B. Brown/Collins/Duguid 1989) betonen die Wichtigkeit authentischer Kontexte beim Lehren und Lernen, da konzeptuelles Wissen nicht so leicht abstrahiert wird. In dieser Sichtweise ist konzeptuelles Wissen eine Menge von Werkzeugen, die in realen Situationen tatsächlich benutzt werden können. Die Bedingungen und Möglichkeiten der Anwendung ergeben sich erst aus dem Handlungskontext. Durch flexibel gestaltbare Visualisierungen und Simulationen wird ermöglicht, dass die Lernenden eigene Experimente in authentischen Kontexten durchführen können.

4. Fragestellungen und Ziele

Aus dem bisherigen Stand der Forschung lässt sich mit gutem Grund annehmen, dass Schüler eine höhere stochastische Kompetenz erreichen, wenn sie sich in Lehr-Lernumgebungen befinden, in denen diese Bedingungen erfüllt sind. Diese Rahmenhypothese wurde ab Januar 2001 systematisch durch außerschulische Trainingsstudien und erste Pilotimplementationen in schulischen Lehr-Lernumgebungen untersucht. Wir geben im Folgenden eine Zusammenfassung von Fragestellungen, Untersuchungsmethoden und ersten Befunden des Forschungsprojekts.

Die Fragestellungen betrafen zunächst eine systematische und individuelle Überprüfung der genannten Bedingungen im Hinblick darauf, ob ihr Vorhandensein grundlegende Kompetenzen im Umgang mit Unsicherheit verbessern kann. Die Implikationen für die Gestaltung der Lehr-Lernumgebung sind:

- 1) Verbesserung der Instrukionsqualität durch intuitive (an individuelle Verarbeitungsprozesse der Schüler angepasste) Darbietungsformen,
- 2) Verstärkte Möglichkeiten der eigenen Lernhandlung von Schülern durch interaktive und flexible, computergestützte Trainingsprogramme,
- 3) Betonung alltagsorientierter und problembasierter Unterrichtsmuster.

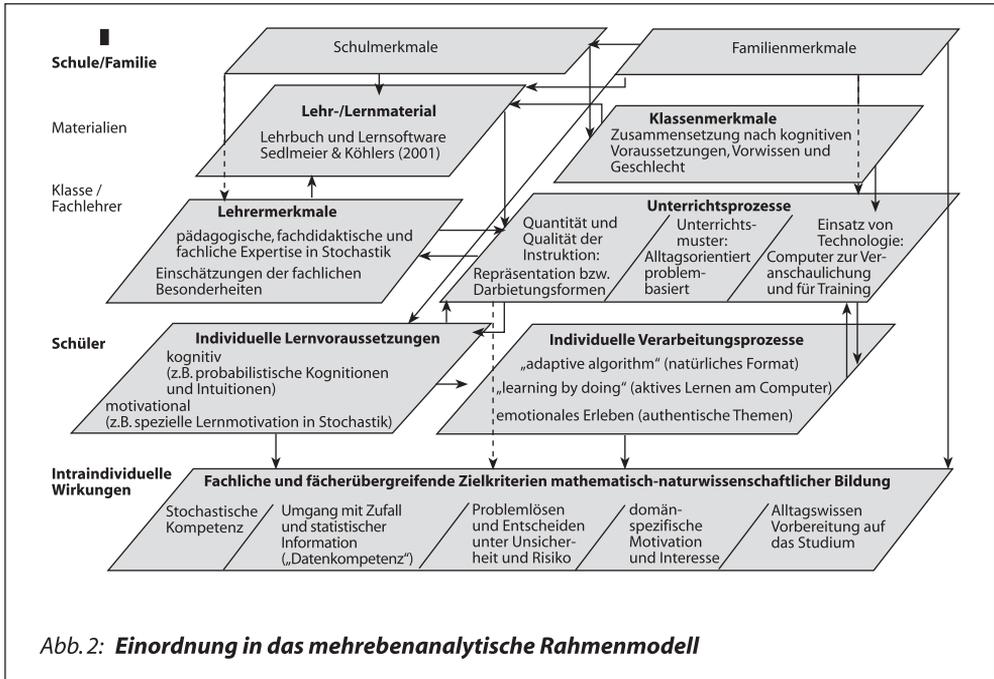


Abbildung 2 zeigt die Einordnung zu untersuchender Bedingungen in ein mehrbenenanalytisches Rahmenmodell des Wirkungsgefüges von Lernvoraussetzungen auf fachliche und fächerübergreifende Zielkriterien von stochastischer Bildung. Der Untersuchungsbereich beginnt in der 2. Ebene (unterhalb von Schule/Familie), wobei zunächst Lehr-Lernmaterialien für die Stochastik und individuelle Verarbeitungsprozesse in ihrer Interaktion untersucht werden sollen. Die neuen Materialien wurden, wie bereits dargestellt, konsequent nach kognitionstheoretischen bzw. empirisch validierten Konzepten entwickelt (Sedlmeier/Köhlers 2001). Auf Schülerebene soll untersucht werden, wie bestimmte Merkmale auf individuelle Verarbeitungsprozesse der Schüler wirken. Die Gestaltung und Auswahl von Lehrmaterialien betrifft in natürlicher Weise den Prozess der Unterrichtsplanung und -vorbereitung und hat nicht nur inhaltliche, sondern natürlich auch Auswirkungen auf Unterrichtsprozesse. Die natürliche Interaktion zwischen Lehrmaterial, Unterricht und individueller Schülerebene muss folgerichtig um die Lehrerdimension ergänzt werden. Lehrervorstellungen üben erheblichen Einfluss auf Lehr-Lernprozesse aus und sind deshalb ein wichtiger Faktor, wenn Bedingungen des Lernens untersucht werden (z.B. Fischler 2001). Eine grundlegende fachliche Besonderheit der

Stochastik ist die Mittlerrolle zwischen der Exaktheit der Mathematik und der Beliebigkeit einer vom Zufall geprägten Welt. So ist etwa eine deutliche Wahr-Falsch-Kategorisierung gewohnter mathematischer Aufgabenstellungen bei stochastischen Aussagen oft nicht möglich. Kompetenz und Ansichten der Lehrer über fachliche Besonderheiten bilden die Basis für die Instruktionsqualität im Unterricht.

Konsistent zur theoretischen Einbettung ergeben sich konkrete Forschungsfragen im Hinblick auf individuelle Verarbeitungs- und Unterrichtsprozesse:

- Welchen Einfluss haben verschiedene Repräsentationsmodelle auf den Lern- und Verstehenserfolg bei grundlegenden Wahrscheinlichkeitsproblemen (z.B. im Bereich bedingte Wahrscheinlichkeit, bayessche Inferenz, statistische Interpretation von Wahrscheinlichkeit, statistische Verteilungen)?
- Welchen Nutzen haben verschiedene Repräsentationen für Transfer- und Formalisierungskompetenzen?
- Hat die Einkleidung in authentische Kontexte Auswirkungen auf Lernmotivation und Verstehen von Wahrscheinlichkeitsproblemen?
- Inwieweit kann der Einsatz von interaktiver Simulations- bzw. Trainingssoftware beim Verstehen und der Durchdringung der Problemlösungen helfen? Welche besondere Rolle spielt dabei die Betonung der Eigenaktivität der Lernenden?

5. Trainingstudien zu individuellen Verarbeitungsprozessen bei unterschiedlichen Repräsentationsmodellen

Untersuchungsmethoden im Projekt betrafen zunächst Trainingsstudien zur systematischen Überprüfung individueller Verarbeitungsprozesse der Schüler im Labor. Wir beschränken uns in diesem „Werkstattbericht“ auf die Darstellung einer repräsentativen Trainingsstudie.

5.1 Untersuchungsgegenstand

In der Trainingsstudie wurde der Einfluss unterschiedlicher Repräsentationsmodelle auf das Verständnis von bedingter Wahrscheinlichkeit bzw. bayesscher Inferenz untersucht. Die Trainingsvarianten unterschieden sich hinsichtlich der verwendeten Repräsentationsformate. In einem Formeltraining (FORM) übten die Schüler, zur Lösung von Wahrscheinlichkeitsaufgaben passende Wahrscheinlichkeitsinformationen in verschiedene Formeln für bedingte Wahrscheinlichkeiten einzusetzen, und in einem Häufigkeitstraining (BAUM) übten sie, diese Aufgaben durch Einsetzen von Häufigkeiten in einen Häufigkeitsbaum zu lösen. In einer weiteren Variante wurde sowohl die Formel-darstellung als auch die Häufigkeitsbaumverwendung trainiert (KOMBI). Die Gesamtdauer aller drei Trainingsvarianten lag bei etwa 1 Stunde.

Beschreibung der Trainings: In allen Versionen erklärte das computergestützte Trainingsprogramm zunächst anhand von Beispielen, wie Probleme aus dem Bereich der bayesschen Inferenz mithilfe des jeweiligen Repräsentationsmodells (Formel oder Häufigkeitsbaum) gelöst werden konnte. Anschließend sollten die Versuchsteilnehmer in einem Übungsmodus das Gelernte anwenden, indem sie weitere Aufgaben schrittweise am Computer lösten. Das Trainingsprogramm gab Feedback, wenn die Versuchsperson Übungsschritte nicht korrekt bzw. Anweisungen falsch verstanden hat (vgl. auch Sedlmeier/Gigerenzer 2001).

5.2 Hypothesen

Wir erwarteten, dass das Häufigkeitstraining (BAUM) zu einem besseren konzeptuellen Verständnis von bedingter Wahrscheinlichkeit bzw. bayesscher Inferenz als das formale Wahrscheinlichkeitstraining (FORM) führt. Bei Training, das zunächst formale Lösungsmethoden einführt und zusätzlich die Übersetzung in das Häufigkeitsmodell beinhaltet (KOMBI) kann etwa das Verständnisniveau des BAUM-Trainings erreicht werden. Das FORM-Training unterstützt die Fähigkeit besser, mathematisch zu abstrahieren, d.h. Lösungsalgorithmen allgemein angeben zu können.

5.3 Erhebungsinstrumente

Die verwendeten Tests wurden aus typischen Aufgaben der Bereiche bedingte Wahrscheinlichkeit und bayessche Inferenz mit verschiedenen inhaltlichen und prozessualen Anforderungen konstruiert. Die kontextbezogenen Problemstellungen sind adaptiert aus vorangegangenen Studien (Sedlmeier 1999). Zusätzlich zu Aufgabenitems, die eine numerische Lösung verlangten, beinhalteten die Tests auch Multiple-Choice-Items. Bei allen Items wurde auch eine genaue schriftliche Erläuterung des Lösungswegs verlangt. Die einzelnen Items wurden nach Anforderungen klassifiziert, sodass die Überprüfung eines differenziellen Lerngewinns in verschiedenen Anforderungsstufen möglich ist. Folgende Anforderungsstufen wurden unterschieden:

- A1: Es soll angegeben werden, welche Informationen aus dem Aufgabentext für die Lösung gebraucht werden bzw. wie einzelne Informationen das Ergebnis beeinflussen können (keine algorithmischen Verknüpfungen nötig).
- A1+: Es soll aus gegebenen Informationen eine „totale Wahrscheinlichkeit“ gebildet werden, d.h. eine Teilstruktur der bayesschen Inferenzaufgabe.
- A2: Es soll eine bayessche Inferenzaufgabe mit den Informationen aus dem Aufgabentext gelöst werden. Diese Anforderung war, wie oben beschrieben, Gegenstand der Trainings.
- A3: Es soll eine bayessche Inferenzaufgabe gelöst werden, die in ihrer Struktur von den trainierten Aufgaben abweicht. Anstatt der üblichen dichotomen Struktur, kom-

men trichotome Ereignisstrukturen vor (z.B. Kriterien für Wetter: gut, schlecht, wechselhaft).

- A4: Es soll eine bayessche Inferenzaufgabe mit extrem wahrscheinlichen bzw. unwahrscheinlichen Werten gelöst werden (z.B. bedingte Wahrscheinlichkeiten im Aufgabentext sind fast gleich 100% bzw. fast gleich 0%).

Während A1 und A1+ Teilanforderungen von A2 darstellen, sind bei A3 und A4 einfache „Transferleistungen“ nötig, d.h. es muss eine Problemstruktur erkannt werden, die von der im Training geübten abweicht. Mögliche unterschiedliche Itemschwierigkeiten wurden durch eine systematische Permutation der Darbietungsreihenfolge der Items in allen Gruppen kontrolliert.

Zur Erfassung der Abstrahierungsfähigkeiten wurde in einem Teilstest überprüft, ob die Schüler fähig sind, allgemein und ohne die kontextbezogene Einbettung zugrunde liegende mathematische Strukturen zu erkennen und zu formulieren. Zusätzliche Fragen erhoben allgemeines mathematisches Vorwissen, subjektive Präferenzen und die Einschätzung der Aufgabenschwierigkeiten.

5.4 Stichproben

In der 1. Studie wurden 47 Schüler (31 männlich/16 weiblich) der 13. Jahrgangsstufe eines Leistungskurses Mathematik (Sekundarstufe II) mit unterrichtlichen Vorkenntnissen im Themengebiet zufällig dem BAUM-Training (BAUM: N = 24, m/w: 16/8) oder formalem Training (FORM: N = 23, m/w: 15/8) zugeordnet. Die Gruppen waren hinsichtlich leistungsbezogener Merkmale weitgehend parallel (Klausur in Mathematik im Themengebiet, in Punkten [Skala 0–15]: M = 8,65, s = 2,87; BAUM M = 8,83, s = 2,59). Es wurde in beiden Gruppen üblicher Unterricht zu diesem Themengebiet durchgeführt.

Die 2. Studie wurde mit 53 Schülern (23 Jungen/30 Mädchen) ohne unterrichtliche Vorkenntnisse in Stochastik der Jahrgangsstufen 9 bis 11 (Sekundarstufe I) durchgeführt. N = 27 erhielten das Training BAUM und N = 26 erhielten ein kombiniertes Training (KOMBI). Die Gruppen waren hinsichtlich personenbezogener Merkmale weitgehend parallel, insbesondere die proportionale Zusammensetzung der Gruppen hinsichtlich Jahrgangsstufen. Die Untersuchung von Schülern der Sekundarstufe I ist auch im Hinblick auf deutliche Empfehlungen von Seiten der Didaktik zur Ausweitung der Stochastikcurricula auf frühere Stufen von Relevanz.

5.5 Ergebnisse und Diskussion

Studie 1: Sekundarstufe II

Quantitative Ergebnisse: Im Vortest (Pre) und unmittelbar nach dem Training gegebenen Nachtest (Post) wurden folgende Testleistungen in den beiden Trainingsgruppen, jeweils im gesamten Test bzw. nach Aufgabenanforderungen gegliedert, erreicht:

Tab. 1: Erreichte Testleistungen (Mittelwerte in% der erreichbaren Testscores) in der Sekundarstufe II mit unterrichtlichem Vorwissen im Gesamttest (ges) und in verschiedenen Anforderungsstufen (A1 bis A4)

Gruppe	Pre, ges.	Pre, A1	Pre, A1+	Pre, A2	Post, ges.	Post, A1+	Post, A2	Post, A3	Post, A4
FORM	47,83%	65,22%	39,13%	36,96%	48,12%	43,48%	64,13%	40,22%	43,48%
BAUM	51,52%	60,42%	50,00%	43,75%	78,33%	91,67%	93,75%	72,92%	58,33%

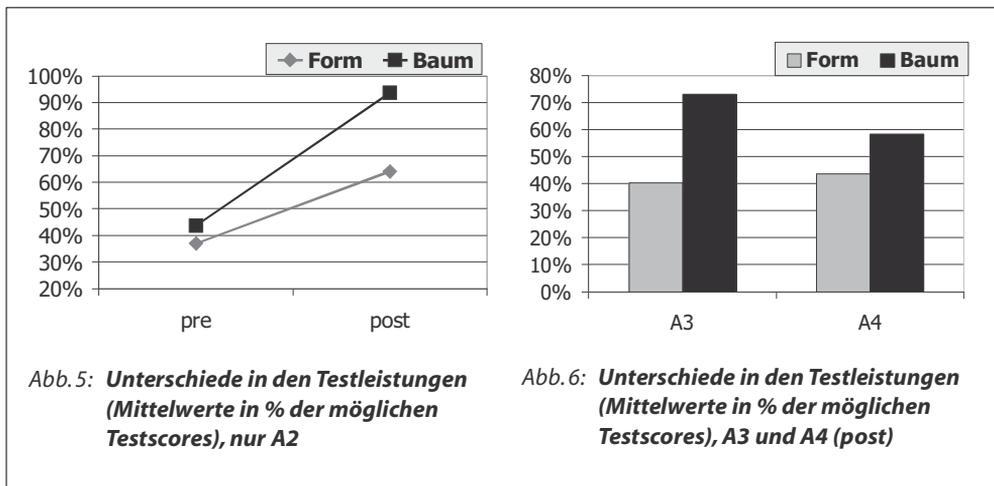
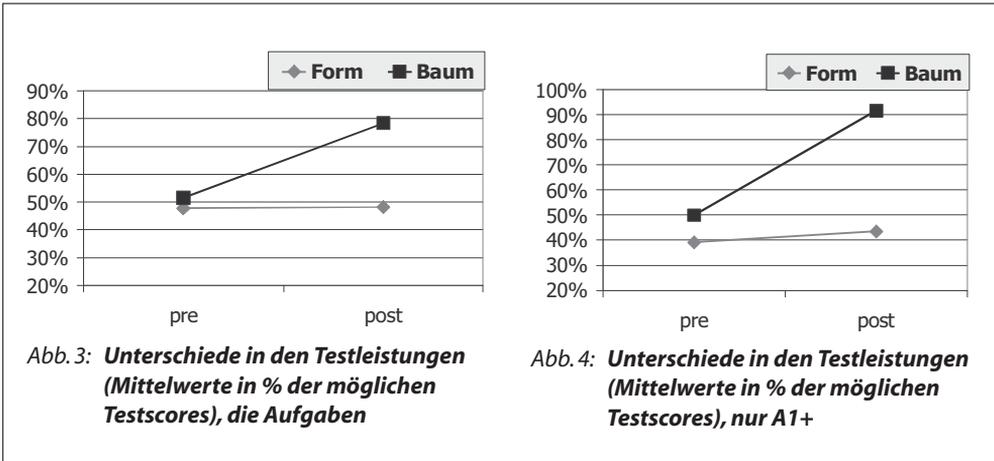
Die Unterschiede im Lernzuwachs durch das Training zwischen den Gruppen sind sehr deutlich festzustellen. Nach FORM-Training erreichen die Schüler in etwa denselben Prozentanteil an erreichbaren Testscores, während nach BAUM-Training der Prozentanteil um etwa 27 Prozentpunkte höher liegt als vorher. Tab. 2 gibt einen Überblick über (differenzielle) Unterschiede im Lernzuwachs beider Gruppen in den verschiedenen Anforderungsstufen:

Tab. 2: Differentielle Unterschiede im Lernzuwachs im Gesamttest und in verschiedenen Anforderungsstufen

Gruppe	Δ_{ges} (s. Abb. 3)	d_{ges}	Δ_{A1+} (s. Abb. 4)	d_{A1+}	Δ_{A2} (s. Abb. 5)	d_{A2}	Δ_{A3} (s. Abb. 6)	d_{A3}	Δ_{A4} (s. Abb. 6)	d_{A4}
FORM	0,29%		6,52%		27,17%		40,22%		43,48%	
BAUM	26,81%	0,67	37,50%	0,84	50,00%	0,46	72,92%	0,92	58,33%	0,355

(Δ = Mittlere Leistungssteigerungen zwischen Vor- und Nachtest in Prozentpunkten; d = Effektgrößenindex für den Unterschied zwischen den mittleren Leistungssteigerungen pro Gruppe)

Im Vortest erreichten die Schüler beider Gruppen im Mittel jeweils etwa 50% der möglichen Testscores, wobei die BAUM-Gruppe etwas bessere Vorleistungen zeigte. Während die mittleren Testleistungen der FORM-Gruppe nach dem Training unter 50% der erreichbaren Testscores blieben, erreichte die BAUM-Gruppe eine deutlich höhere Testleistung von fast 80%. Der Nachtest enthielt im Gegensatz zum Vortest keine A1-Aufgaben, da erfahrungsgemäß nach beiden Trainings alle Schüler diese Anforderungsstufe erreichen konnten. Außerdem enthielt der Nachtest auch die komplizierter strukturierten Aufgaben vom Typ A3 bzw. A4 (siehe Tab.1). Insofern ist davon auszugehen, dass das Anforderungsniveau des Nachtests insgesamt höher war als das des Vortests. Es ist zu beachten, dass nicht das absolute Leistungsniveau, sondern der (differenzielle) Unterschied im Lernzuwachs entscheidend ist. Differenziert nach Anforderungsstufen ergeben sich mittlere Unterschiedseffekte ($d = 0,46$) im Lernzuwachs bei A2-Aufgaben (siehe Abb. 5), die im Training geübt wurden. Das Häufigkeitstraining ermöglichte gegenüber dem Wahrscheinlichkeitstraining bei bereits im Themengebiet unterrichteten Schülern eines Leistungskurses Mathematik deutlich höhere Lernzuwächse. Schüler der BAUM-Gruppe konnten bei A2-Aufgaben nach dem Training im Mittel fast 95% der möglichen Punkte erreichen.



Noch größere Unterschiedseffekte in den Lernzuwächsen (bis $d = 0,92$, siehe Tab. 2) ergeben sich für die Anforderungsstufen, die nicht explizit Trainingsinhalt waren.

- A1+: Aus didaktischer Sicht scheint gerade in diesem Bereich eine besondere Schwäche der formalen Repräsentation zu liegen, da die hier nötige Umordnung nach der anderen Bedingung und Bildung der totalen Wahrscheinlichkeit nicht deutlich wird (vgl. Wassner/Krauss/Martignon 2002).
- A3 und A4: Das trainierte Lösungsverfahren mit Häufigkeitsbäumen kann offensichtlich flexibler zur Bearbeitung unterschiedlicher Aufgabenstrukturen eingesetzt werden als das formale Verfahren. Beim Auftreten „extremer“ Wahrscheinlichkeitswerte (z.B. 100%) in Anforderung A4 ist der Unterschiedseffekt jedoch geringer ($d = 0,355$).

Insgesamt zeigt sich also ein deutlicher Vorteil des Häufigkeitstrainings im Vergleich zum Wahrscheinlichkeitstraining.

Qualitative Analyse: Den Schülern wurden keine expliziten Vorgaben zur Verwendung einer bestimmten Repräsentationsform oder eines numerischen Formates in den freien Überlegungsteilen gegeben. Es zeigte sich eine gemischte Verwendung im Vortest, je nach gelernter bzw. erinnerter Lösungsmethode aus dem vorangegangenen Mathematikunterricht. Es gibt eine Präferenz für grafische Repräsentationen (FORM: 65%, BAUM: 74%) und für Wahrscheinlichkeitsformate (65% in beiden Gruppen). Naturgemäß wurde dieser Verwendungsmix von Repräsentationen durch das Training stark beeinflusst: Die Gruppe BAUM benutzte bei allen Anforderungsstufen (auch A3 und A4) im Posttest zu 96% den trainierten Häufigkeitsbaum zur Lösung. Die Gruppe FORM demgegenüber verwendete bei A2-Aufgaben im Posttest nur zu 74% die trainierte formale Lösungsmethode. Bei höheren Anforderungsstufen sank die Verwendungsrate ab (A3: 39%, A4: 52%).

Der Zusammenhang zwischen der Verwendung einer speziellen Repräsentation und ihrer korrekten Erstellung zeigt ebenfalls deutliche Unterschiede zwischen den Gruppen. In Gruppe BAUM führte ein im Posttest verwendeter Häufigkeitsbaum bei A2-Aufgaben zu 96% zu einem korrekten Ergebnis. In Gruppe FORM wurde bei der Verwendung der trainierten Formeldarstellung in A2-Aufgaben im Posttest nur zu 53% ein korrektes Ergebnis erreicht.

In einem ex-post Fragebogen schätzte die Gruppe BAUM die Aufgabenschwierigkeit im Posttest leichter ein als die Gruppe FORM (4-stufige Likert-Skala, „leicht“ und „eher leicht“: FORM A2: 65%, A3: 39%, A4: 35%. BAUM A2: 83%, A3: 75%, A4: 50%), natürlich auch als Ergebnis einer hohen Korrelation mit den höheren Lernerfolgen. Schüler der BAUM-Gruppe finden die Häufigkeitsbaumrepräsentation zu 78% einfach und verständlich. Schüler der FORM-Gruppe geben zu 48% an, die formale Repräsentation gerne zu verwenden, Schüler der BAUM-Gruppe zu 96%. 48% aus der BAUM-Gruppe, aber auch 42% aus FORM-Gruppe – wohlgemerkt nach dem formalen Training und Posttest – halten die mathematische Formel von Bayes für schwierig zu erinnern.

Hohe Verwendungsraten im Nachtest und persönliche Einschätzungen (ex post) lassen folgern, dass die Schüler der BAUM-Gruppe vom Nutzen der Häufigkeitsbäume überzeugt waren. Die qualitative Analyse der Lösungen mit dem Häufigkeitsbaum lässt eine weitgehende Einsicht in die Bearbeitung der Probleme erkennen. Bei der FORM-Gruppe hingegen wurde die trainierte Lösungsmethode mit der Formel deutlich weniger oft eingesetzt. Das spiegelt sich in den Einschätzungen wieder, wie sie in der ex-post Befragung erkennbar waren.

Die intuitive Verständlichkeit des Häufigkeitsbaumes und seine flexible Übertragbarkeit auf andere Aufgabenkontexte bzw. strukturen scheint für die auffällig hohen Lösungsraten verantwortlich zu sein. Die korrekte Verwendung formaler Repräsentation bereitet hingegen selbst nach intensivem Training Schwierigkeiten, nicht zuletzt, weil sie sich einer naiven Intuition zu entziehen scheint (vgl. Wassner/Krauss/Martignon 2002).

Studie 2: Sekundarstufe I

Quantitative Ergebnisse: Im Vortest (Pre) und unmittelbar nach dem Training gegebenen Nachtest (Post) wurden folgende Testleistungen in den beiden Trainingsgruppen, jeweils im gesamten Test bzw. nach Aufgabenanforderungen gegliedert, erreicht:

Tab. 3: Erreichte Testleistungen (Mittelwerte in% der erreichbaren Testscores) in der Sekundarstufe I ohne unterrichtlichem Vorwissen im Gesamttest (ges) und in verschiedenen Anforderungsstufen (A1 bis A3)

Gruppe	Pre, ges.	Pre, A1	Pre, A1+	Pre, A2	Post, ges.	Post, A1+	Post, A2	Post, A3
KOMBI	12,59%	25,96%	11,54%	0	43,32%	20,51%	52,88%	31,73%
BAUM	18,52%	33,33%	22,22%	0,93%	58,09%	45,68%	63,27%	51,85%

Tab.4 gibt einen Überblick über (differenzielle) Unterschiede im Lernzuwachs beider Gruppen im Gesamttest und in den verschiedenen Anforderungsstufen:

Tab. 4: Differentielle Unterschiede im Lernzuwachs im Gesamttest und in verschiedenen Anforderungsstufen

Gruppe	Δ_{ges}	d_{ges}	Δ_{A1+}	d_{A1+}	Δ_{A2}	d_{A2}	Δ_{A3}	d_{A3}
KOMBI	30,73%	0,39	8,97%	0,31	52,88%	0,32	31,73%	0,63
BAUM	39,57%		23,46%		62,34%		51,85%	

(Δ = Mittlere Leistungssteigerungen zwischen Vor- und Nachtest in Prozentpunkten;
 d = Effektgrößenindex für den Unterschied zwischen den mittleren Leistungssteigerungen pro Gruppe)

Die Testleistungen in beiden Gruppen stiegen durch das Training sehr deutlich an. Das Anforderungsniveau des Nachtests ist insgesamt höher zu bewerten als das des Vortests (vgl. oben). Differenziert nach Anforderungsstufen ergeben sich kleinere bis mittlere Unterschiedseffekte zwischen den Trainingsgruppen ($d \approx 0,3$) im Lernzuwachs bei A1+ und A2-Aufgaben. Die relativ hohen Leistungszuwächse bei Aufgaben der Aufgabenanforderung A2 zeigen, dass nicht in Stochastik unterrichtete Schüler ab Klasse 9 erfolgreich mit Problemen der Wahrscheinlichkeitsrechnung konfrontiert werden können. Die Schüler der BAUM-Gruppe konnten ihre Leistung vom Nullniveau auf ca. 63% der zu erreichenden Punkte steigern. Deutlichere Unterschiedseffekte zwischen den Trainingsvarianten ergeben sich für die Anforderungsstufe A3. Wir finden einen deutlichen Effekt ($d = 0,63$) zugunsten der BAUM-Trainingsgruppe. A4-Aufgaben waren im Test der Sekundarstufe I nicht enthalten. Insgesamt zeigt sich also ein deutlicher Vorteil des Häufigkeits- gegenüber dem Wahrscheinlichkeitstraining.

Abstrahierungsfähigkeit

Bei diesen zusätzlichen Testitems sollten die Schüler ohne unterrichtliches Vorwissen (Studie 2) von den konkreten kontextbezogenen Problemen ausgehend in mehreren Abstraktionsschritten Angaben zu allgemeinen mathematischen Zusammenhängen oh-

ne Einbettung in Aufgabenkontexte geben. Aus den Vorfragebögen zum allgemeinen mathematischen Vorwissen geht hervor, dass kein Schüler über derartiges abstraktes probabilistisches Wissen verfügte. Im Unterschied zum BAUM-Training beinhaltete das FORM-Training eine allgemeine Herleitung der Bayesformel ohne Einbettung in einen Aufgabenkontext. Es wurde daher als Hypothese angenommen, dass formales Training die Abstrahierungsfähigkeiten deutlich besser unterstützt.

Den allgemeinen Zusammenhang zwischen bedingter Wahrscheinlichkeit und konjunktiver Verknüpfung $P(A) \cdot P(B | A) = P(A \cap B)$ konnten so die Schüler der ursprünglich formal trainierten KOMBI-Gruppe zu 68,4% angeben. Doch auch nach reinem BAUM-Training konnten diesen Zusammenhang 67,2% der Schüler allgemein formulieren. Bei der Aufstellung der kompletten Bayesformel hingegen konnten wir einen klareren Vorteil des KOMBI-Trainings finden: 44,7% können die allgemeine Bayesformel für zwei Ereignisse angeben im Unterschied zu 28,9% nach BAUM-Training.

5.6 Implikationen

Aus den Ergebnissen dieser Trainingsstudie mit Schülern und Erkenntnissen früherer Studien (v.a. von Gigerenzer und Sedlmeier) kann gefolgert werden, dass Repräsentations- und Darbietungsformen entscheidenden Einfluss auf individuelle Verarbeitungsprozesse bei bayesscher Inferenz und bedingter Wahrscheinlichkeit haben. Durch ein einstündiges, computerbasiertes Training mit Darbietungsformen, die auf Häufigkeitsformaten und grafischen Baummodellen basieren, konnten nicht nur die höchsten Lern- und Verständniszuwächse erreicht werden, sondern auch eine deutliche Verbesserung der flexiblen Anwendung und eine Steigerung der Nachhaltigkeit des Gelernten. Im Hinblick auf Lehr-Lernprozesse lassen die Vorleistungen der getesteten Schülern den Schluss zu, dass bisher üblicher Stochastikunterricht (z.B. im LK Mathematik im Themenbereich bedingte Wahrscheinlichkeit) noch nicht zu optimalen Lernerfolgen führt. Die Bereicherung der Instruktion durch die Verwendung „intuitiver“ Darbietungsformen, die im Alltag leichter zu benutzen sind, wie etwa Häufigkeitsbäume, ist dringend anzuraten. Substanzielle fachliche und fächerübergreifende Kompetenzen im Umgang mit statistischer Information und realen Wahrscheinlichkeitsproblemen sind leichter und nachhaltiger zu erreichen. Eine Kombination mit auf Wahrscheinlichkeitsformaten basierendem formalem Training ist sinnvoll, wenn abstrakte mathematische Modellierung das Lernziel ist. Die erreichbaren hohen Lernzuwächse bei Schülern der Sekundarstufe I bestärken darüberhinaus Forderungen eines früheren Beginns der Stochastik in allen Mathematiklehrplänen.

Literatur

- Anderson, J.R./Corbett, A./Koedinger, K./Pelletier, R. (1995): Cognitive tutors: Lessons learned. In: *The Journal of the Learning Sciences* 4, S. 167–207.
- Bea, W. (1995): Stochastisches Denken – Analysen aus kognitionspsychologischer und didaktischer Perspektive. In: Crott, H.W./Scholz, R.W. (Hrsg.): *Psychologie des Entscheidungsverhaltens und des Konflikts*. Frankfurt a.M.: Europäischer Verlag der Wissenschaften.
- Biehler, R. (1997): Students' difficulties in practising computer supported data analysis – Some hypothetical generalizations from results of two explanatory studies. In: Garfield, J./Burrill, G. (Hrsg.): *Research on the role of technology in teaching and learning statistics*. Voorburg: International Statistical Institute, S. 169–190.
- Biehler, R. (2001): Statistische Kompetenz von Schülerinnen und Schülern – Konzepte und Ergebnisse empirischer Studien am Beispiel des Vergleiches von statistischen Verteilungen. In: Borovcnik, M./Engel, J./Wickmann, D. (Hrsg.): *Anregungen zum Stochastikunterricht*. Hildesheim: Franzbecker, S. 97–114.
- Borovcnik, M. (1992): *Stochastik im Wechselspiel von Intuitionen und Mathematik, Lehrbücher und Monographien zur Didaktik der Mathematik, Band 10*. Mannheim: BI-Wissenschaftsverlag.
- Brown, J.S./Collins, A./Duguid, P. (1989): Situated Cognition and the Culture of Learning. In: *Educational Researcher* 18 (1), S. 32–42.
- Cosmides, L. (1989): The logic of social exchange: Has natural selection shaped how humans reason?. In: *Cognition* 31, S. 187–276.
- Cosmides, L./Tooby, J. (1996): Are Humans Good Intuitive Statisticians after all? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty. In: *Cognition* 58, S. 1–73.
- Cummins, D.D. (1998): Social norms and other minds: the evolutionary roots of higher cognition. In: Cummins, D.D./Allen, C. (Hrsg.): *The evolution of mind*. New York: Oxford University Press. S. 31–50.
- Engel, J. (2001): Datenorientierte Mathematik und beziehungshaltige Zugänge zur Statistik: Konzepte und Beispiele. In: Borovcnik, M./Engel, J./Wickmann, D. (Hrsg.): *Anregungen zum Stochastikunterricht*. Hildesheim: Franzbecker, S. 63–82.
- Fischler, H. (2001): Verfahren zur Erfassung von Lehrer-Vorstellungen zum Lehren und Lernen in den Naturwissenschaften. In: *Zeitschrift für Didaktik der Naturwissenschaften* 7, S.105–120.
- Freudenthal, H. (1983): *Didactical phenomenology of mathematical structures*. Dordrecht : Reidel.
- Garfield, J. (1993): Teaching Statistics Using Small-Group Cooperative Learning. In: *Journal of Statistics Education* 1 (1).
- Gigerenzer, G. (1993): Die Repräsentation von Information und ihre Auswirkung auf statistisches Denken. In: Hell, W./Fiedler, K./Gigerenzer, G. (Hrsg.): *Kognitive Täuschungen – Fehl-Leistungen und Mechanismen des Urteilens, Denkens und Erinnerns*. Heidelberg: Spektrum, S. 99–127.
- Gigerenzer, G. (1999): Mentale Fakultäten, methodische Rituale und andere Stolpersteine. In: *Zeitschrift für Psychologie* 207, S. 287–297.
- Gigerenzer, G. (in press): *Calculated Risk: What Numbers really tell us about our lives*. New York: Simon & Schuster.
- Gigerenzer, G./Hoffrage, U. (1995): How to improve Bayesian reasoning without instruction: Frequency formats. In: *Psychological Review* 102, S. 684–704.
- Heitele, D. (1976): *Didaktische Ansätze zum Stochastikunterricht in Grundschule und Förderstufe*. Dissertation, Dortmund.
- Kahneman, D./Slovic, P./Tversky, A. (1982): *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Krauss, S./Hertwig, R. (2000): Muss DNA-Evidenz schwer verständlich sein? Der Ausweg aus einem Kommunikationsproblem. In: *Monatsschrift für Kriminologie und Strafrechtsreform* 3, S. 155–162.

- Martignon, L. (2000): Repräsentation von Information in mathematischen Kontexten: eine kognitionspsychologische Vision. In: Neubrand, M./Jahnke, T (Hrsg.): Beiträge zum Mathematikunterricht 2000. Hildesheim: Franzbecker.
- Martignon, L./Wassner, C. (2001): Repräsentation von Information in der Wahrscheinlichkeitstheorie. In: Borovcnik, M./Engel, J./Wickmann, D. (Hrsg): Anregungen zum Stochastikunterricht. Hildesheim: Franzbecker, S. 163–170.
- NCTM – National Council of Teachers of Mathematics (2000): Principles and Standards for school mathematics. Reston, Va.: National Council of Teachers of Mathematics.
- Piaget, J./Inhelder, B. (1951/1975): The origin of the idea of chance in children (L. Leake, Jr./P. Burrell/H.D. Fischbein, Transl.). New York: Norton.
- Piattelli-Palmarini, M. (1994): Inevitable illusions: How mistakes of reason rule our minds. New York: Wiley.
- Riemer, W. (1985): Neue Ideen zur Stochastik, Lehrbücher und Monographien zur Didaktik der Mathematik, Band 3. Mannheim: Bibliographisches Institut.
- Sedlmeier, P. (1997): BasicBayes: A tutor system for simple Bayesian inference. In: Behavior Research Methods, Instruments, & Computers 29, S. 328–336.
- Sedlmeier, P. (1998): The distribution matters: Two types of sample-size tasks. In: Journal of Behavioral Decision Making 11, S. 281–301.
- Sedlmeier, P. (1999): Improving statistical reasoning: Theoretical models and practical implications. Mahwah: Lawrence Erlbaum Associates.
- Sedlmeier, P. (2000a): How to improve statistical thinking: Choose the task representation wisely and learn by doing. In: Instructional Science 28, S. 227–262.
- Sedlmeier, P. (2000b): Wie kann der gymnasiale Statistikunterricht von der psychologischen Urteilsforschung profitieren? In: Neubrand, M./Jahnke, T (Hrsg.): Beiträge zum Mathematikunterricht 2000. Hildesheim: Franzbecker, S. 599–602.
- Sedlmeier, P. (2001): Statistik ohne Formeln. In: Borovcnik, M./Engel, J./Wickmann, D. (Hrsg): Anregungen zum Stochastikunterricht. Hildesheim: Franzbecker, S. 83–95.
- Sedlmeier, P./Gigerenzer, G. (2000): Was Bernoulli wrong? On intuitions about sample size. In: Journal of Behavioral Decision Making 13, S. 133–139.
- Sedlmeier, P./Gigerenzer, G. (2001): Teaching Bayesian reasoning in less than two hours. In: Journal of Experimental Psychology: General 130, S. 380–400.
- Sedlmeier, P./Köhlers, D. (2001): Wahrscheinlichkeiten im Alltag. Braunschweig: Westermann-Schulbuchverlag.
- Sedlmeier, P./Wettler, M. (1998): Was muss ein Tutorsystem „wissen“? In: Zeitschrift für Pädagogische Psychologie 12, S. 219–235.
- Stern, E./Aprea, C./Ebner, H. (in press): Improving cross-content representation. Learning and Instruction.
- Wassner, C./Krauss, S./Martignon, L. (2002): Muss der Satz von Bayes schwer verständlich sein? In: Praxis der Mathematik 44/1, S. 12–16.
- Wickmann, D. (1990): Bayes-Statistik – Einsicht gewinnen und entscheiden bei Unsicherheit, Lehrbücher und Monographien zur Didaktik der Mathematik, Band 4. Mannheim: BI-Wissenschaftsverlag.

Anschrift der Autoren:

Christoph Wassner, Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, 14195 Berlin.

Laura Martignon, Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, 14195 Berlin.

Peter Sedlmeier, Technische Universität Chemnitz, Institut für Psychologie, Wilhelm-Raabe-Straße 43, 09120 Chemnitz.

Kristina Reiss/Frank Hellmich/Joachim Thomas

Individuelle und schulische Bedingungs- faktoren für Argumentationen und Beweise im Mathematikunterricht¹

1. Theoretischer Hintergrund

1.1 Beweisen und Begründen in der mathematikdidaktischen Diskussion

Die Frage, welches Verständnis Schülerinnen und Schüler von mathematischem Begründen und Beweisen haben, wird in der mathematikdidaktischen Forschung vielfach diskutiert. Korrektes Beweisen und die Fähigkeit zu einem logisch begründeten Vorgehen bei Problemlösungen werden als wichtige Ziele des Mathematikunterrichts angesehen. Diese Fähigkeiten bilden das Fundament der Fortentwicklung der Mathematik, sie zählen darüber hinaus zu den Basisqualifikationen, die für eine mathematische Grundbildung unerlässlich sind.

Es würde allerdings zu kurz greifen, die Fähigkeit zum Begründen, also zum logisch konsistenten Argumentieren, nur unter mathematischen Aspekten zu sehen. Vielmehr handelt es sich dabei um eine Fähigkeit, die auch fachübergreifend bedeutsam ist. Begründen, Beweisen und sachangemessen logisch Argumentieren sind zentrale Aspekte wissenschaftlichen Denkens und Arbeitens. Darüber hinaus verbergen sich hinter diesen Begriffen aber auch prototypische Lernziele für einen diskursiven Unterricht. Ein solcher Unterricht betont einen rational argumentativen Dialog zwischen allen Beteiligten, also zwischen den Schülern und dem Lehrer sowie zwischen Schülern untereinander. Argumentationen und Beweise werden damit als Aspekt des Mathematikunterrichts betrachtet, der nicht nur in einzelnen, isolierten Unterrichtseinheiten thematisiert werden darf, sondern in allen Klassenstufen gepflegt werden sollte.

Die mathematikdidaktische Forschung hat in den letzten Jahren unterschiedliche Fragestellungen zu diesem Themenbereich untersucht. So gibt es Arbeiten über die Fähigkeit zum logischen Begründen bei mathematikbezogenen Problemstellungen (Healy/Hoyles 1998; Reiss/Klieme/Heinze 2001) oder über individuelle Schemata (Harel/Sowder 1998). Darüber hinaus werden Aspekte des Beweizens auch in Untersuchungen über mathematikbezogene Beliefs einbezogen (Törner 1998; Pehkonen/Törner 1999).

In der Bundesrepublik rückte die Frage nach dem Beweis- und Begründungsverständnis durch die TIMS-Studie in die Diskussion. Die Ergebnisse der Studie zeigen, dass das durchschnittliche Leistungsniveau von Schülerinnen und Schülern der Jahrgangsstufen 7 und 8 in der Bundesrepublik nur knapp unter dem Fähigkeitsniveau liegt,

¹ Das Projekt wird von der DFG gefördert im Rahmen des DFG-Schwerpunktprogramms »Bildungsqualität von Schule« (Geschäftszeichen: RE 1247/4-1).

das hinreichend für ein sicheres Verständnis von mathematischen Konzepten und Verfahren ist, wobei allerdings eine große Streuung in den Leistungen zu beobachten ist (Baumert u.a. 1997). Im Rahmen der PISA-Studie wurden diese Ergebnisse erst kürzlich bestätigt (Klieme u.a. 2001). Das hier geforderte mathematische Wissen erschöpft sich gerade nicht in abfragbaren Algorithmen, sondern beinhaltet wesentlich Prozesse des Modellierens, des Argumentierens und des Begründens (Neubrand u.a. 2001). Deutsche Schüler haben demnach Schwierigkeiten bei der Bearbeitung von Items, die argumentative Fähigkeiten erfordern.

Aber auch in Ländern wie z.B. Großbritannien (das im Rahmen von PISA dem oberen Leistungsbereich zugeordnet werden konnte) zeigte sich, dass Schüler Defizite im mathematikbezogenen deduktiven Schließen haben. Dies bestätigt eine Untersuchung von Healy/Hoyles (1998) mit mehr als 2400 Probanden. Die Autorinnen kommen zu dem Ergebnis, dass das Beweisverständnis bei Schülern der zehnten Klasse gering ist. Es zeigte sich, dass sie kaum in der Lage waren, einen Beweis zu führen oder in ihrer Argumentation deduktive Folgerungen zu verwenden. Stattdessen wurden häufig empirische Begründungen benutzt. Bei der eigenen Beweisführung bevorzugten sie eher einen narrativen als einen formalen Stil. Formal geführte Beweise unterscheiden sich dabei von narrativen Beweisen insbesondere durch eine verstärkte Verwendung der mathematischen Symbolsprache.

1.2 *Beliefs von Schülern und Lehrern über Mathematik*

Im Bereich der Mathematikdidaktik hat sich ein Forschungszweig entwickelt, der unter dem Stichwort *Beliefs* beschrieben werden kann (Törner 1998). Ergebnisse dieser Forschung zeigen einen weiteren möglichen Begründungszusammenhang für ein schwaches Abschneiden bei mathematischen Aufgabenstellungen auf. So hängen danach Erfolge beim Problemlösen nicht ausschließlich vom mathematischen Wissen und Können ab, sondern die Vorstellung von Mathematik („Belief“) muss als erklärender Faktor mit einbezogen werden.

Wenn die mathematischen Beliefs der Schüler einen zentralen Kontext für die Entwicklung mathematischer Fähigkeiten und Fertigkeiten darstellen, so ergibt sich daraus die grundlegende Fragestellung, ob Schülervorstellungen durch Lehrer bzw. durch Unterrichtsverfahren gefördert oder sogar vermittelt werden. Erste Ergebnisse, dass die Vorstellungen von Lehrern zur Mathematik den Unterricht maßgeblich prägen, findet man bei Jacobs u.a. (1997), die den amerikanischen und japanischen Unterricht miteinander vergleichen. Darüber hinaus zeigen die Untersuchungen zu Vorstellungen über Mathematik bei Mathematiklehrern in Deutschland von Grigutsch/Raatz/Törner (1998), dass für Lehrer die Entwicklung mathematischen Wissens in einem aufgabenbezogenen Verstehens- und Erkenntnisprozess ein wesentliches Merkmal ist. Dennoch ist die dahinter liegende Mathematik ein eher statisch gesehenes Objekt, das in geeigneter Weise aufbereitet werden muss. Somit ist das Produkt zentrales Anliegen, der Prozess spielt nur als adäquater Lernprozess eine Rolle. Mathematikbezogene Beliefs sind also

einerseits eine Determinante, die die konkrete Unterrichtsgestaltung beeinflusst. Sie sind darüber hinaus ein individueller Bedingungsfaktor, der in Beziehung zur mathematischen Kompetenz gesehen werden muss.

1.3 Ein Phasenmodell mathematischen Beweisens

Es besteht die Vermutung, dass mathematikbezogene Beliefs auch das Begründungsverständnis und das Führen von Beweisen beeinflussen. Dabei könnte ein möglicher Zusammenhang zwischen den Schwierigkeiten, die Schüler bei Beweisaufgaben haben, und ihren Vorstellungen über den Charakter des Beweisens in einem inadäquaten mentalen Modell des Beweisführungsprozesses liegen. Es liegt daher nahe, Problemlöseprozesse von Schülern mit denen mathematischer Experten zu vergleichen.

Ein Modell, das den Beweisprozess von Experten zusammenfasst, beschreibt Boero (1999). Er unterscheidet sechs Phasen der Beweisführung: (1) Exploration der Problemstellung, Entwicklung einer Hypothese, Identifikation möglicher Argumente; (2) Formulierung dieser Hypothese gemäß den Konventionen; (3) Exploration der Hypothese und möglichen Argumentverknüpfungen; (4) Auswahl von Argumenten und Verknüpfung in einer Kette von Deduktionsschlüssen; (5) Organisation der Argumente in einen Beweis, der den mathematischen (Publikations-)Standards entspricht; (6) Annäherung an einen formalen Beweis.

Dieser Prozess, insbesondere die explorativen Schritte, bleiben für manchen Mathematiklehrer und dementsprechend auch für Schüler weitgehend intransparent. Der Schüler sieht nur das (im Schulbuch gegebene) Endergebnis im Sinne einer eindeutigen Schlussfolgerungskette mit definiertem Anfangs- und Endzustand. Ein Einblick in das Problemlöseverhalten des Experten mit seinen explorativen Komponenten und Irrwegen bleibt ihm hingegen versagt. So gelangt der Schüler und oft schon der Lehrer zu einem idealisierten mentalen Modell der Beweisführung, das letztendlich die Entwicklung geeigneter Problemlöseheuristiken verhindert.

1.4 Mathematisches Arbeiten und das wissenschaftliche Grundverständnis

Es lassen sich hinsichtlich mathematischer Problemlöseprozesse verschiedene Parallelen zur Entwicklung des wissenschaftlichen Denkens aufzeigen. Beispielsweise konnten entsprechende Forschungen zeigen, dass Fehler bei der Entwicklung und beim Prüfen von Hypothesen auf ein inadäquates metawissenschaftliches Verständnis zurückgeführt werden können. Induktionsschlüsse aus empirisch gewonnenen Daten werden danach im Sinne eines deduktiven wissenschaftlichen Modells überinterpretiert (Thomas 1997).

Betrachtet man Beweisen als Problemlöseprozess (Koedinger 1998), so findet man auch im Expertenverhalten Induktionsschlüsse. Die Sichtweise von Schülern zeigt aber die Besonderheit, dass diese induktiv gewonnenen Erkenntnisse nicht mehr durch ein deduktives Vorgehen im Sinne einer hypothetisch-deduktiven Strategie abgesichert werden (Reiss/Thomas 2000). In mehreren Studien zum wissenschaftlichen Denken im Ju-

gendalter wurde der Einfluss inhaltlicher und formaler Plausibilität auf die Leistungen bei der Konstruktion von Hypothesen und ihrer Prüfung untersucht (Thomas/ Schillig 1996; Thomas 1997). Schüler, die sich in ihren Schlussfolgerungen auf inhaltliche Plausibilität stützten, suchten im Rahmen der Hypothesenprüfung ausschließlich Information, die ihre Hypothesen scheinbar bestätigte, auch wenn diese Information redundant war. Schüler, die in ihren Schlussfolgerungen von Inhalten abstrahierten, suchten hingegen gezielt Information, die ihre Hypothesen falsifizieren konnte. Nicht untersucht ist bisher der Zusammenhang zwischen mathematischen Kompetenzen und Aspekten des wissenschaftlichen Denkens. Es ist zu vermuten, dass beide Kompetenzbereiche nicht unabhängig voneinander sind.

1.5 Forschungsfragen und Ziele der Studie

Die bisherige Forschung zur Genese des Beweisens und Begründens im Mathematikunterricht zeigt die Bedeutung der Einbettung diesbezüglicher Ergebnisse in einen allgemeineren Kontext unterrichtlicher Bedingungen. Da die aktuelle Forschung zu Bedingungsfaktoren schulischer Leistungen von der multiplen Interaktion individueller und schulischer Bedingungsfaktoren sowie von einer Multikriterialität der Effekte ausgeht, ergibt sich die Notwendigkeit einer integrativen Sichtweise der verschiedenen Einflussfaktoren.

Auf dieser Grundlage sind die Ziele der Studie eine Beschreibung wesentlicher individueller kognitiver und nicht-kognitiver Determinanten der Lösungskompetenz bei Beweisaufgaben, eine Erfassung von Unterrichtsvariablen, die für die Entwicklung des Beweisverständnisses von zentraler Bedeutung sind, und die Prüfung der Interdependenzen zwischen Unterrichtsvariablen und individuellen Variablen. Damit verbunden sind die folgenden Forschungsfragen:

- Über welche argumentativen Kompetenzen und über welches deklarative Wissen verfügen Schülerinnen und Schüler?
- Welche Rolle spielt die Kompetenz zur Beurteilung von Beweisen?
- Welche Rolle spielt das wissenschaftliche Grundverständnis für die argumentative Kompetenz?
- Haben die mathematikbezogenen Beliefs von Schülern und Lehrern einen Einfluss auf die Entwicklung des Beweisverständnisses?

2. Beschreibung der Untersuchung

2.1 Stichprobe

An der Untersuchung nahmen 358 Schülerinnen und 301 Schüler der siebten bzw. achten Jahrgangsstufe an Gymnasien aus 27 Schulklassen teil. Die Lehrerinnen und Lehrer der Klassen waren in die Untersuchung einbezogen.

2.2 Untersuchungsmethode

Die Schülerinnen und Schüler nahmen am Ende der siebten Jahrgangsstufe an einem Vortest teil. Dabei wurden Items zur Geometrie, zur Methodenkompetenz, zum wissenschaftstheoretischen Grundverständnis und zu den mathematikbezogenen Beliefs vorgegeben. Der Leistungstest für die siebte Klasse enthielt sechs Aufgaben zu Basisqualifikationen im Bereich der Geometrie, die vornehmlich elementares Begriffswissen und einfache Anwendungen abfragten. Es handelte sich insbesondere um Aufgaben zur Kongruenz und zu Winkelbeziehungen. Außerdem waren in diesem Test sieben Aufgaben zum Argumentieren und Begründen enthalten, die von der Konstruktion eines Begründungsschrittes bis hin zu der Notation mehrerer Argumentationsschritte sowie einer adäquaten Verknüpfung reichten. Die einzelnen Items wurden in Anlehnung an den gängigen Kanon der Rahmenrichtlinien und Schulbücher auf der Grundlage des Beweismodells von Boero (1999) konzipiert.

Die Aufgaben zur Erhebung der Methodenkompetenz im Hinblick auf das Begründen und Beweisen wurden bereits in einer Pilotstudie zu dieser Untersuchung verwendet. Die Idee dazu stammte aus einer Untersuchung von Healy/Hoyles (1998). Die Aufgaben wurden auf die Belange der vorliegenden Studie zugeschnitten und dem gültigen Curriculum angepasst. In diesem Test werden die Schüler aufgefordert, vier verschiedene Lösungen zu einer Beweisaufgabe (Winkelsumme im Dreieck) zu beurteilen. Eine der Aufgaben beruht auf einem empirischen Argument (Nachmessen) und ist kein mathematischer Beweis. Ebenfalls nicht korrekt ist eine Lösung, die einen Zirkelschluss enthält. Zwei andere Lösungen sind hingegen korrekt, wovon eine in eher narrativer Form und die andere in eher formaler Form aufgeschrieben ist.

Zur Bewertung von Kompetenzen im Bereich des wissenschaftlichen Denkens wurden insgesamt acht Aufgaben zum hypothetisch-deduktiven Denken konzipiert. Die Aufgaben bestanden jeweils aus zwei Teilen. Im ersten Teil der Aufgaben wurden die Schüler dazu aufgefordert, aus einer vorgegebenen Information Schlussfolgerungen bezüglich der Reduktion eines gegebenen Problemraums zu ziehen (Klahr/Dunbar 1993). Im zweiten Teil sollten sie eine Informationsanordnung auswählen, die zu einer weiteren Reduktion des Suchraums führen kann.

Zur Erfassung der Beliefs über Mathematik von Schülerinnen und Schülern wurde ein Fragebogen mit 24 Items konzipiert, der sich ganz wesentlich an die von Grigutsch u.a. (1998) entwickelten und von Klieme/Ramseier (2001) verfeinerten Instrumente anlehnt. Dabei fanden im Wesentlichen die von Grigutsch u.a. (1998) identifizierten Kategorien *Anwendung*, *Prozess*, *Schema* und *Formalismus* im Fragebogen Berücksichtigung.

Auch die Lehrer wurden zu ihren mathematikbezogenen Beliefs befragt. Dabei fand im Wesentlichen der gleiche Fragebogen Verwendung, der auch der Schülerbefragung zugrunde lag. Zusätzlich wurde nach dem bevorzugten Unterrichtsstil gefragt. Eingesetzt wurde dabei ein auf die spezifischen Belange der Untersuchung zugeschnittener Fragebogen zum Lehren und Lernen von Mathematik (Peterson u.a. 1989; Stern/Staub 2000). Der Fragebogen beinhaltete Aussagen über die Grundkonzeption von Unterricht. Die Lehrerinnen und Lehrer sollten Einschätzungen hinsichtlich ihrer Einstellung und

ihres Wissens über Instruktion (fragend-entwickelnd bzw. diskursiv) sowie über das Lernen abgeben.

3. Ergebnisse

3.1 Individuelle Bedingungen und Leistungsvariablen in Klasse 7/8

Basiskompetenzen, Beweisverständnis und argumentative Kompetenz

Im Leistungstest Geometrie lösen die Schüler im Mittel Aufgaben zu Basiskompetenzen ($M = 7,4$; $s = 2,5$; maximal 12 Punkte) deutlich besser als Aufgaben zu Kompetenzen im Bereich des Argumentierens und Begründens ($M = 5,2$; $s = 3,7$; maximal 14 Punkte). Bei den Aufgaben, die eher Argumentieren bzw. Beweisen verlangen, können noch einmal zwei Typen entsprechend der Schwierigkeit der Aufgabenstellung unterschieden werden. So werden Aufgaben, bei denen die Argumentation eine einfache Identifizierung und Anwendung gelernter Begriffe verlangt (z.B. des Begriffs „Scheitelwinkel“ und einer Rechnung damit), noch relativ gut gelöst. Ein eigenständiges Zusammenfügen verschiedener Fakten in eine Argumentationskette wird hingegen eher nicht geleistet.

In Anlehnung an die Stufen mathematischer Kompetenz, die Klieme (2000) bei der Bearbeitung von TIMSS-Items in der Sekundarstufe II herausstellen konnte, lassen sich hinsichtlich der beiden Subtests zu geometrischen Qualifikationen in unserer Untersuchung drei Kompetenzstufen identifizieren. Es zeigt sich, dass die theoretische Einordnung der Items in die oben beschriebenen Kompetenzstufen (Basiswissen, einfache anwendungsbezogene Argumentationen, komplexere Argumentationsketten) durch die empirischen Daten gestützt wird. Dabei erfolgte die Einordnung der Items in die Kategorien aufgrund mathematikdidaktischer Kriterien. In Bezug auf die Mittelwerte erhält man die in Tabelle 1 dargestellte Einteilung. Die Zahlen bedeuten jeweils die Lösungswahrscheinlichkeiten über alle Items der entsprechenden Stufe.

	Kompetenzstufe I Einfaches Anwenden von Regeln und elementares Schlussfolgern $M = 0,69$	Kompetenzstufe II Argumentieren und Begründen (einschrittig) $M = 0,56$	Kompetenzstufe III Argumentieren und Begründen (mehrschrittig) $M = 0,24$
unteres Leistungsdrittel $N = 238$	0,51	0,22	0,00
mittleres Leistungsdrittel $N = 215$	0,72	0,61	0,18
oberes Leistungsdrittel $N = 206$	0,85	0,89	0,50

In der ersten Kompetenzstufe werden keinerlei formale Argumentationen gefordert. Vielmehr steht das Anwenden von Begriffen und Regeln im Vordergrund. Diese Stufe zeigt in allen Leistungsgruppen einen gewissen Anteil korrekter Lösungen, insbesondere stellen die Aufgaben auch für Schüler im unteren Leistungsdrittel keine Überforderung dar. Es wird bei Aufgaben der Kompetenzstufe I einfaches Basiswissen im Bereich der Geometrie vorausgesetzt, zum Teil müssen elementare arithmetische Operationen angewendet werden.

Die Aufgaben der zweiten Kompetenzstufe werden insbesondere im unteren Leistungsdrittel von deutlich weniger Schülern gelöst. Die erfolgreiche Bearbeitung dieser Items setzt bereits Begriffs- und Faktenwissen im Bereich der Mittelstufengeometrie voraus, das darüber hinaus auf Problemsituationen angewendet und passend notiert werden muss. Auch hier zeigt das obere Leistungsdrittel eine deutliche Sicherheit im Umgang mit den Anforderungen.

Die Aufgaben der dritten Kompetenzstufe werden fast ausschließlich von Schülern im oberen Leistungsdrittel angemessen bearbeitet. Neben der adäquaten Notation von Argumenten und Begründungen zu einem Beweisschritt, ist diese Teilstichprobe in der Lage, Argumente sinnvoll zu verknüpfen. Auf dieser Stufe wird eigenständiges, zum Teil auch kreatives Problemlösen und Argumentieren verlangt.

Ein wesentliches Ergebnis in Bezug auf den Leistungstest sind die erheblichen Unterschiede zwischen einzelnen Klassen. Bei einer Punktzahl von insgesamt 26 Punkten (je zwei maximal erreichbare Punkte pro Aufgabe) liegen die Mittelwerte der Klassenergebnisse zwischen $M_{\min} = 5,67$ und $M_{\max} = 17,60$ Punkten. Insgesamt liegt der Mittelwert über alle 27 Klassen bei $M = 12,42$ ($s = 3,07$). In den drei schwächsten Klassen wurden (mit Ausnahme nur einer Person) von den Schülerinnen und Schülern jeweils weniger als die Hälfte der möglichen Punkte erreicht. Es gibt in diesen Klassen somit keine vergleichsweise starken Schüler ($M = 6,13$; $s = 3,02$; Minimum = 1; Maximum = 14), wohingegen in den drei leistungsstärksten Klassen eine breite Fächerung der Leistungen gegeben ist ($M = 16,66$; $s = 4,25$; Minimum = 5,5; Maximum = 24,5). Die einfacheren Aufgaben werden von den stärkeren Klassen gut bis sehr gut, von den schwächeren Klassen immer noch zufriedenstellend gelöst. Unterschiede ergeben sich aber bei Aufgaben, deren Lösung Argumentationen erfordert. Dies entspricht den Aufgaben der Kompetenzstufen II und III. Hier gibt es bei den Schülern der schwächeren Klassen kaum noch korrekte Lösungen. Hochsignifikante Unterschiede im argumentativen und begründenden Verhalten liegen zwischen leistungsschwächeren und leistungsstärkeren Klassen vor ($t = 13,81$; $df = 436$; $p < 0,001$). Die Art der Antworten zeigt, dass die Inhalte des Tests allen Schülern vertraut sind, die Inhalte also in allen Klassen im Unterricht behandelt worden waren.

Unterschiede zwischen leistungsstärkeren und leistungsschwächeren Klassen konnten weiterhin im argumentativen Verhalten in Bezug auf einfache Rechenaufgaben festgestellt werden. Bei prinzipiell korrekt gelösten Aufgaben, in denen der Aufgabentext nur nach einem Ergebnis, nicht aber nach einer Herleitung verlangte, begründeten Schüler aus leistungsstärkeren Klassen ihre Rechnungen signifikant häufiger als Schüler aus leistungsschwächeren Klassen ($t = 10,02$; $df = 408$; $p < 0,001$). Schüler, die auch ein-

fache Rechnungen begründeten, lösten andererseits die schwierigeren Aufgaben signifikant besser als Schüler, die nur ein (korrektes) Ergebnis nannten ($t = 48,52$; $df = 254$; $p < 0,001$). Die Ergebnisse deuten an, dass ein Klassenklima, in dem argumentative Prozesse gefördert werden, zu besseren Leistungen führt.

Methodenkompetenz im Bereich des Beweisens

Bei den Lösungen der Aufgaben zur Methodenkompetenz zeigt sich, dass es den Schülern insbesondere schwer fällt, nicht korrekte bzw. nicht allgemein gültige Argumentationen als (im Sinne eines mathematischen Beweises) nicht korrekt zu identifizieren. Während korrekte Beweise mit einer Lösungswahrscheinlichkeit von $w_k = 0,67$ als korrekt identifiziert werden, werden nicht korrekte Argumentationen nur mit einer Wahrscheinlichkeit von $w_f = 0,35$ als fehlerhaft bzw. empirisch erkannt.

Darüber hinaus wird allein aus den Unterschieden in den Mittelwerten deutlich, dass die Beurteilung von Beweisen wesentlich einfacher als eigenständiges Formulieren von Beweisen ist. In Bezug auf normierte Werte ergibt sich für die Methodenkompetenz ein Mittelwert von $M = 0,51$ mit $s = 0,21$. Für die Kompetenz im Bereich des Argumentierens und Begründens (betrachtet als Lösungswahrscheinlichkeit für Aufgaben der dritten Kompetenzstufe) war der entsprechende Mittelwert $M = 0,24$ mit $s = 0,26$. Selbst wenn man die Lösungswahrscheinlichkeiten der Aufgaben in der zweiten und dritten Kompetenzstufe zusammenfasst, zeigt sich bei einem Mittelwert von $M = 0,37$ mit $s = 0,26$ ein klarer Unterschied zu den Aufgaben zur Methodenkompetenz. Es können damit sowohl Ergebnisse von Healy/Hoyles (1998; 10. Schuljahr) als auch eigene Ergebnisse (Reiss/Klieme/Heinze 2001; 13. Schuljahr) für die 7. Jahrgangsstufe repliziert werden.

Kompetenzen im wissenschaftlichen Denken

Wie bereits in anderen Studien gezeigt wurde, lassen sich Schüler bei der Lösung von Aufgaben zum wissenschaftlichen Denken häufig von plausiblen Argumenten leiten, die nicht logisch und zwingend aus der Aufgabenstellung geschlossen werden können (Thomas 1997). Dies wurde in der vorliegenden Untersuchung bestätigt. So konnten die Schüler vor allem Aufgaben richtig lösen, bei denen die korrekte Lösung auch inhaltlich plausibel war. Ein großer Teil der Schülerinnen und Schüler versagte bei Aufgaben, bei denen die inhaltliche Plausibilität keine Lösungshilfe darstellte. Es gab kaum Schüler, die Aufgaben mit verfremdeten Inhalten lösen konnten, aber bei Aufgaben versagten, die plausible Inhalte aufwiesen ($\chi^2 = 144,94$; $p < 0,01$). Tabelle 2 zeigt die Prozentanteile in Bezug auf die Verteilung der Lösungen bei Aufgaben mit verfremdeten und plausiblen Inhalten. Danach werden die verfremdeten Aufgaben nur von einem Viertel der Schüler vollständig korrekt gelöst, während mehr als die Hälfte mit diesen Aufgaben nichts anfangen kann. Nur ein geringer Teil löst verfremdete Aufgaben korrekt und macht gleichzeitig Fehler bei den plausiblen Aufgabenstellungen.

Tab. 2: Verteilung der Lösungen bei Aufgaben mit verfremdeten und plausiblen Inhalten (in %)

	verfremdete Aufgaben nicht gelöst	verfremdete Aufgaben teilweise gelöst	verfremdete Aufgaben vollständig gelöst
plausible Aufgaben nicht gelöst	24,7	3,3	1,7
plausible Aufgaben teilweise gelöst	16,8	8,2	6,5
plausible Aufgaben gelöst	12,1	8,5	18,1

Ein weiterer Aspekt der Untersuchung war die Betrachtung des Zusammenhangs der Fähigkeiten zum wissenschaftlichen Denken auf die Leistungen der Schüler beim Begründen und Beweisen. Wie Tabelle 3 zeigt, sind Schüler, die über formale Strategien im wissenschaftlichen Denken verfügen, den Schülern, die sich im Wesentlichen von inhaltlichen Plausibilitätserwägungen leiten lassen, auf allen Kompetenzstufen überlegen. Besonders deutlich wird dieser Unterschied bei Aufgaben der Kompetenzstufe III. Die Leistungsunterschiede sind moderat, aber stabil.

Tab. 3: Lösungsraten der einzelnen Kompetenzstufen des Geometrieleistungstests in Abhängigkeit von der Verfügbarkeit formaler Lösungsstrategien bei den Aufgaben zum wissenschaftlichen Denken

	Kompetenzstufe I Einfaches Anwenden von Regeln und elementares Schlussfolgern M = 0,69	Kompetenzstufe II Argumentieren und Begründen (einschrittig) M = 0,56	Kompetenzstufe III Argumentieren und Begründen (mehrschrittig) M = 0,24
keine formale Strategie verfügbar N = 354	0,67	0,54	0,20
formale Strategie in Ansätzen verfügbar N = 132	0,66	0,55	0,21
formale Strategie vollständig verfügbar N = 173	0,74	0,60	0,33

Schülerbeliefs über Mathematik

In der Auswertung des Fragebogens per Faktorenanalyse ergaben sich fünf Faktoren. Die Klassifikation der Items repliziert im Wesentlichen die von Klieme/Ramseier (2001) sowie die von Grigutsch/Raatz/Törner (1998) gefundene Struktur. Als Hauptfaktoren

werden die mathematikbezogenen Beliefs Anwendung, Exaktheit und Prozess identifiziert.

Ein signifikanter Zusammenhang zwischen Schülerleistungen und Schülerbeliefs lässt sich für die Aspekte Anwendungs- und Prozessorientierung nachweisen, er ist allerdings eher gering ($r_\tau = 0,055$; $p < 0,05$). Dieses Ergebnis korrespondiert mit den Ergebnissen der TIMS-Studie, die sich auf Schüler der Abschlussklassen beziehen (Köller u.a. 2000). Die Ergebnisse von TIMSS legen nahe, dass die Beliefs nicht direkt, sondern vor allem über Mediatorvariablen wie Interesse und Motivation auf den Unterricht wirken.

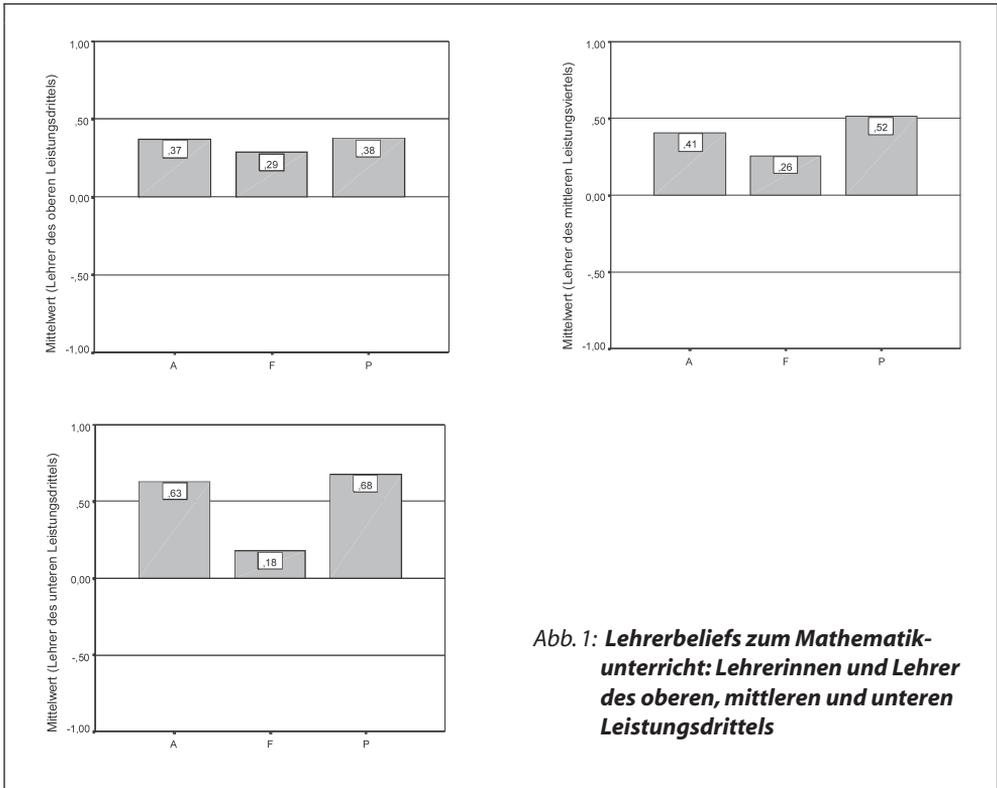
Insgesamt sind die Beliefs zum Mathematikunterricht positiv geprägt. Dominierend ist dabei die Auffassung, dass Mathematik wesentlich durch Exaktheit bzw. Formalismus geprägt ist. Weniger bestimmend ist die Meinung, dass Mathematik durch ihre Anwendungen geprägt ist, noch weniger wird das Prozesshafte in der Mathematik von den Schülerinnen und Schülern gesehen. Diese Auffassungen finden sich in ganz ähnlicher Ausprägung durchgängig in allen untersuchten Klassen. Sie sind unabhängig von Ergebnissen des Leistungstests, d.h. stärkere und schwächere Klassen unterscheiden sich in dieser Beziehung wenig.

Ordnet man den jeweiligen, faktorenanalytisch identifizierten Beliefs Werte zwischen -1 (völlige Ablehnung) und 1 (völlige Zustimmung) zu, dann ergeben sich für die Aspekte Anwendungsbezug (A), Exaktheit (E) und Prozessorientierung (P) im Durchschnitt die Werte $A = 0,34$, $E = 0,55$ und $P = 0,25$. Für Schüler aus Klassen im unteren Leistungsdrittel sind die Werte $A = 0,33$, $E = 0,57$ und $P = 0,25$, für Schüler aus Klassen im mittleren Leistungsdrittel bekommt man $A = 0,33$, $E = 0,57$ und $P = 0,27$ und für Schüler aus Klassen im oberen Leistungsdrittel ergibt sich $A = 0,35$, $E = 0,53$ und $P = 0,21$. Damit besteht sowohl bei den stärkeren als auch bei den schwächeren Klassen ein recht einheitliches Bild von Mathematik, das wesentlich durch Formalismus und Exaktheit geprägt ist.

3.2 *Unterrichtsbezogene Bedingungsfaktoren in Klasse 7/8*

Lehrerbeliefs über Mathematik

Im Wesentlichen wurde für die Befragung der Lehrerinnen und Lehrer der gleiche (jedoch an diese Population angepasste) Fragebogen verwendet, wie er auch der Schülerbefragung zugrunde lag. Bei Betrachtung der mathematikbezogenen Beliefs der Lehrer der 7./8. Klassen ergeben sich deutliche Unterschiede zu den Schülerbeliefs, dies betrifft insbesondere die Aspekte Formalismus ($t = 4,77$; $df = 679$; $p < 0,001$) sowie den Prozesscharakter von Mathematik ($t = 3,03$; $df = 679$; $p < 0,01$). Lehrer, die im oberen Drittel, also in den leistungstärkeren Klassen unterrichten, zeigen relativ ähnliche Werte in den drei Bereichen Anwendungsbezug, Exaktheit und Prozessorientierung, nämlich $A = 0,37$, $E = 0,29$ und $P = 0,38$. Für Lehrer des mittleren Leistungsdrittels ergeben sich die Werte $A = 0,41$, $E = 0,26$ und $P = 0,52$, und für Lehrer des unteren Leistungsdrittels ergeben sich $A = 0,63$, $E = 0,18$ und $P = 0,68$. Abbildung 1 zeigt die Unterschiede.



Lehrer der Klassen im mittleren und unteren Leistungsdrittel sehen in stärkerem Maße die Anwendungsorientierung und den Prozessaspekt der Mathematik, wohingegen die Exaktheit von Mathematik eher in den Hintergrund tritt. Demgegenüber vertreten die Lehrer der leistungsstärkeren Klassen ein relativ ausgewogenes Bild von Mathematik, bei dem alle drei Aspekte als bedeutsam gesehen werden. Diese deutlichen Unterschiede zwischen den Gruppen sind wegen der geringen Zahl der beteiligten Lehrer nicht signifikant. Lehrer des unteren Drittels und Lehrer des oberen Drittels unterscheiden sich allerdings auf einem Niveau von $p < 0,05$ hinsichtlich ihrer Meinung zu der Prozesshaftigkeit von Mathematik voneinander ($t = 2,89$; $df = 12$; $p = 0,014$).

Unterrichtsstil

Die Ergebnisse zeigen eine deutliche Tendenz zu einem problemlöseorientierten Unterricht ($M = 0,61$; auf einer Skala, bei der 1 Zustimmung und -1 Ablehnung bedeutet), fragend-entwickelnde Verfahren werden von den meisten Lehrern abgelehnt ($M = -0,25$). Der Unterschied zwischen diesen beiden Positionen ist deutlich zu erkennen. Auch in diesem Zusammenhang wurden die Lehrer, die in den verschiedenen Leistungsgruppen unterrichten, getrennt betrachtet. Die Ergebnisse zeigen keine Unterschiede zwischen den Gruppen. Auch diese Ergebnisse werden zur Zeit durch Unterrichtsbeobachtungen validiert. Erste Auswertungen zeigen allerdings, dass problemlöse-

orientierter Unterricht eher die Ausnahme ist. Vielmehr werden lehrerzentrierte Unterrichtsformen angewendet, bei denen häufig das fragend-entwickelnde Unterrichtsgespräch das Geschehen im Klassenzimmer dominiert.

4. Diskussion

Die Studie zeigt, dass argumentative Kompetenzen bei den Schülerinnen und Schülern der siebten Klasse wenig ausgeprägt sind. Obwohl im Unterricht dieser Klasse die Thematik des Beweisens und Begründens behandelt wird, ist die Lösung entsprechender Aufgaben mit großen Schwierigkeiten verbunden. Weit besser sind zu diesem Zeitpunkt die notwendigen Basisfähigkeiten im Bereich der Geometrie erworben. Einfache mathematische Begriffe und Konzepte werden von einem großen Teil korrekt angewendet, mathematische Fakten können abgerufen werden. Diese Befunde bestätigen bisherige Forschungsergebnisse in verschiedener Hinsicht. So passen sie sich für den Bereich mathematischen Argumentierens in die Ergebnisse der PISA-Studie ein (Klieme/Neubrand/Lüdtke 2001). Dabei konnte insbesondere auch das im Rahmen von TIMSS (Klieme 2000) aufgestellte und im Rahmen von PISA (Klieme u.a. 2001) weiter ausgearbeitete Kompetenzstufenmodell auf diesen Anforderungsbereich übertragen werden. Mathematische Kompetenz im Begründen und Beweisen baut damit zunächst auf einem sicheren Umgang mit (deklarativem) Basiswissen auf, das durch Kenntnisse in der Verwendung geeigneter Arbeitsmethoden ergänzt wird. Andererseits konnten Ergebnisse von Healy und Hoyles (1998) und Reiss u.a. (2001) auf die relativ junge Population von Schülern der siebten Jahrgangsstufe übertragen werden.

In Bezug auf die Methodenkompetenz konnte gezeigt werden, dass es für Schüler der Sekundarstufe I schwieriger ist, nicht korrekte gegenüber korrekt geführten Beweisen zu identifizieren. Diese Ergebnisse ergänzen andere Befunde. Sie replizieren Befunde mit Schülern der Sekundarstufe II (Reiss/Klieme/Heinze 2001) sowie auch solche von Healy/Hoyles (1998) mit Schülern am Ende der Sekundarstufe I. Unklar bleibt, warum korrekte und nicht-korrekte Beweise unterschiedlich beurteilt werden. Hier ist zu vermuten, dass der Charakter mathematischer Beweise im Unterricht nicht deutlich geworden ist. Darüber hinaus kann vermutet werden, dass Beweisen in der Schule vom Lehrer präsentiert und als fehlerfreier Prozess dargestellt wird. Erste Auswertungen der Unterrichtsbeobachtungen machen eine solche Interpretation jedenfalls wahrscheinlich. Entsprechend ist zu vermuten, dass die bei Boero (1999) geschilderten und für den Beweisprozess des Experten unabdingbaren explorativen Phasen im Mathematikunterricht keine wesentliche Rolle spielen.

Das wissenschaftliche Denken ist auch in dieser Altersgruppe eng mit der Plausibilität von Aussagen verknüpft. Hypothetisches Denken und deduktives Schließen auf der Basis von Hypothesen wird nur von wenigen Schülerinnen und Schülern geleistet. Für die mathematische Kompetenz scheinen diese Fähigkeiten eine Rolle zu spielen, die Effekte sind allerdings moderat. Mathematische Kompetenz scheint durch ein Faktorbündel bestimmt zu sein, dessen Zusammenwirken wenig geklärt ist.

Die Rolle der mathematikbezogenen Beliefs kann aus den Daten der vorliegenden Studie nur teilweise diskutiert werden. Auffällig ist, dass die Schüler unabhängig von ihrer persönlichen Kompetenz ein Bild von Mathematik haben, dass durch Exaktheit und Formalismus bestimmt ist. Ganz anders sehen die Lehrer dieser Klassen Mathematik. Auch wenn die kleine Zahl der beteiligten Lehrer der Interpretation enge Grenzen setzt, erstaunt doch der systematische Unterschied zwischen den Gruppen. Er würde nahe legen, dass eine ausgewogene Berücksichtigung der verschiedenen Aspekte mathematischen Arbeitens mit dem Unterrichtserfolg zusammen hängt. Es stellt sich allerdings die Frage, ob die geäußerten Beurteilungen der Mathematik den konkreten Unterricht der Lehrkräfte beeinflussen.

Diese Frage gewinnt auch vor dem Hintergrund der Widersprüche zwischen den konkreten Unterrichtsbeobachtungen, die im Wesentlichen einen lehrerzentrierten, fragend-entwickelnd geführten Unterricht zeigen, und der von den Lehrern geäußerten eindeutigen Bevorzugung eines problem-lösenden Unterrichts vor dem fragend-entwickelnden Unterrichtsgespräch an Bedeutung. Das Ergebnis muss wohl zunächst dahingehend interpretiert werden, dass den Lehrern sehr wohl bewusst ist, welcher Unterrichtsstil in der mathematikdidaktischen Diskussion präferiert wird. Unklar ist, ob ihnen die Diskrepanz zwischen der eigenen Unterrichtsgestaltung und diesem gewünschten Ideal deutlich ist.

Die bedeutende Rolle des Lehrers und seines Unterrichtsstils scheint allerdings auch aus den Ergebnissen dieser Untersuchung ableitbar zu sein. Dafür sprechen nicht nur die deutlichen Leistungsunterschiede zwischen einzelnen Klassen. Dafür spricht auch, dass in den besseren Klassen ganz allgemein das Argumentieren und Begründen eine wichtige Rolle zu spielen scheint. Man kann vermuten, dass es so etwas wie ein diskursiv orientiertes Klassenklima gibt, das einen positiven Einfluss auf die mathematische Leistung hat. Weitere Unterrichtsbeobachtungen und ihre Auswertung im Rahmen der hier vorgestellten Studie sollen zur Klärung dieser Vermutung beitragen.

Literatur

- Baumert, J./Lehmann, R. u.a. (Hrsg.) (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Boero, P. (1999): Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. In: International Newsletter on the Teaching and Learning of Mathematical Proof, S. 7–8.
- Grigutsch, S./Raatz, U./Törner, G. (1998): Mathematische Weltbilder bei Mathematiklehrern. In: Journal für Mathematikdidaktik 19, S. 3–45.
- Harel, G./Sowder, L. (1998): Students' Proof Schemes: Results from Exploratory Studies. In: Schoenfeld, A.H./Kaput, J./Dubinsky, E. (Hrsg.): Research in Collegiate Mathematics Education. Providence, RI: American Mathematical Society, S. 234–283.
- Healy, L./Hoyle, C. (1998): Justifying and proving in school mathematics. Technical Report on the Nationwide Survey. Mathematical Science. London: Institute of Education, University of London.
- Jacobs, J.K./Yoshida, M./Stigler, J.W./Fernandez, C. (1997): Japanese and American teachers' evaluations of mathematics lessons: a new technique for exploring beliefs. In: The Journal of Mathematical Behaviour 16, H. 1, S. 7–24.

- Klieme, E. (2000): Fachleistungen im voruniversitären Mathematik- und Physikunterricht: Theoretische Konzepte, Kompetenzstufen und Unterrichtsschwerpunkte. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): TIMSS – Mathematisch-naturwissenschaftliche Bildung am Ende der Sekundarstufe II. Opladen: Leske & Budrich, S. 57–128.
- Klieme, E./Ramseier, E. (2001): The impact of school context, student background, and instructional practice. Vortrag auf der Tagung der EARLI, Fribourg, Schweiz.
- Klieme, E./Neubrand, M./Lüdtke, O. (2001): Mathematische Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich, S. 139–190.
- Klahr, D./Dunbar, K. (1988): Dual space search during scientific reasoning. In: Cognitive Science 12, S. 1–55.
- Koedinger, K.R. (1998): Conjecturing and argumentation in high-school geometry students. In: Lehrer, R./Chazan, D. (Hrsg.): Designing learning environment for developing understanding of geometry and space. Mahwah, NJ: Earlbaum, S. 319–347.
- Kölller, O./Baumert, J./Neubrand, J. (2000): Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske + Budrich, S. 229–269.
- Neubrand, M./Biehler, R./Blum, W./Cohors-Fresenborg, E./Flade, L./Knoche, N./Lind, D./Löding, W./Müller, G./Wynands, A. (2001): Grundlagen der Ergänzung des internationalen PISA-Mathematik-Tests in der deutschen Zusatzerhebung. In: Zentralblatt für Didaktik der Mathematik 33, H. 2, S. 45–59.
- Pehkonen, E./Törner, G. (1999): Mathematical Beliefs and their Impact on teaching and Learning of Mathematics. Schriftenreihe des Fachbereichs Mathematik. Duisburg: Universität.
- Peterson, P.L./Fennema, E./Carpenter, T.P./Loef, M. (1989): Teachers` pedagogical content beliefs in mathematics. In: Cognition and Instruction 6, H. 1, S. 1–40.
- Reiss, K./Thomas, J. (2000): Wissenschaftliches Denken beim Beweisen in der Geometrie. Ergebnisse einer Studie mit Schülerinnen und Schülern der gymnasialen Oberstufe. In: Mathematica didactica 23, S. 96–112.
- Reiss, K./Klieme, E./Heinze, A. (2001): Prerequisites for the understanding of proofs in the geometry classroom. In: Van den Heuvel-Panhuizen, M. (Hrsg.): Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education. Vol. 4. Utrecht: Utrecht University, S. 97–104.
- Stern, E./Staub, F. (2000): Mathematik lernen und verstehen: Anforderungen an den Unterricht. In: Kahlert, J./Inckermann, E./Speck-Hamdan, A. (Hrsg.): Beiträge zur Schulentwicklung: Grundschule: Sich Lernen leisten. Theorie und Praxis. Neuwied/Kriftel: Luchterhand, S. 90–100.
- Thomas, J./Schillig, S. (1996): Die Entwicklung hypothetisch-deduktiven Denkens im Jugendalter. In: Schumann-Hengsteler, R./Trautner, M. (Hrsg.): Entwicklung im Jugendalter. Göttingen: Hogrefe.
- Thomas, J. (1997): Wissenschaftliches Denken im Jugendalter. Habilitationsschrift. Johannes Gutenberg-Universität Mainz.
- Törner, G. (Hrsg.) (1998): Current State of Research on Mathematical Beliefs VI. Proceedings of the MAVI Workshop. Duisburg: Schriftenreihe des Fachbereichs Mathematik der Universität.

Anschrift der Autoren:

- Prof. Dr. Kristina Reiss, Universität Augsburg, Lehrstuhl Didaktik der Mathematik, 86135 Augsburg.
 Frank Hellmich, Carl von Ossietzky Universität Oldenburg, Fachbereich Mathematik, Postfach 2503, 26111 Oldenburg.
 Prof. Dr. Joachim Thomas, Katholische Universität Eichstätt-Ingolstadt, Lehrstuhl Psychologic V, Ostenstr., Waisenhaus, 85072 Eichstätt.

Ingmar Hosenfeld/Andreas Helmke/Friedrich-Wilhelm Schrader

Diagnostische Kompetenz: Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte¹ in der Unterrichtsstudie SALVE

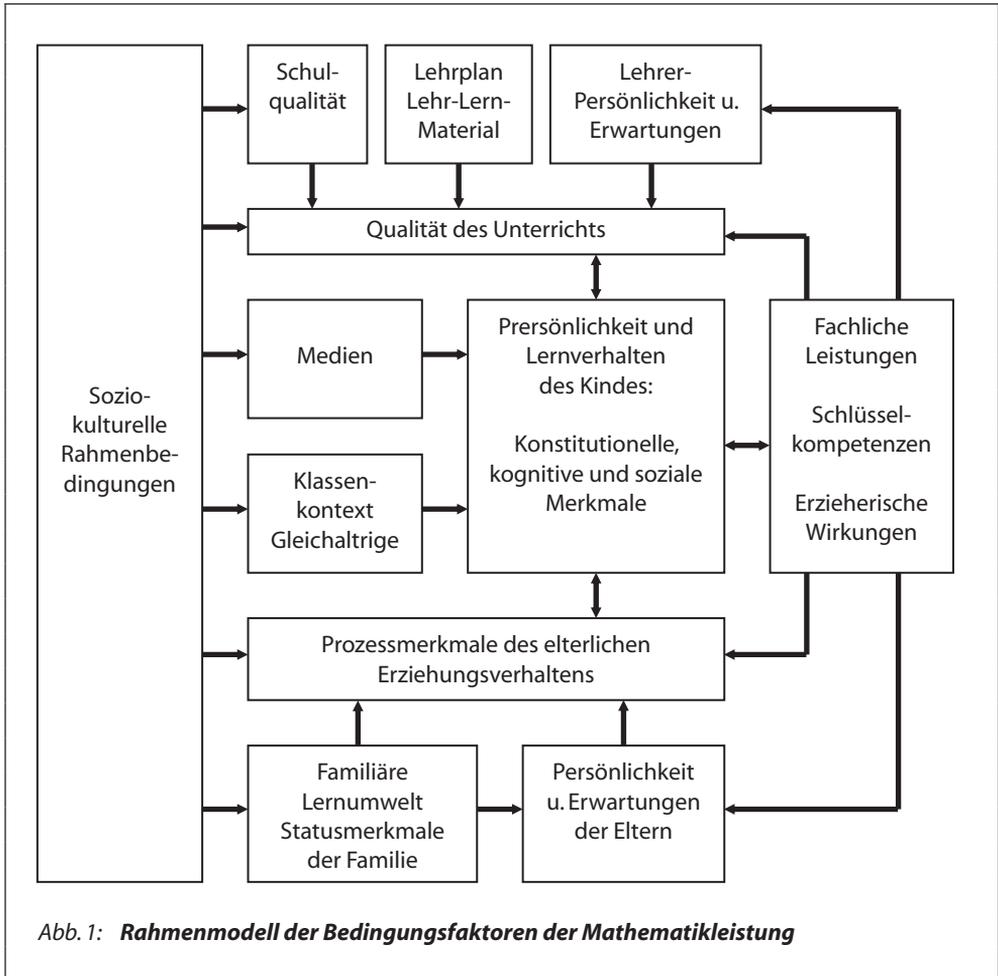
1. Ziele und Fragestellungen

Ziel des Projekts „Systematische Analyse des Lernverhaltens und des Verständnisses in Mathematik: Entwicklungstrends und Fördermöglichkeiten“ (SALVE) ist die Beschreibung und Erklärung der Entwicklung mathematischer Leistungen, des Verständnisses im Fach Mathematik und der lernbezogenen Motivation in Abhängigkeit von der Qualität des Unterrichts, vom Klassenkontext und von den individuellen Lernvoraussetzungen. Gegenstand der Untersuchung sind zentrale Determinanten der schulischen Leistung, wobei den unterrichtlichen Bedingungen ein besonderer Stellenwert zukommt. Zu diesem Themenbereich liegen verschiedene Übersichtsarbeiten vor (vgl. etwa Helmke/Weinert 1997; Helmke/Schrader 2001a, b; Einsiedler 1997). In Abbildung 1 (S. 66) ist eine allgemeine Übersicht über die Bedingungsfaktoren der Leistung im Fach Mathematik dargestellt.

Die Fachleistung einzelner Schüler² sowie ganzer Klassen hängt neben allgemeinen soziokulturellen Rahmenbedingungen vor allem ab (1) vom Schul- und Klassenkontext (z.B. Einzugsgebiet der Schule, Klassengröße, Niveau und Homogenität des Vorwissens), (2) dem familiären Kontext (z.B. Bildungsschicht, Wohnsituation), (3) individuellen Schülermerkmalen (z.B. Intelligenz, Lernmotivation) und (4) der Qualität des Unterrichts (z.B. Strukturiertheit, Verständlichkeit). Als *Kontextmerkmale* bezeichnet man in diesem Zusammenhang solche Faktoren, die die Leistung von Schülern determinieren, aber dem Einfluss der unterrichtenden Lehrkraft entzogen sind (Medley 1982; Hosenfeld u.a. 2002). Diese außerunterrichtlichen Leistungsdeterminanten begrenzen die Möglichkeiten und die Wirksamkeit der Unterrichtsgestaltung und müssen daher berücksichtigt werden, wenn der Einfluss des Unterrichts untersucht werden soll. Ohne eine Einbettung der Leistungsentwicklung und ihrer individuellen und unterrichtlichen Determinanten in den klassenbezogenen und familiären Kontext wäre ein Projekt mit dem Ziel einer Erklärung von Leistungsunterschieden innerhalb und zwischen Klassen wenig aussichtsreich (s. Helmke/Hosenfeld/Schrader 2002).

1 Die Studie wurde gefördert durch Mittel der DFG (Aktenzeichen HE 1873/1-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

2 Aus Platzgründen wird im Rahmen dieses Beitrags nur die männliche Form verwendet, gemeint sind jeweils beide Geschlechter.



Einen besonderen Stellenwert für die Vorhersage und Erklärung der Leistung von Klassen und individuellen Schülern hat der Unterricht, da er der eigentliche Ansatzpunkt für den gezielten und systematischen Wissenserwerb und die Vermittlung von Kompetenzen bei den Schülern ist, und da unterrichtliche Bedingungen im Unterschied zu Kontext- und vielen Schülermerkmalen durch die Lehrkraft prinzipiell veränderbar sind (Ditton 2000; Weinert 2001). Die Unterrichtsforschung hat belegt, dass Unterschiede in der Quantität und Qualität des Unterrichts einen wichtigen Beitrag für die Erklärung von Leistungsunterschieden zwischen Klassen und einzelnen Schülern leisten (vgl. Brophy/Good 1986; Shuell 1996; Einsiedler 1997).

Einen wichtigen Ausgangspunkt für SALVE stellen die Befunde der Münchner Hauptschulstudie des MPI für psychologische Forschung (Classroom Environment Study, CES; Helmke/Schneider/Weinert 1986; Helmke 1992; Helmke/Schrader 1993) zu Bedingungsfaktoren der Mathematikleistung dar. Gegenstand dieses Projekts, das deutschlandweit zu den umfassendsten Projekten zur Unterrichtsqualität gehört, war

die Leistungsentwicklung im Fach Mathematik und die Entwicklung motivationaler Merkmale über den Zeitraum vom Beginn der fünften bis zum Ende der sechsten Klassenstufe. Zu den wichtigsten Ergebnissen der Münchner Studie gehören der Nachweis der Wirksamkeit eines lehrergesteuerten, gut strukturierten und unterstützenden Unterrichts (Helmke/Schrader 1993) sowie die Katalysatorfunktion der diagnostischen Kompetenz von Lehrkräften (Schrader 1989, 2001).

In SALVE soll das Zusammenspiel zwischen individuellen Bedingungsmerkmalen, der Unterrichtsqualität und dem Kontext im Hinblick auf Verständnis und Lernmotivation als den beiden zentralen Zielkriterien analysiert werden. Neben dem Klassenkontext und der Schulform wird dabei der häuslichen Lernumwelt besondere Aufmerksamkeit geschenkt. Ein erstes Ziel von SALVE ist es, grundlegende Befunde der Münchner Studie zu replizieren. Weiterführende Fragen beziehen sich vor allem auf die folgenden Punkte:

- 1) Innerhalb der Mathematik gilt unser Interesse primär dem *Verstehen*, also eben jenem Aspekt, der sich bei TIMSS als eine besondere Schwachstelle im deutschen Mathematikunterricht herausgestellt hat (vgl. Baumert 1999). Letztlich geht es um den Erwerb intelligenten Wissens im Sinne von Weinert (Weinert 1998, 1999), also den Aufbau einer gut organisierten und flexibel nutzbaren Wissensbasis, die eine effektive Auseinandersetzung mit Problemstellungen verschiedener Art erlaubt.
- 2) Neben dem Verständnis ist die *Lernmotivation* das zweite zentrale Zielkriterium. Von Interesse ist insbesondere auch die Frage der Vereinbarkeit der beiden Zielkriterien „Verständnis“ und „Lernmotivation“.
- 3) Ein weiteres Ziel von SALVE besteht darin, den *Unterricht* so umfassend und detailliert zu untersuchen, dass die für Verständnis und Lernmotivation wichtigen Unterrichtsmerkmale identifiziert werden können. Dafür werden Videoaufzeichnungen des Unterrichts verwendet. Diese bieten den Vorteil, den Unterricht unter unterschiedlichen theoretischen Perspektiven und Fragestellungen analysieren zu können. Neben den in TIMSS verwendeten und vom Kieler Projekt (s. Prenzel u.a., in diesem Band) weiterentwickelten Kategorien zur Beschreibung des Unterrichts werden vor allem die bereits in der Münchner Studie und in SCHOLASTIK verwendeten Ratings (Weinert/Helmke 1997) eingesetzt, die für SALVE noch einmal überarbeitet und verfeinert wurden. Darüber hinaus sollen unter anderem Unterrichtsausschnitte besonders in den Blick genommen werden, die für Verständnisleistungen wichtig sind, etwa der Umgang mit Fehlern. Schließlich soll mit den Videoaufzeichnungen versucht werden, bei Auswertungen Gesichtspunkte der allgemeinen Lehr-Lernforschung stärker mit fachdidaktischen Gesichtspunkten zu verknüpfen. Zu diesem Zweck wurde in Zusammenarbeit mit dem Projekt Oldenburg/Gießen (s. Reiss u.a., in diesem Band) ein fachdidaktisch orientiertes Kategoriensystem entwickelt.
- 4) Ein vierter Schwerpunkt ist die *Diagnosekompetenz*: Für eine effektive Unterrichtsgestaltung ist es von zentraler Bedeutung, dass eine möglichst gute Passung zwischen dem Unterrichtsangebot und den Lernvoraussetzungen der Schüler (aktueller Wis-

sensstand, Aufmerksamkeits- und Verständnisniveau usw.) vorliegt. Dafür, dass eine solche diagnostische Sensibilität für den Unterrichtserfolg bedeutsam ist, gibt es erste Belege (s. Schrader 1989, 2001; Weinert/Lingelbach 1995).

2. Fragestellungen

Grundlegendes Ziel des Projekts SALVE ist die Analyse von Bedingungsfaktoren, die für das Verständnis und die Lernmotivation im Fach Mathematik bedeutsam sind. Eine solche Bedingungsanalyse kann zum gegenwärtigen Zeitpunkt noch nicht geleistet werden, da die abschließenden Erhebungen (insbesondere die Messung der am Ende des Unterrichts vorhandenen Ausprägungen beider Zielkriterien) noch nicht erfolgt sind. Ziel dieses Beitrags ist daher zunächst eine deskriptive Darstellung der Ausprägung einiger grundlegender Bedingungsfaktoren. Dies ist auf Seiten der individuellen Lernvoraussetzungen zunächst das *Vorwissen*, das zu den für den Lernerfolg der Schüler wichtigsten Lernvoraussetzungen gehört. Das Vorwissen stellt gleichzeitig auch eine wichtige Ausgangsbedingung dar, die für eine optimale Unterrichtsgestaltung zu berücksichtigen ist. Neben diesem kognitiven Schülermerkmal sind auch relativ überdauernde motivationale Konstrukte, z.B. Leistungsangst und Lernfreude, maßgeblich für die Fähigkeit und Bereitschaft der Schüler, das Unterrichtsangebot zu nutzen. In diesem Beitrag wird exemplarisch die *Freude am Fach Mathematik*, die die tätigkeitsspezifischen Anreize einer Beschäftigung mit diesem Fach umfasst, betrachtet. Schließlich stellen im Unterricht feststellbare Schülermerkmale wie *Aufmerksamkeit*, *Verständnis* und *subjektive Schwierigkeit* Indikatoren für die Angemessenheit und die tatsächliche Nutzung des Unterrichtsangebots dar. Die Qualität der Unterrichtsgestaltung hängt nicht zuletzt davon ab, wie gut Lehrkräfte über diese Schülermerkmale orientiert und damit in der Lage sind, ihren Unterricht den aktuellen Erfordernissen anzupassen. Deshalb soll gleichzeitig auch die Einschätzung dieser Schülermerkmale durch die Lehrkräfte untersucht und die Übereinstimmung dieser Einschätzungen mit den Merkmalen selbst analysiert werden. Damit ergibt sich ein erster Zugang zur Frage der diagnostischen Sensibilität, deren systematische Analyse eines der künftigen Projektziele sein wird.

Daraus ergeben sich im Einzelnen folgende Fragestellungen, zu denen erste Ergebnisse dargestellt werden sollen:

- Wie sehen die Eingangsleistungen der Schüler aus? Welche Unterschiede gibt es zwischen Klassen?
- Wie sind die motivationalen Eingangsbedingungen (Freude an Mathematik) der Klassen beschaffen?
- Wie schätzen Schüler unterrichts- und lernrelevante Merkmale (Interesse, Aufmerksamkeit, Verständnis, Unterforderung) ein?
- Wie gut sind die Lehrkräfte über diese verschiedenen Aspekte orientiert, d.h. wie genau stimmen ihre Einschätzungen mit den Schülerangaben bzw. -leistungen überein?

3. Methode

3.1 Allgemeiner Rahmen und Ablauf der Untersuchung

Um die verschiedenen Einflussfaktoren adäquat abzubilden, werden Daten unterschiedlichen Typs und aus verschiedenen Quellen kombiniert: Kern der Untersuchung ist die längsschnittliche Erfassung der Mathematikleistungen und motivationaler Merkmale der Schüler sowie eine Beobachtung des Unterrichts. Zusätzlich zur Testung und Befragung der Schüler fanden auch schriftliche Befragungen der Mathematiklehrkräfte und der Eltern statt. Der Mathematikunterricht wurde videografisch aufgezeichnet, so dass zusätzlich zu den Informationen aus Sicht der „Beteiligten“ auch Beurteilungen des unterrichtlichen Geschehens aus der Sicht geschulter Experten verfügbar sind. Die Datenerhebung begann im Frühjahr 2001 mit der 1-stündigen Videografie einer Einführungsstunde in 30 Klassen der 5. Jahrgangsstufe (alle Schularten), die ein Interview mit der Lehrkraft und eine schriftliche Kurzbefragung der Schüler einschloss. Ende des Schuljahres folgten eine Befragung der Eltern und Lehrkräfte sowie eine Leistungsmessung und Befragung der Schüler (2 Schulstunden). Im Herbst 2001 wurden in 13 dieser Klassen die ersten 3 Stunden der Einführung des Bruchzahlbegriffs aufgezeichnet. Zuvor bearbeiteten die Schüler einen Bruch-Vorwissenstest und am Ende der Aufzeichnung einen erweiterten Schüler-Kurzfragebogen. Die Datenerhebungen für den zweiten Messzeitpunkt (2 Schulstunden), die erneut einen Mathematiktest sowie Fragebögen für Schüler, Eltern und Lehrkräfte umfassen, werden im Juni 2002 stattfinden.

3.2 Stichprobe

Die Stichprobe der Studie umfasst 30 Klassen der 5. Klassenstufe mit insgesamt 654 Schülern (Mädchenanteil: 50.2%). Einzelheiten dazu sind Tabelle 1 zu entnehmen.

Schulart	Klassen	Schüler/innen	Lehrkräfte	% weiblich
Hauptschule (HS)	9	161	9	55.6
Realschule (RS)	5	135	5	60.0
Gymnasium (GY)	9	251	7	42.9
Integrierte Systeme* (IS)	7	107	6	66.7
Gesamt	30	654	27	55.6

* Als Integrierte Systeme sind hier alle in Rheinland-Pfalz existierenden bildungsgangübergreifende Organisationsformen zusammengefasst. Es sind dies: Duale Oberschulen (DOS), Regionale Schulen (RegS), Integrierte Gesamtschulen (IGS) sowie gemeinsame Orientierungsstufen Kooperativer Gesamtschulen (GOS).

3.3 Instrumente

Mathematiktest

Der Mathematiktest (Test 1) umfasst 2 Testheftversionen à 40 Items, die über 21 gemeinsame Items verankert sind. Die Aufgaben entstammen zu etwa gleichen Teilen der CES (Weinert/Helmke 1984) und dem QuaSUM-5-Test³ (Lehmann u.a. 2000). Zur Bearbeitung der 40 Aufgaben standen 40 Minuten zur Verfügung. Die gemeinsame Skalierung der 59 Items beider Testhefte erfolgte unter Zugrundelegung des dichotomen Rasch-Modells. Die Skalenmetrik wurde wie folgt festgelegt: $M = 500$, $s = 100$. Die auf der Grundlage der Rasch-Skalierung berechnete Reliabilität des Tests beträgt $R = .98$ (zur Berechnung s. Rost 1996, S. 354).

Schülerfragebogen

Der Schülerfragebogen (insgesamt 98 Items) umfasst neben demographischen Angaben und Noten Items zu folgenden drei zentralen Bereichen: Lernmotivation für Mathematik, häusliches Umfeld und Beurteilung des Mathematikunterrichts. Für die vorliegenden Analysen wurde eine Skala zur *Freude an Mathematik* ($\alpha = .75$) herangezogen, deren 3 Items dem Potsdamer-Motivations-Inventar (PMI; vgl. Rheinberg/Wendland/Zarse 2000) und der CES (Weinert/Helmke 1984) entstammen und je 5 Antwortalternativen von „trifft zu“ (5) bis „trifft nicht zu“ (1) aufwiesen.

Lehrerfragebogen

Der Lehrerfragebogen erfasst neben demographischen Angaben auch Angaben zur untersuchten Klasse, Informationen zur Ausübung von mathematikbezogenen Funktionen, die Unterrichts- und Erziehungsziele, epistemologische Überzeugungen, Aspekte der Unterrichtsgestaltung sowie die *Beurteilung der Bereiche und Aufgaben des Tests durch die Lehrkraft*. Für die vorliegenden Analysen wird nur auf die letztgenannten Daten zurückgegriffen. Zu diesem Zweck wurden die Lehrkräfte gebeten, zu beurteilen, wie viele Schüler ihrer Klasse 12 Aufgaben, die Bestandteil des Mathematiktests sind und verschiedene Inhaltsbereiche repräsentieren, jeweils lösen können.

Lehrerinterview zur videografierten Unterrichtsstunde

Unmittelbar vor Beginn und nach Ende der Videoaufzeichnung wurden die Lehrkräfte zur Planung und zum Verlauf der Unterrichtsstunde interviewt. In der Nachbefragung wurden auch Einschätzungen erbeten, wie viele Schüler jeweils (a) den Stoff der Stunde gut nachvollzogen und alles verstanden haben, (b) wirklich aufmerksam waren, (c) mit der Stunde unterfordert waren, (d) mit der Stunde überfordert waren und (e) die Stunde interessant fanden. Für die Analysen in diesem Beitrag wird auf diese von den Lehrkräften vorgenommenen *Einschätzungen der Schüler* zurückgegriffen.

- 3 QuaSUM: Qualitätsuntersuchungen an Schulen zum Unterricht in Mathematik. Eine im Land Brandenburg durchgeführte repräsentative Untersuchung mit dem Ziel der Erhebung von Lernständen am Ende der Klassenstufe 5 und 9 im Fach Mathematik sowie der Analyse von inner- und außerschulischen Bedingungen zur Erklärung von Lernständen.

Schüler-Kurzfragebogen zur videografierten Unterrichtsstunde

Die Schüler wurden unmittelbar nach Ende der Videoaufzeichnung gebeten, einen kurzen Fragebogen zur Stunde zu bearbeiten. Er enthielt folgende Fragen:

- 1) Was wir in der Stunde heute durchgenommen haben, habe ich ... verstanden. (Antwortmöglichkeiten: sehr gut – gut – einigermaßen – nicht so gut – schlecht – sehr schlecht);
- 2) War Dir klar, worauf der Lehrer/die Lehrerin hinauswollte? (Antwortmöglichkeiten: völlig klar – ziemlich klar – einigermaßen klar – nicht so klar – ziemlich unklar – völlig unklar);
- 3) Heute habe ich in der Mathestunde ... aufgepasst. (Antwortmöglichkeiten: sehr gut – gut – einigermaßen – nicht so gut – schlecht – sehr schlecht);
- 4) Heute war der Matheunterricht für mich ... (Antwortmöglichkeiten: viel zu leicht – etwas zu leicht – genau richtig – eher etwas zu schwer – viel zu schwer);
- 5) Was wir heute in Mathe durchgenommen haben, fand ich ... (Antwortmöglichkeiten: sehr interessant – ziemlich interessant – einigermaßen interessant – nicht so interessant – ziemlich uninteressant – völlig uninteressant).

Im vorliegenden Beitrag wird auf die *Selbsteinschätzungen zu Aufmerksamkeit, Interesse, Verständnis* und *Unterforderung* zurückgegriffen. Dazu wurde jeweils pro Klasse ermittelt, welcher Prozentsatz der Schüler den jeweiligen Aussagen zustimmt, d.h. es wurden jeweils die beiden ersten Antwortalternativen als Zustimmung zusammengefasst. Für *Verständnis* wurden anschließend die Zustimmungsprozentsätze der ersten beiden Items gemittelt.

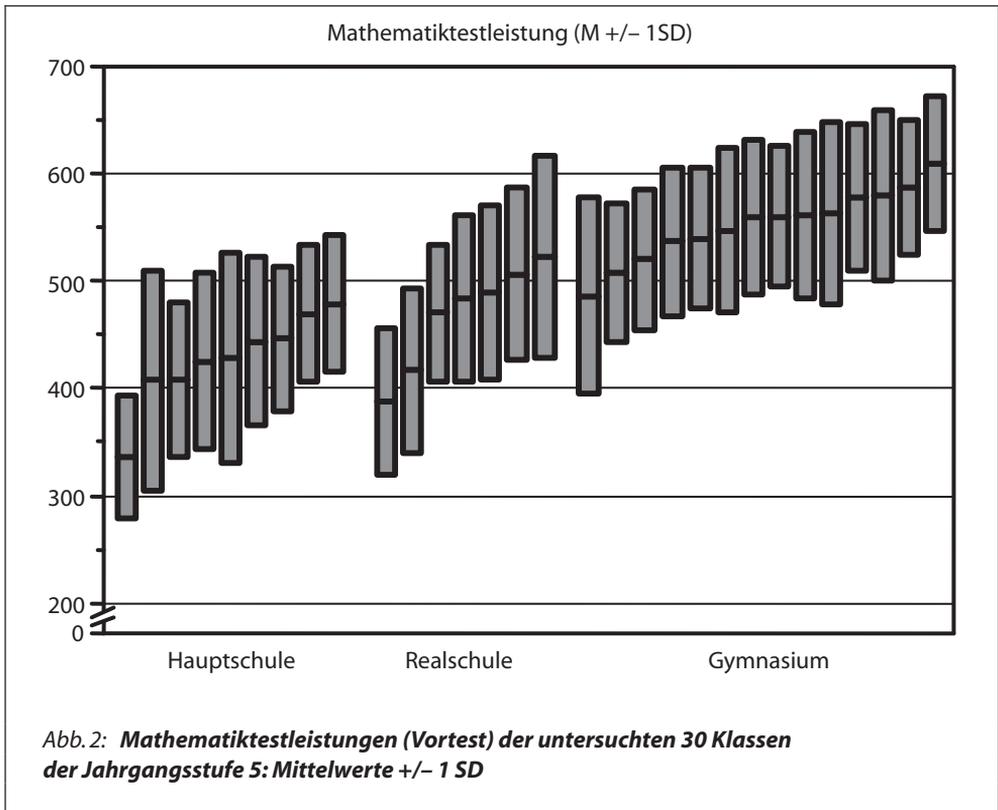
4. Ergebnisse

4.1 Vortestleistungen (Gesamttestwert)

Zur Orientierung über das Vorwissen sind in der folgenden Abbildung 2 (S. 72) die Mittelwerte und Standardabweichungen der Vortestleistungen der Klassen dargestellt. Dabei sind die Klassen nach Schularten⁴ und mittlerer Testleistung sortiert.

Deutlich zu erkennen sind die Mittelwertsunterschiede zwischen den verschiedenen Schularten (Mittelwerte auf Klassenebene: Hauptschule: 427, Realschule: 468, Gymnasium: 552). Zugleich wird jedoch auch deutlich, dass erhebliche Überlappungen der schulartspezifischen Leistungsbereiche existieren. So erzielt beispielsweise die leistungsstärkste Realschulklasse eine höhere Durchschnittsleistung als die leistungsschwächste Gymnasialklasse. Schließlich sind deutliche Unterschiede der klasseninternen Leistungsstreuungen erkennbar. Die klassenspezifischen Standardabweichungen reichen von

4 Die in SALVE untersuchten 7 Klassen integrierter Systeme wurden der Übersichtlichkeit halber der Realschule (2 RegS-Klassen) und dem Gymnasium (5 Klassen GOS) zugeordnet.



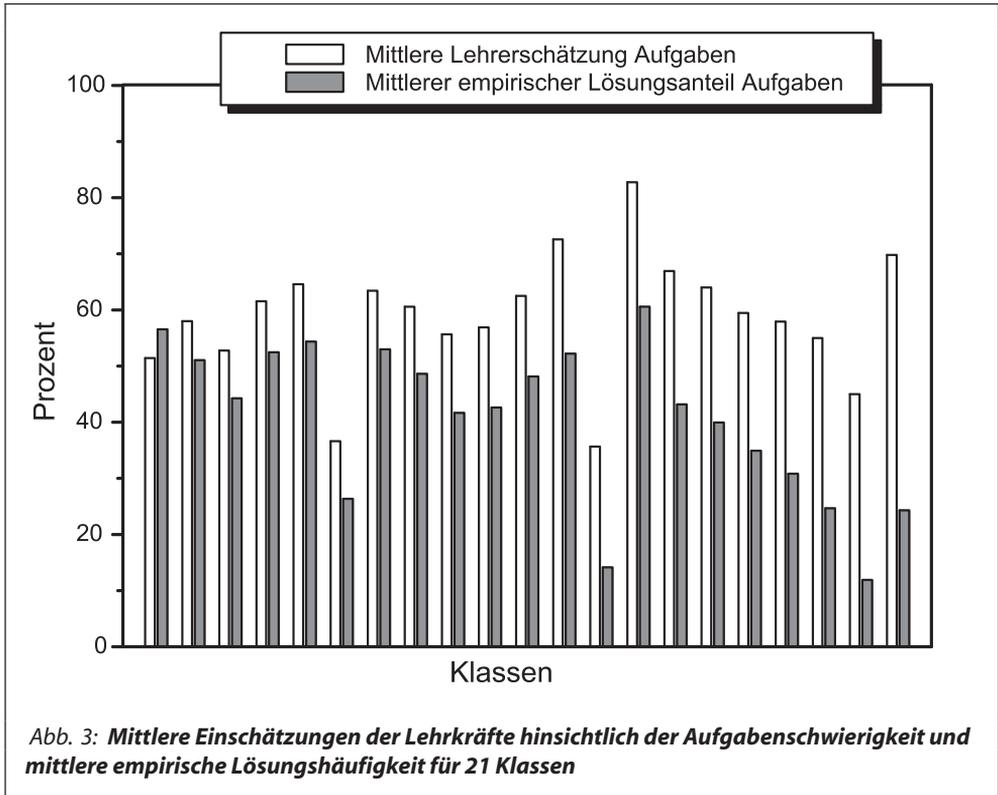
minimal 57 bis maximal 103, d.h. in diesem Fall weist eine Klasse von 19 Schülern ein genauso große Leistungsheterogenität auf, wie die Gesamtstichprobe der Studie.

4.2 Vortestleistungsniveau bei einzelnen Aufgaben

Nach dieser Darstellung der Gesamtttestleistungen werden im Folgenden die Leistungen der Klassen für die 12 ausgewählten Aufgaben betrachtet, zu denen auch die Lehrkräfte befragt wurden⁵. Abbildung 3 stellt den über die 12 Aufgaben gemittelten empirischen Lösungsprozentsätzen der Klassen die entsprechenden Schätzungen der Lehrkräfte gegenüber. Dabei sind die Klassen so angeordnet, dass die Differenzen zwischen Lehrerurteil und empirischem Wert von links nach rechts ansteigen.

Die Einschätzung der Aufgabenschwierigkeit zeigt, dass die Lehrkräfte durchschnittlich von einem Anteil von 58,7% der Schüler ausgehen, welche die zu beurteilenden 12 Aufgaben lösen können. Die Spanne der Einschätzungen reicht dabei von 35,7% bis zu

5 Die folgenden Analysen beruhen auf den Daten der 21 Klassen, für die hierzu vollständige Angaben im Lehrerfragebogen vorliegen.

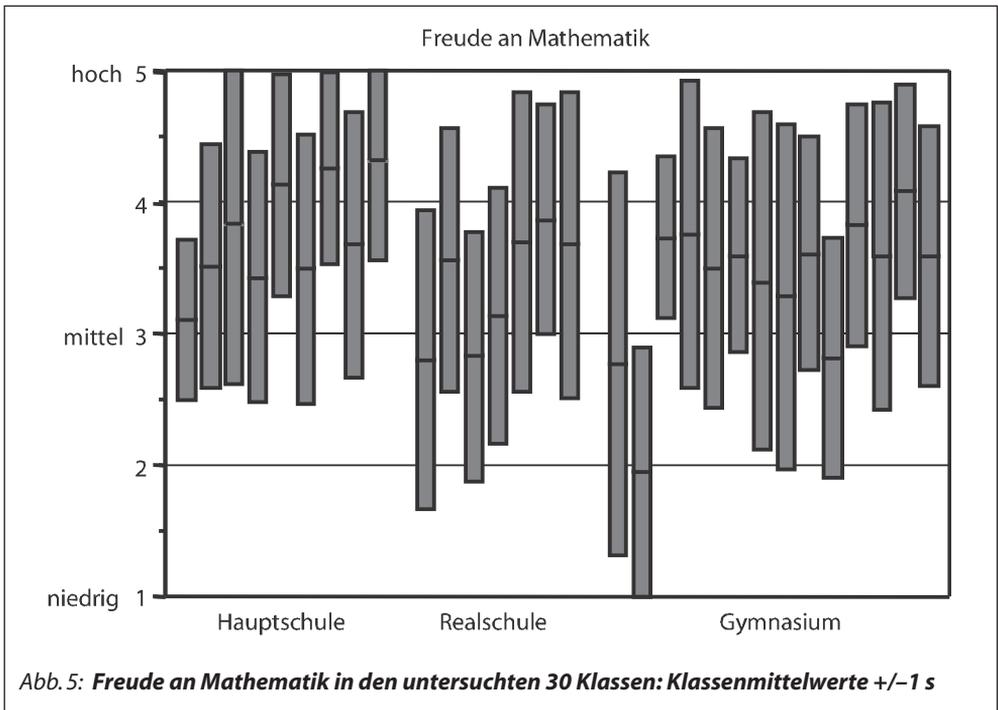
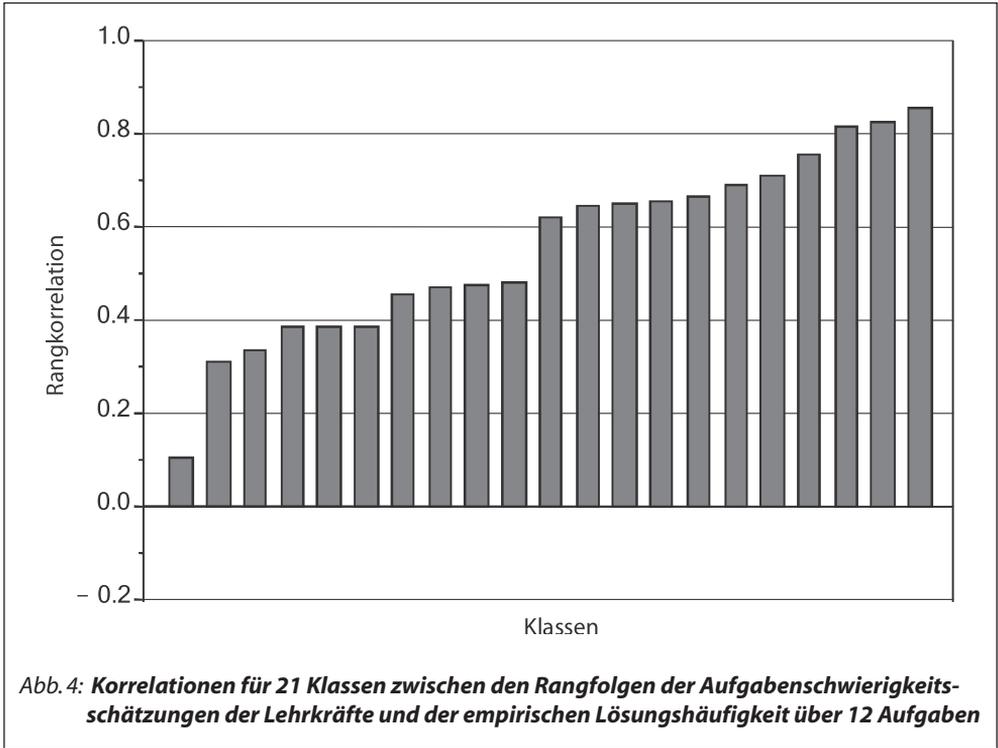


einem Maximum von 82,8%. Die empirischen Lösungshäufigkeiten der Klassen liegen (mit Ausnahme einer Klasse, in der Abbildung ganz links dargestellt) *unter* den Schätzungen der Lehrkräfte. Minimal wurden die 12 Aufgaben von 11,9% der Schüler gelöst, das Maximum liegt bei 60,6%. Im Durchschnitt überschätzen die Lehrkräfte ihre Klassen um 18%, das Maximum liegt hier bei 45,5%.

Rangkorrelationen

Neben Informationen zum Niveau der Aufgabenschwierigkeit lassen sich über die 12 zu beurteilenden Aufgaben hinweg auch die Schwierigkeits-Rangfolgen zur Gewinnung eines Indikators der diagnostischen Sensibilität nutzen. Dazu werden die Rangfolgen der Lehrkraftschätzungen mit den empirischen Rangfolgen innerhalb der jeweiligen Klassen korreliert. Die resultierenden Spearman-Korrelationskoeffizienten der 21 Klassen mit hinreichenden Angaben im Lehrerfragebogen sind in der folgenden Abbildung 4 (S. 74) in aufsteigender Reihenfolge dargestellt.

Die Korrelationen aller 21 Klassen sind positiv, sie reichen von minimal $r = .11$ bis maximal $r = .86$. Im Mittel über die 21 Klassen ergibt sich ein Zusammenhang von $r = .56$ zwischen der von den Lehrkräften geschätzten Schwierigkeitsrangfolge der zwölf Aufgaben mit der Rangfolge der empirischen Lösungshäufigkeiten in der jeweiligen Klasse.



4.3 Freude an Mathematik

Abbildung 5 zeigt die Klassenmittelwerte und Streuungen der Skala „Freude an Mathematik“. Dabei sind die Klassen innerhalb der Schularten aufsteigend von links nach rechts entsprechend der durchschnittlichen Vortestleistung angeordnet, d.h. die Reihenfolge der Klassen ist hier identisch zur Reihenfolge der Abbildung 2.

Im Unterschied zur Vortestleistung sind die Schulartunterschiede in der Freude an Mathematik nur gering (Mittelwerte auf Klassenebene: Hauptschule: 3,8, Realschule: 3,4, Gymnasium: 3,4). Es ist erkennbar, dass innerhalb der Schularten Vortestleistung und Freude an Mathematik der Klassen zusammenhängen: diese Merkmale korrelieren an Realschulen und Gymnasien mit $r = .57$, in Hauptschulen mit $r = .68$. Weiter fällt auf, dass es erhebliche Unterschiede in der Heterogenität der Freude an Mathematik zwischen den Klassen gibt. Die minimale Streuung der Freude an Mathematik innerhalb einer Klasse beträgt $s = 0,61$, die Standardabweichung der Klasse mit der maximalen Heterogenität beträgt mit $s = 1,46$ deutlich mehr als das Doppelte.

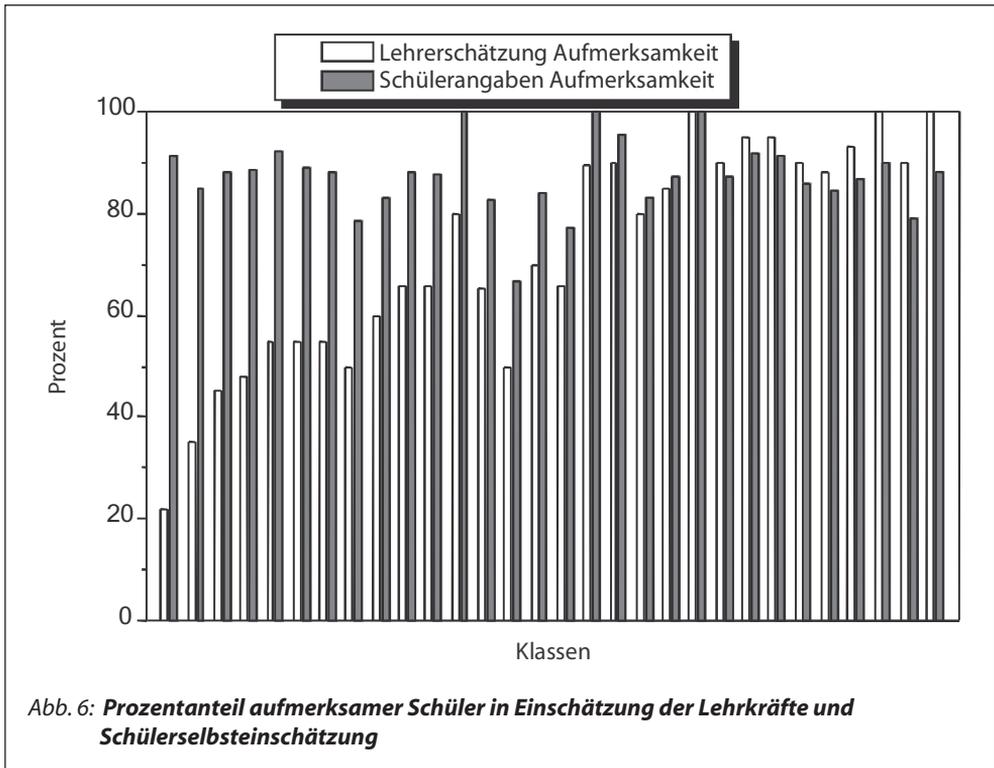
4.4 Einschätzungen zu den videografierten Unterrichtsstunden

Für den Lernerfolg sind neben dem Vorwissen auch zahlreiche weitere, teilweise stark situationsspezifische Merkmale der Schüler bedeutsam. In SALVE liegen dazu unmittelbar nach der Videografie einer Mathematikstunde eingeholte Selbsteinschätzungen der Schüler und Einschätzungen der Lehrkräfte zur Aufmerksamkeit, dem Interesse, der Unter- bzw. Überforderung und dem Verständnis des behandelten Stoffs vor.

Aufmerksamkeit

Die Mehrzahl der Lehrkräfte schätzt, dass mehr als die Hälfte der Schüler der Stunde aufmerksam gefolgt sei; durchschnittlich gehen die Lehrkräfte von 72,5% aufmerksamen Schülern aus. Die minimale Schätzung liegt mit gut einem Fünftel sehr niedrig, die Variabilität zwischen den Lehrkräften ist hoch ($s = 21,5$, Range = 78,3). Die Schüler berichten hingegen zu erheblich größeren Anteilen, in der Stunde sehr gut oder gut aufgepasst zu haben. Das Klassenminimum liegt bei gut zwei Dritteln aufmerksamer Schüler, der Durchschnitt bei 87,5%. Es ist auch in der Abbildung 6 (S. 76) deutlich zu erkennen, dass die Variabilität der Schüleraussagen zwischen den Klassen ($s = 6,9$, Range = 33,3) erheblich geringer ist als im Lehrerurteil.

Zugleich wird deutlich, dass die Mehrzahl der Lehrkräfte die Aufmerksamkeit der Schüler unterschätzt (maximal um knapp 70%), während Überschätzungen der Schüleraufmerksamkeit eher selten sind und mit maximal 11,5% eher gering ausfallen. Für etwa ein Viertel der Klassen (8 von 30) ergeben sich hohe Korrespondenzen zwischen der Einschätzung der Lehrkraft und den Angaben der Schüler (Differenzwerte $\leq 5\%$).



Interesse

Bei der Beurteilung, welcher Prozentsatz der Klasse die Stunde interessant fand, sind die Lehrkräfte zurückhaltender. Im Durchschnitt gehen sie von 63% interessierter Schüler aus, das Minimum der Lehrerschätzung liegt hier bei einem Drittel. Im Gegensatz zur Schätzung der Aufmerksamkeit variieren die Angaben der Lehrkräfte zum Interesse ($s = 19,8$, Range = 67) in vergleichbarem Ausmaß wie die klassenweise gemittelten Schülerangaben ($s = 14,9$, Range = 63).

Auch hier unterschätzt die Mehrzahl der Lehrkräfte ihre Klassen, allerdings in deutlich geringerem Umfang, maximal um 47,7%. Überschätzungen treten etwas seltener auf und fallen mit maximal 34,5% noch etwas moderater aus als die Unterschätzungen. Durchschnittlich unterschätzen die Lehrkräfte den Anteil der Schüler, die die aufgezeichnete Stunde als sehr oder ziemlich interessant bezeichnen, um 5,6%. Das Minimum der Schülerangaben liegt bei 37%. Eine hohe Korrespondenz zwischen Schüler- und Lehrerangaben (Differenz $\leq 5\%$) ergibt sich für 8 Klassen.

Verständnis

Bei der Beurteilung, wie viele Schüler der Klasse den in der Stunde behandelten Stoff verstanden haben, ergibt sich ein ähnliches Muster wie bei der Beurteilung der Aufmerksamkeit. Die Einschätzungen der Lehrkräfte reichen von 20% bis zu 96,2%, im Durchschnitt gehen die Lehrkräfte davon aus, dass knapp 2/3 der Klasse (65,7%) den

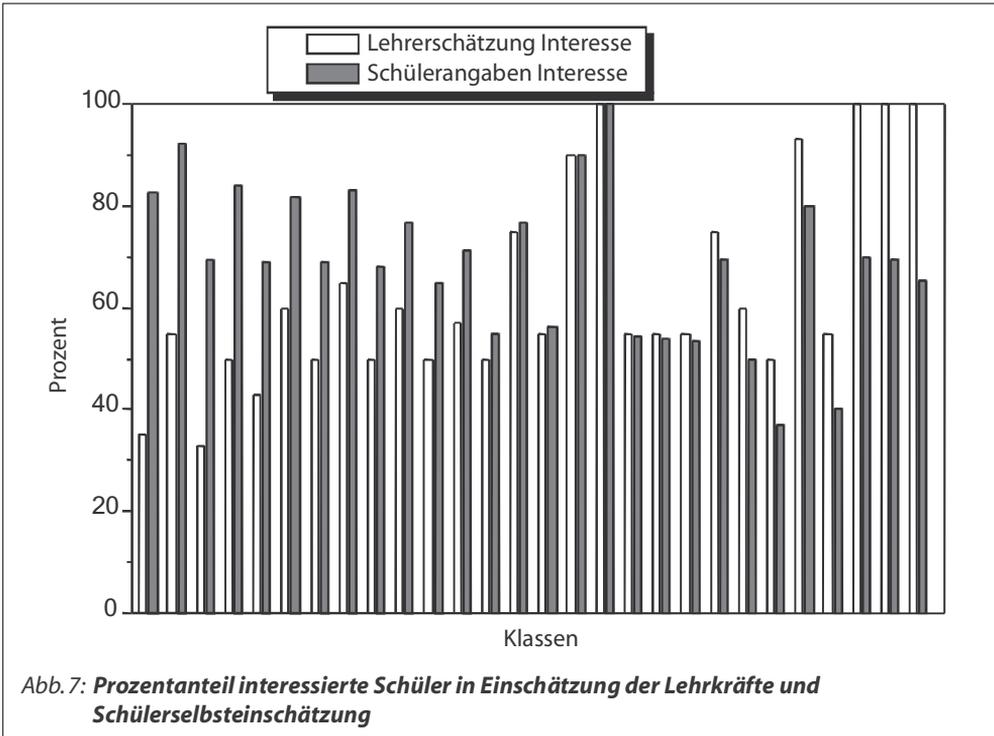


Abb. 7: Prozentanteil interessierte Schüler in Einschätzung der Lehrkräfte und Schülerelbsteinschätzung

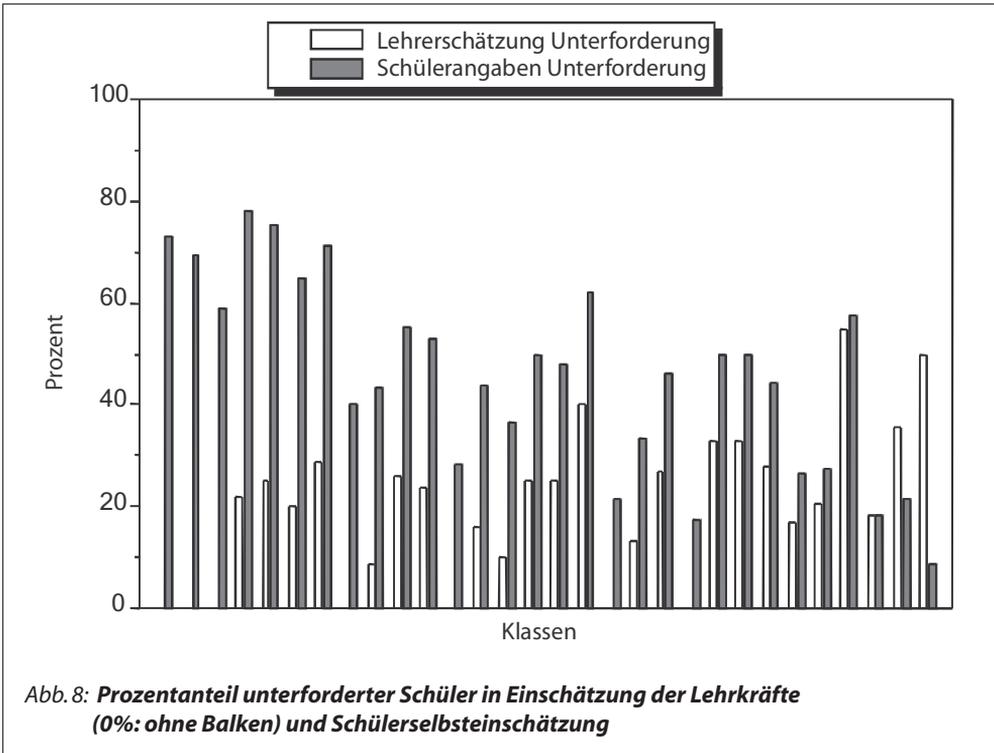


Abb. 8: Prozentanteil unterforderter Schüler in Einschätzung der Lehrkräfte (0%: ohne Balken) und Schülerelbsteinschätzung

behandelten Stoff wirklich verstanden haben. Auch hier liegen die Angaben der Schüler auf höherem Niveau, durchschnittlich geben 80,2% an, den Stoff der Stunde sehr gut oder gut verstanden zu haben und die Zielsetzung der Stunde völlig oder ziemlich klar erkannt zu haben. Das Minimum der Schülerangaben liegt hier bei 54,6%. Die Unterschiede zwischen den Klassen sind dabei im Schülerurteil abermals geringer ($s = 11,1$) als in der Einschätzung der Lehrkräfte ($s = 21,5$). Betrachtet man die direkte Korrespondenz der Angaben, so ergibt sich auch hier, dass ca. $\frac{2}{3}$ der Lehrkraftschätzungen unter den von den Schülern gemachten Angaben liegen, im Extremfall bis zu 74% darunter. Überschätzungen sind seltener und fallen mit maximal 18,5% deutlich moderater aus. Im Durchschnitt ergeben sich Differenzen zwischen Lehrereinschätzung und Schülerangaben von knapp 15%, in 5 der 30 Klassen sind die Differenzwerte $\leq 5\%$.

Unterforderung

Den Anteil der Schüler, die in der aufgezeichneten Mathematikstunde unterfordert waren, schätzen die Lehrkräfte mit durchschnittlich 20% gering ein, das Maximum der Lehrerschätzung liegt hier bei 55%. Die Selbsteinschätzungen der Schüler liegen auch für dieses Merkmal deutlich darüber: Im Durchschnitt befanden 45,8% der Schüler die Stunde als viel oder etwas zu leicht, das Minimum liegt hier bei 8,4%, das Maximum bei immerhin 78,2%. Die direkte Gegenüberstellung der beiden Angaben zeigt, dass die Lehrkräfte den Anteil der unterforderten Schüler um durchschnittlich 25,9% unterschätzen, die Extremwerte liegen bei 73,1% Unter- und 41,6% Überschätzung. Interessanterweise ergibt sich für dieses Merkmal über die Klassen hinweg eine geringere Variation zwischen den Lehrerurteilen ($s = 15$) als zwischen den Schülerangaben ($s = 19$), wobei sich dieser Unterschied aufgrund der geringen Power jedoch statistisch nicht sichern lässt ($F = 1,94$, $p = .169$). Der Anteil von Klassen mit hoher Korrespondenz zwischen den Angaben (Differenzen $\leq 5\%$) liegt bei nur $\frac{1}{15}$.

5. Diskussion

Die Ergebnisse lassen deutliche Unterschiede im *Vorwissen* zwischen den untersuchten Klassen erkennen. Außerdem zeigen sich die zu erwartenden Schulartunterschiede, wobei allerdings erhebliche Überlappungen zwischen den Schularten festzustellen sind. Solche Überlappungen konnten auch bereits in der landesweiten rheinland-pfälzischen Evaluationsstudie MARKUS⁶ an 1862 Klassen der 8. Klassenstufe im Fach Mathematik (Hosenfeld u.a. 2001; Helmke/Hosenfeld/Schrader 2002) belegt werden. Die gefundenen Vorkenntnisunterschiede verweisen noch einmal nachdrücklich auf die Notwendigkeit, die kognitiven Eingangsvoraussetzungen zu kontrollieren, um Aussagen über

6 MARKUS: Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Es handelt sich um eine Vollerhebung in Rheinland-Pfalz zu den Mathematikleistungen in der 8. Jahrgangsstufe und zu den Unterrichtsmerkmalen sowie zu Lernvoraussetzungen und zum persönlichen Hintergrund der Schüler.

die Bedeutung unterschiedlicher Determinanten des Lernerfolgs von Klassen, insbesondere der Unterrichtsqualität, zu machen. Erhebliche Klassenunterschiede zeigen sich auch bei den Lösungshäufigkeiten für einzelne Aufgaben (Aufgabenschwierigkeiten).

Für die *Freude an Mathematik* ergibt sich ein etwas anderes Bild als beim Vorwissen. Auch hier gibt es beträchtliche Klassen-, aber im Mittel nur geringe Schulartunterschiede. Entgegen landläufigen Vorstellungen ist die Freude an Mathematik bei den Hauptschülern nicht geringer als in den anderen Schularten. In der landesweiten MARKUS-Studie hatte sich für die 8. Klassenstufe sogar gezeigt, dass sich Hauptschüler bei motivationalen Merkmalen günstiger einschätzen als die Schüler der anderen Schularten (Schrader/Helmke 2002). Die relativ hohen Korrelationen zwischen Vorwissen und Freude an Mathematik könnten darauf hindeuten, dass beide Zielkriterien weitgehend miteinander verträglich sind: Ein leistungseffektiver Unterricht scheint also nicht unbedingt auf Kosten der Freude am Fach zu gehen. Auffallend ist die große Heterogenität innerhalb der Klassen. Hier wird der Frage nachzugehen sein, ob es systematische Unterschiede zwischen bestimmten Schülergruppen gibt, und wie die Zusammenhänge zwischen Leistung und Freude am Fach innerhalb der Klassen aussehen. Dabei ist es aus praktischer Sicht besonders wichtig zu prüfen, ob es eine relativ homogene Gruppe wenig motivierter Schüler gibt.

Deutlich geringere Klassenunterschiede finden sich bei den *unterrichtsbezogenen Selbsteinschätzungen* der Schüler für Aufmerksamkeit, Interesse und Verständnis, die als mögliche Mediationsvariablen für die Erklärung von Lernerfolgen angesehen werden können: Schüler aller Klassen tendieren dazu, sich als aufmerksam und interessiert einzuschätzen und sich ein hohes Maß an Verständnis zuzuschreiben. Eine interessante Frage ist, warum es bei diesen unterrichtsnah erfassten Merkmalen im Gegensatz zu den habituellen Merkmalen (Leistung, Lernmotivation) so geringe Klassenunterschiede gibt. Zu vermuten ist, dass der Unterricht bei verschiedenen Schülergruppen (z.B. leistungsschwache vs. -starke; ängstliche vs. wenig ängstliche) differenzielle Effekte hat.

Wie sehen die *Lehrereinschätzungen* zu den einzelnen Schülermerkmalen aus? Bei der Aufgabenschwierigkeitsschätzung zeigt sich, dass Lehrkräfte die Klassenleistung im Durchschnitt deutlich überschätzen. Ein Grund für die durchgängig zu findenden Überschätzungen könnte sein, dass Lehrkräfte eher die *Kompetenz* einschätzen – also die prinzipielle Fähigkeit ihrer Schüler, eine Aufgabe zu lösen – und nicht die *Performanz*, also die tatsächliche Lösung in der Testsituation. Dies würde darauf hindeuten, dass die Lehrkräfte leistungsmindernde Faktoren (wie das Vergessen bereits durchgenommenen und früher beherrschten Stoffs; die Begrenzung der Bearbeitungszeit in einem Test; Flüchtigkeitsfehler; mangelnde Anstrengung; Aufregung und Leistungsangst usw.) nicht hinreichend in Rechnung stellen.

Übereinstimmungen im Niveau von Einschätzungen und tatsächlichen Leistungen sind allerdings nur ein sehr grober Indikator für die Diagnosegenauigkeit von Lehrkräften. Für Niveaudiskrepanzen sind vermutlich auch relativ überdauernde Merkmale und Eigenschaften der Lehrpersonen, wie Anspruchsniveau, Erwartungen, Unterrichts- und Erziehungsziele, relevant (vgl. Schrader/Helmke 1987). Ob es sich bei den hier berichteten Diskrepanzen zwischen empirischer Aufgabenschwierigkeit und Lehrerschätzung

um echte Fehleinschätzungen handelt, und ob diese dann auch für die Unterrichtsgestaltung, etwa im Sinne von Über- oder Unterforderungen der Schüler bedeutsam sind, muss in weiterführenden Untersuchungen geklärt werden. Ein besserer Indikator für die diagnostische Kompetenz von Lehrkräften sind klassenspezifische Korrelationen zwischen den Aufgabenschwierigkeiten und den Lehrereinschätzungen: Sie sind ein Maß dafür, inwieweit Lehrkräfte unabhängig vom absoluten Niveau Schwierigkeitsunterschiede zwischen den Aufgaben erkennen. Hier zeigt sich, dass die Lehrerurteile im Mittel recht zutreffend sind, dass es aber deutliche Unterschiede zwischen einzelnen Lehrkräften gibt. Inwieweit sich solche Unterschiede absichern lassen und welche Bedeutung sie für die Unterrichtsgestaltung und den Lernerfolg der Schüler haben, wird Gegenstand weiterführender Untersuchungen sein.

Für die Lehrereinschätzungen zu Aufmerksamkeit, Verständnis und Interesse zeigt sich ein klares Muster: Im Gegensatz zu den Leistungen bei den Aufgaben unterschätzen Lehrkräfte in den meisten Fällen Aufmerksamkeit, Verständnis und in geringerem Maße auch das Interesse ihrer Schüler. Auf Seiten der Schüler könnten selbstwertdienliche Einflüsse mitverantwortlich für die gefundenen Diskrepanzen zwischen Schüler- und Lehrerurteil sein, während Lehrkräfte gerade im Kontext von Beobachtungssituationen möglicherweise eher konservativ urteilen. Darüber hinaus könnte es sein, dass Lehrkräfte Aufmerksamkeit und Verständnis der Klasse bevorzugt an Verhaltensindikatoren festmachen. Schüler können sich dagegen bei ihren Selbsteinschätzungen stärker auf Informationen stützen, die Außenstehenden nicht zugänglich sind, was jedoch nicht zwangsläufig heißt, dass sie zutreffender sind. Gerade beim Verständnis könnten die aus der Metakognitionsforschung bekannten Tendenzen zur Überschätzung wirksam werden: Schüler dieses Alters sind möglicherweise größtenteils noch nicht in der Lage, ihr Verständnis mehr als nur sehr oberflächlich zu überwachen.

Besonders erstaunlich ist jedoch die Unterschätzung des Schülerinteresses durch die Lehrkräfte, da hier selbstwertdienliche Wahrnehmungsverzerrungen auf Seiten der Schüler unplausibel sind. Natürlich ist auch zu berücksichtigen, dass beide Datenquellen unterschiedlich zuverlässig sind: Auf Schülerseite liegen die Angaben ganzer Klassen zugrunde; die Lehrerangaben bestehen dagegen nur aus Einzeleinschätzungen. Auffallend ist, dass die Einschätzungen der Lehrkräfte sehr viel heterogener ausfallen als die Selbsteinschätzungen der Schüler. Inwieweit hierfür Unterschiede in der diagnostischen Kompetenz maßgeblich sind, muss in weiterführenden Analysen geklärt werden.

Literatur

- Baumert, J. (1999): Befunde der internationalen Leistungsvergleiche zum mathematisch-naturwissenschaftlichen Unterricht und fachdidaktische Konsequenzen. In: Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg.): Weiterentwicklung des mathematisch-naturwissenschaftlichen Unterrichts. Stuttgart: Ministerium für Kultus, Jugend und Sport Baden-Württemberg, S. 11–20.
- Brophy, J.E./Good, T.L. (1986): Teacher behavior and student achievement. In: Wittrock, M.C. (Hrsg.): Handbook of research on teaching. London: Macmillan, S. 328–375.

- Ditton, H. (2000): Qualitätskontrolle und -sicherung in Schule und Unterricht. Ein Überblick über den Stand der empirischen Forschung. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule. Weinheim: Beltz, S. 73–92.
- Einsiedler, W. (1997): Unterrichtsqualität und Leistungsentwicklung. Literaturüberblick. In: Weinert, F.E./Helmke, A. (Hrsg.): Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union, S. 225–240.
- Helmke, A. (1992): Selbstvertrauen und schulische Leistungen. Göttingen: Hogrefe.
- Helmke, A./Hosenfeld, I./Schrader, F.-W. (2002): Unterricht, Mathematikleistung und Lernmotivation. In: Helmke, A./Jäger, R.S. (Hrsg.): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Grundlagen und Perspektiven. Landau: Verlag Empirische Pädagogik, S. 413–480.
- 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 2, S. 1–18.
- Helmke, A./Schrader, F.-W. (1993): Was macht erfolgreichen Unterricht aus? Ergebnisse der Münchner Studie. In: Praxis Schule 5–10, S. 11–13.
- Helmke, A./Schrader, F.-W. (2001a): Determinanten der Schulleistung. In: Rost, D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: Beltz, S. 81–90.
- Helmke, A./Schrader, F.-W. (2001b): School achievement, cognitive and motivational determinants. In: Smelser, N.J./Baltes, P.B. (Hrsg.): International encyclopedia of the social and behavioral sciences. Vol. 20. Oxford: Elsevier, S. 13552–13556.
- Helmke, A./Weinert, F.E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Weinert, F.E. (Hrsg.): Enzyklopädie der Psychologie. Psychologie des Unterrichts und der Schule. Pädagogische Psychologie, Vol. 3. Göttingen: Hogrefe, S. 71–176.
- Hosenfeld, I./Helmke, A./Ridder, A./Schrader, F.-W. (2001): Eine mehrbenenanalytische Betrachtung von Schul- und Klasseneffekten. In: Empirische Pädagogik 15, S. 513–534.
- Hosenfeld, I./Helmke, A./Ridder, A./Schrader, F.-W. (2002): Die Rolle des Kontextes. In: Helmke, A./Jäger, R.S. (Hrsg.): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Grundlagen und Perspektiven. Landau: Verlag Empirische Pädagogik, S. 155–256.
- Lehmann, R.H./Peek, R./Gänsfuß, R./Lutkat, S./Mücke, S./Barth, I. (2000): Qualitätsuntersuchungen an Schulen zum Unterricht in Mathematik (QuaSUM). Potsdam: Ministerium für Bildung, Jugend und Sport des Landes Brandenburg (MBSJ).
- Medley, D.M. (1982): Teacher effectiveness. In: Mitzel, H.E./Best, J.H./Rabinowitz, W. (Hrsg.): Encyclopedia of educational research. Vol. 4. New York: Free Press, S. 1894–1903.
- Rheinberg, F./Wendland, M./Zarse, U. (2000): Itemzuordnung des PMI-M. Arbeitsmaterial des Projekts „Veränderung der Lernmotivation in Mathematik und Physik: Eine Komponentenanalyse und der Einfluss elterlicher sowie schulischer Kontextfaktoren“.
- Rost, J. (1996): Lehrbuch Testtheorie, Testkonstruktion. Bern: Huber.
- Schrader, F.-W. (1989): Diagnostische Kompetenzen von Lehrern und ihre Bedeutung für die Gestaltung und Effektivität des Unterrichts. Frankfurt a.M.: Lang.
- Schrader, F.-W. (2001): Diagnostische Kompetenz von Eltern und Lehrern. In: Rost, D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: Beltz, S. 68–71.
- Schrader, F.-W./Helmke, A. (1987): Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. In: Empirische Pädagogik 1, S. 27–52.
- Schrader, F.-W./Helmke, A. (2002): Motivation, Lernen und Leistung. In: Helmke, A./Jäger, R.S. (Hrsg.): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Grundlagen und Perspektiven. Landau: Verlag Empirische Pädagogik, S. 257–324.

- Shuell, T. (1996): Teaching and learning in a classroom context. In: Berliner, D.C./Calfée, R. (Hrsg.): Handbook of Educational Psychology. New York: Macmillan, S. 726–764.
- Weinert, F.E. (1998): Neue Unterrichtskonzepte zwischen gesellschaftlichen Notwendigkeiten, pädagogischen Visionen und psychologischen Möglichkeiten. In: Bayerisches Staatsministerium für Unterricht, Kultur, Wissenschaft und Kunst (Hrsg.): Wissen und Werte für die Welt von morgen – Dokumentation zum Bildungskongress am 29./30. April 1998. München: Bayerisches Staatsministerium für Unterricht, Kultur, Wissenschaft und Kunst, S. 101–125.
- Weinert, F.E. (1999): Bedingungen für mathematisch-naturwissenschaftliche Leistungen in der Schule und die Möglichkeiten ihrer Verbesserung. In: Ministerium für Kultus, Jugend und Sport in Baden Württemberg (Hrsg.): Weiterentwicklung des mathematisch-naturwissenschaftlichen Unterrichts. Stuttgart: Ministerium für Kultus, Jugend und Sport in Baden Württemberg, S. 21–32.
- Weinert, F.E. (2001): Schulleistungen – Leistungen der Schule oder der Schüler. In: F.E. Weinert (Hrsg.) Leistungsmessungen in Schulen. Weinheim: Beltz, S. 73–86.
- Weinert, F.E./Helmke, A. (1984): Zwischenbericht für das DFG-Projekt „Unterrichtsqualität und Leistungszuwachs bei Formen direkter Instruktion im Mathematikunterricht fünfter Hauptschulklassen“. Max-Planck-Institut für psychologische Forschung, München.
- Weinert, F.E./Helmke, A. (Hrsg.) (1997): Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union.
- Weinert, F.E./Lingelbach, H.C. (1995): Teaching expertise: Theoretical conceptualizations, empirical findings, and some consequences for teacher training. In: Hoz, R./Silberstein, M. (Hrsg.): Partnerships of schools and institutions of higher education in teacher development. Beer-Shera, Israel: Ben Gurion University of the Negev Press, S. 293–302.

Anschrift der Autoren:

Dr. Ingmar Hosenfeld, Universität Koblenz-Landau, Campus Landau, Fachbereich Psychologie, Im Fort 7, 76829 Landau in der Pfalz.

Prof. Dr. Andreas Helmke, Universität Koblenz-Landau, Campus Landau, Fachbereich Psychologie, Im Fort 7, 76829 Landau in der Pfalz.

Dr. Friedrich-Wilhelm Schrader, Universität Koblenz-Landau, Campus Landau, Fachbereich Psychologie, Im Fort 7, 76829 Landau in der Pfalz.

Rudolf vom Hofe/Reinhard Pekrun/Michael Kleine/Thomas Götz

Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Konstruktion des Regensburger Mathematikleistungstests für 5. bis 10. Klassen¹

In diesem Beitrag wird die Konstruktion des Regensburger Mathematikleistungstest für 5.–10. Klassen dargestellt, und es werden erste Befunde zur Entwicklung der mathematischen Grundbildung in der Sekundarstufe I berichtet, die mit diesem Test gewonnen wurden. Die vorgestellten Testanalyse- und Entwicklungsbefunde stammen aus der ersten Voruntersuchung zu einer Längsschnittuntersuchung, welche die Entwicklung mathematischer Grundbildung bei deutschen Schülern analysiert (Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA); Projektleiter: R. Pekrun, Universität München; R. vom Hofe, Universität Regensburg; W. Blum, Universität Kassel). Aus den dargestellten Befunden ergeben sich gleichzeitig auch erste Hinweise zu didaktischen Konsequenzen.

1. Zielsetzung und Konzeption der Längsschnittstudie PALMA

Vergleichende Evaluationsstudien zu mathematischen und naturwissenschaftlichen Schülerleistungen belegen, dass solche Leistungen international und intranational eine erhebliche Varianz zeigen. Deutsche Schüler schneiden dabei in der Regel eher mäßig ab. Allerdings sind Evaluationsstudien dieser Art meist auf deskriptiv orientierte, querschnittliche Designs beschränkt. Geleistet wird damit ein Systemmonitoring des schulischen Bildungswesens. Erwartungen von Politik und Öffentlichkeit hingegen, über Indikatoren des Leistungsstandes von Schülern und Schulsystemen hinaus Aufschlüsse zu den Ursachen von Leistungen und zu Maßnahmen der Leistungssteigerung und der Prävention von Defiziten zu erhalten, können solche Studien aufgrund der Beschränkungen ihrer Designs kaum erfüllen (vgl. Pekrun 2002).

Ziel dieses Projekts ist es deshalb, anhand einer prospektiv-längsschnittlichen Erweiterung der OECD-Studie PISA an einer deutschen Schülerkohorte Entwicklungen und Bedingungen von Mathematikleistungen im Laufe der Sekundarstufe I zu analysieren (5.–10. Klassenstufe). Die einbezogene Schülerkohorte wird in der 5. Klassenstufe (2002) erstmals untersucht, in der 9. Klassenstufe werden die Erhebungen mit dem dritten Zyklus der PISA-Erhebungen verschaltet (2006). Mit der inhaltlichen und methodi-

¹ Diese Studie wurde im Rahmen des DFG-Schwerpunktprogramms BIQUA durch Mittel der DFG gefördert.

schen Konzeption des Längsschnitts wird angestrebt, eine Aufklärung der Entwicklungsverläufe und Bedingungen zu erreichen, die den PISA-Mathematikleistungen deutscher Schüler zugrunde liegen.

Analysiert werden (1) die Entwicklungsverläufe von mathematischen Leistungen, (2) Schülervoraussetzungen solcher Leistungen und (3) Kontextbedingungen in Unterricht, Schulklasse und Elternhaus. Dies entspricht der Konzeption der PISA-Erhebungen, wobei sich das vorliegende Projekt auf die Untersuchung des mathematischen Kompetenzerwerbs konzentriert. Bei PISA (Programme for International Student Assessment) handelt es sich um eine international und national vergleichende Schülerleistungsstudie der OECD, die eine Sequenz querschnittlicher Leistungserhebungen vorsieht (zunächst in den Jahren 2000, 2003 und 2006). Erfasst werden Schülerleistungen in den Bereichen Mathematik, Naturwissenschaften, Lesen und fächerübergreifende Kompetenzen, darüber hinaus aber auch Individual- und Kontextvoraussetzungen von Schülerleistungen (vgl. Baumert u.a. 2001).

In den Bereichen Mathematikleistungen, Schülervoraussetzungen und Kontexte werden im vorliegenden Projekt jeweils spezifische Schwerpunkte gesetzt, die mit den Planungen für die PISA-Erhebungen der nächsten Zyklen abgestimmt sind. Im Bereich mathematischer Leistungen liegt ein Schwerpunkt auf der Analyse von Kompetenzen, die auf den mathematischen *Grundvorstellungen* von Schülern beruhen (s.u. 2.). Die Erhebungen zu Schülervoraussetzungen beziehen sich auf Variablen des selbstregulierten Lernens von Schülern im Fach Mathematik, wobei der Erhebung der *Mathematikemotionen* von Schülern ein wesentlicher Stellenwert zukommt. Hintergrund ist, dass zu solchen Emotionen kaum etwas bekannt ist, obwohl ihnen für Motivation, Problemlösen und Leistungen im Fach Mathematik ein entscheidender Stellenwert zukommen dürfte (belegt ist dies bisher nur für Mathematikangst; vgl. Götz 2002; Ma 1999; und allgemein zur Leistungsbedeutung von Emotionen Aspinwall 1998; Pekrun 1992, 1998, 2000; Pekrun u.a. im Druck).

Im Bereich der Kontexte schließlich ist zu konstatieren, dass sich viele Studien der Lehr-Lernforschung auf instruktionale Variablen konzentriert haben, während die Rolle sozialer Kontexte vernachlässigt wurde (vgl. Pekrun 1997, 2001; Wild 1999). Vorgesehen ist deshalb, in Kooperation mit den Projekten von E. Wild (vgl. Beitrag in diesem Band) über die Kontexte *Mathematikunterricht* und *Schulklasse* hinaus auch das *Elternhaus* und den Umgang der Eltern mit den schulischen Leistungsanforderungen im Fach Mathematik zu berücksichtigen.

2. Mathematische Grundbildung und die Ziele des Regensburger Mathematikleistungstests

PISA folgt einer Konzeption mathematischer Grundbildung, die sich im Sinne von „mathematical literacy“ vor allem auf Kompetenzen zu einem realitätsorientierten, flexibelmodellierenden Umgang mit mathematischen Problemen konzentriert. Für solche Kompetenzen sind aus mathematikdidaktischer Perspektive problemlösebezogene men-

tale Repräsentationen mathematischer Inhalte zentral, die häufig als mathematische „Grundvorstellungen“ bezeichnet werden. Es mangelt bisher an Instrumenten zur Erfassung dieser Kompetenzen und an Untersuchungen zu ihrer Entwicklung im Laufe der Schulzeit.

Im Folgenden wird zunächst das an PISA anknüpfende Konzept von mathematischer Grundbildung beschrieben, auf dem der Regensburger Mathematikleistungstest basiert, wobei insbesondere die Bedeutung von Modellierungsaktivitäten und der für solche Aktivitäten erforderlichen Grundvorstellungen erläutert werden. Anschließend werden die Ziele der Konstruktion des Regensburger Mathematikleistungstests dargestellt.

2.1 Mathematische Grundbildung

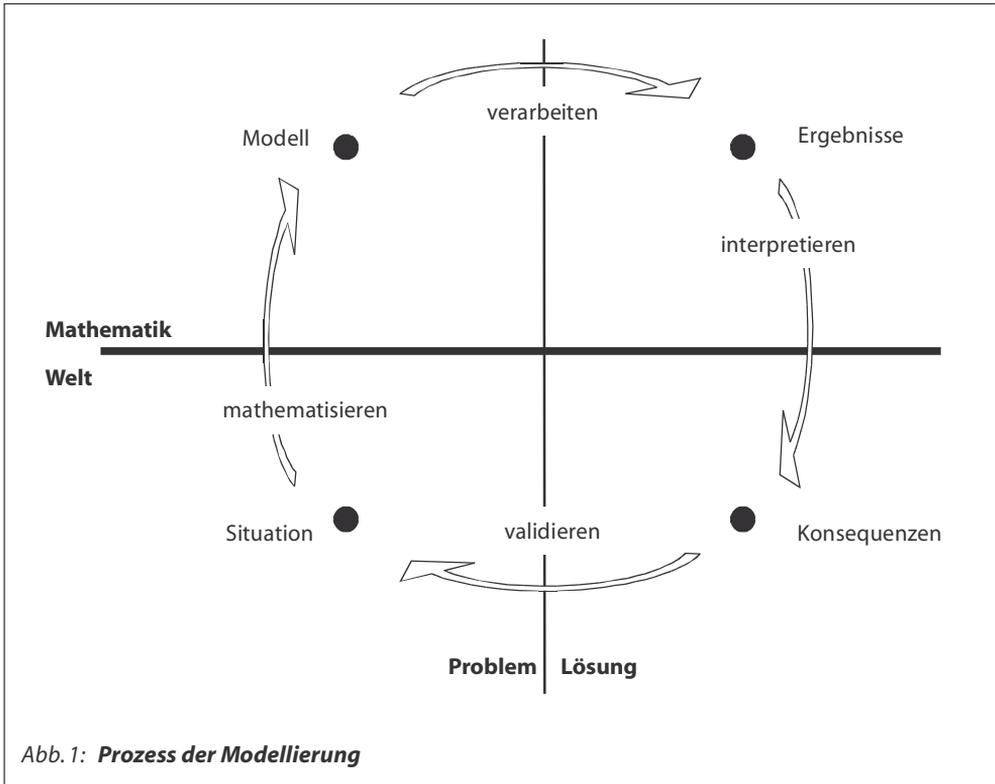
Auffassungen zu mathematischer Grundbildung haben sich in den letzten zwei Jahrzehnten unter den Einflüssen konstruktivistischer Lerntheorien einerseits und der weltweit wirksamen Operationalisierungen durch internationale Schulleistungsstudien andererseits zunehmend konkretisiert und eine heute international weitgehend akzeptierte Form angenommen, die sich im PISA-Konzept der „mathematical literacy“ widerspiegelt. Darin werden die Ansätze von TIMSS fortgesetzt, mathematische Fähigkeiten nicht über Formelanwendungen oder technische Rechenverfahren zu erfassen, sondern über die Rolle, die Mathematik als Werkzeug zur Modellierung und geistigen Gestaltung der Realität zukommt. Dementsprechend wird mathematische Grundbildung als Fähigkeit einer Person definiert, „die Rolle, die Mathematik in der Welt spielt, zu erkennen und zu verstehen, fundierte mathematische Urteile abzugeben und sich auf eine Weise mit der Mathematik zu befassen, die den Anforderungen des gegenwärtigen und künftigen Lebens einer Person als eines konstruktiven, engagierten und reflektierten Bürgers entspricht“ (Klieme/Neubrand/Lüdtke 2001, S. 143).

Diese Auffassung basiert wesentlich auf dem Lebenswerk des niederländischen Mathematikers und Didaktikers Hans Freudenthal, der zu denjenigen zählte, die schon lange vor TIMSS und PISA die Dominanz von kalkülhaftem Rechnen und schematischen Verfahren im Mathematikunterricht kritisierten (vgl. insbesondere Freudenthal 1977, 1983). Damit Lernende etwas von der Kraft und dem Wesen vom Mathematik als gedanklichem Werkzeug erfassen können, müsse sich Lehren und Lernen an der Phänomenologie mathematischer Begriffe orientieren und nicht etwa an vorgefertigten Sätzen oder Definitionen. Freudenthals Ansatz lässt sich beschreiben als ein Plädoyer für einen *genetischen Mathematikunterricht mit eigenständigem und aktiv-entdeckendem Lernen in inner- und außermathematischen Problemkontexten*.

Diese Haltung prägt nicht nur die PISA-Konzeption im Bereich Mathematik, sie entspricht darüber hinaus einem breiten Konsens der zeitgenössischen internationalen Mathematikdidaktik (siehe etwa Schubring 1978; Wagenschein 1983; Brousseau 1983; Sierpinska 1992 zum genetisch-epistemologischen Ansatz sowie Blum 1996; de Lange 1996; Usiskin 1991 zum anwendungs- bzw. realitätsorientierten Ansatz).

2.2 Der Prozess des Modellierens

Der Umgang mit Mathematik lässt sich aus dieser Sichtweise im Wesentlichen als mathematische Modellbildung beschreiben, wie sie in Abbildung 1 in der Terminologie von Schupp (1988) bildlich dargestellt wird. Sie zeigt zum einen die verschiedenen Schritte des Modellierungsprozesses auf, zum anderen wird dessen zyklische Struktur deutlich:



Zunächst wird ein Problem aus der realen Welt mathematisiert, d. h. es wird ein mathematischer Begriff oder ein Verfahren gesucht, um die Sachsituation auf der mathematischen Ebene darzustellen. Als Nächstes werden innerhalb der Mathematik Ergebnisse ermittelt, die dann im Hinblick auf die Sachsituation interpretiert werden. Und schließlich muss geprüft werden, ob die aus dem mathematischen Modell abgeleiteten Konsequenzen tatsächlich für die Lösung des Sachproblems geeignet sind oder ob ein neuer Durchlauf in diesem Zyklus – möglicherweise mit einem geänderten mathematischen Modell – erforderlich ist.

Zentrale kognitive Tätigkeiten beziehen sich hierbei auf das *Übersetzen zwischen Realität und Mathematik*, wenn beispielsweise zu einer Sachsituation eine angemessene Mathematisierung gefunden werden muss oder wenn ein mathematisches Ergebnis wie-

der im Hinblick auf die Sachsituation interpretiert werden soll. Hierfür braucht man Vorstellungen davon, welche mathematischen Inhalte oder Verfahren zu einer bestimmten Sachsituation passen könnten bzw. umgekehrt, welche Situationen sich mit bestimmten mathematischen Inhalten modellieren lassen.

Dies fängt bereits bei sehr einfachen Aufgabenkontexten an: Wer z.B. mit dem Verfahren der *Subtraktion* nicht die Vorstellung des *Abtrennens* oder *Wegnehmens* verbindet, kann selbst einfachste Textaufgaben zur Subtraktion nicht erfassen; wer mit dem Wachstumsverhalten von *Exponentialfunktionen* nicht die Vorstellung des *prozentualen Wachstums* verbindet, kann nicht entscheiden, warum für die Modellierung eines biologischen Wachstumsprozesses eher eine Exponentialfunktion in Frage kommt, und nicht etwa eine lineare oder quadratische Funktion.

2.3 Mathematische Grundvorstellungen

Wichtig für die Vermittlung zwischen Mathematik und Realität ist daher die Ausbildung tragfähiger *mentaler Modelle für mathematische Begriffe* – wie sie Freudenthal nennt – (vgl. Freudenthal 1983) oder kurz: die Ausbildung von *Grundvorstellungen* mathematischer Begriffe und Verfahren (vgl. vom Hofe 1995).

In Analysen zu Ausprägungen solcher Grundvorstellungen und ihrer Bedeutung für die Entwicklung mathematischer Grundbildung liegt ein zentrales Forschungsinteresse unseres Projekts. Die bisher vorliegenden Arbeiten hierzu sind vor allem didaktisch-konstruktiver und präskriptiver Art (z.B. Blum/Törner 1983; Bender 1991; Maier 1990; vom Hofe 1992, 1995, 1998a; Wiegand 1998, 2000; Malle im Druck). Ferner liegen qualitative Einzelfallstudien und Feldstudien vor (vgl. insbesondere Fischbein 1983, 1989, 1990 bzw. vom Hofe 1998b, 1999, im Druck). Hingegen mangelt es an Instrumenten zur quantitativen Erfassung der Ausprägung solcher Vorstellungen und ihrer Einflüsse auf die Entwicklung mathematischer Fähigkeiten im Laufe der Schulzeit.

Als wesentlicher Schritt auf dem Wege zu einer quantitativen Erfassung der Rolle von Grundvorstellungen können Arbeiten innerhalb der deutschen PISA-Experten-gruppe Mathematik angesehen werden, in denen der Grundvorstellungsbegriff in den letzten beiden Jahren als aufgabenanalytisches Kriterium weiterentwickelt und ein Klassifizierungssystem zur Analyse der Grundvorstellungskomplexität von Mathematikaufgaben konstruiert wurde. Die Kategorien dieses Systems werden gegenwärtig für Detailanalysen zu den Erhebungen von PISA 2000 verwendet; erste Befunde sind Ende 2002 zu erwarten. Analysiert wird u.a., welche Rolle die Grundvorstellungskomplexität für die Schwierigkeit von Aufgaben spielt (vgl. Blum/vom Hofe im Druck). Allerdings ermöglicht auch dieses Klassifizierungssystem noch keine Erfassung der Ausprägung von Grundvorstellungen bei individuellen Schülern.

2.4 Ziele der Konstruktion des Regensburger Mathematikleistungstests

Eingangs wurde erläutert, dass Fragen nach der Genese mathematischer Grundbildung von Vergleichsstudien wie PISA aufgrund ihres querschnittlich-deskriptiven Untersuchungsdesigns nicht befriedigend beantwortet werden können, da in der Regel nur punktuelle Leistungen, nicht aber Entwicklungen dieser Leistungen über die Schuljahre hinweg erfasst werden.

Der hier zu konstruierende Mathematikleistungstest soll dazu beitragen, dieses Defizit zu beheben, und es ermöglichen, Entwicklungsverläufe mathematischer Fähigkeiten im Zeitraum der Jahrgangstufen 5–10 längsschnittlich zu erfassen. Damit sollen Erkenntnisse darüber gewonnen werden, wie und unter welchen Bedingungen sich mathematische Kompetenzen im Laufe der Schulzeit entwickeln bzw. mögliche Fehlkonzepte herausbilden. Unser Ziel ist es auch, zu erfassen, wie sich die Ausprägung von Grundvorstellungen bei spezifischen mathematischen Inhalten bzw. Inhaltsbereichen im Laufe der Sekundarstufe I entwickelt. Wesentlich sind dabei insbesondere die folgenden Fragen:

- Verläuft die Entwicklung der mathematischen Grundbildung *gleichmäßig* oder lassen sich *Phasen, Stufen oder Sprünge* identifizieren? Stehen Phasen mit positivem Lernzuwachs solche mit Einbrüchen mathematischer Kompetenzen (negativer Lernzuwachs) entgegen?
- Wie verläuft die Fähigkeitsentwicklung bei spezifischen *mathematischen Inhaltsbereichen* (z.B. Arithmetik, Algebra oder Geometrie), *mathematischen Schlüsselbegriffen* (z.B. Proportionalität, Prozent- oder Funktionsbegriff) und *mathematischen Grundvorstellungen* (z.B. zum Anteils- oder Verhältnisbegriff)?

Ziel der Testkonstruktion ist es damit, ein genaueres Bild der Genese mathematischer Grundbildung zeichnen zu können. Dies soll dem übergeordneten Ziel dienen, konkretere Ansatzpunkte für unterrichtliche, methodisch-didaktische und stofflich-curriculare Konsequenzen zu gewinnen.

3. Methode

3.1 Methodische Konzeption der Längsschnittstudie

Im Folgenden wird zunächst die Gesamtkonzeption der Längsschnittstudie und ihrer Voruntersuchungen skizziert, anschließend wird die Konstruktion des Regensburger Mathematikleistungstests dargestellt. Projektbeginn war im November 2000. Zum Zeitpunkt der Erstellung dieses Manuskripts ist die erste Voruntersuchung abgeschlossen, die Daten der zweiten Voruntersuchung befinden sich in der Auswertung und der erste Messzeitpunkt des Längsschnitts in der Implementierungsphase.

Untersuchungsanlage und Stichproben

Dem Längsschnitt sind zwei Feldtests vorgeschaltet, die der Entwicklung des Regensburger Tests und weiterer Instrumente für den Längsschnitt dienen (Voruntersuchungen I und II; $N = 784 / 1.643$ Schüler aus 5.–10. Klassen bayerischer Hauptschulen, Realschulen und Gymnasien). Der Längsschnitt wird gegen Ende des Schuljahres 2001/02 eine Kohorte von Schülern in der 5. Klassenstufe aufgreifen und in jährlichen Erhebungen bis zur 10. Klassenstufe verfolgen (Ausgangsstichprobe: $N = 2.100$ Schüler sowie deren Eltern und Mathematiklehrer; 84 Klassen aus Hauptschulen, Realschulen und Gymnasien). Die Stichprobenkalkulation erfolgte anhand von Power-Analysen unter Berücksichtigung der zu erwartenden Klumpungseffekte und Effektgrößen für Entwicklungs- und Bedingungeffekte (vgl. Cohen 1988). Die Gewinnung der Stichproben folgt PISA-üblichen Prinzipien der Stichprobenziehung, wobei in Übereinstimmung mit den nationalen Planungen für PISA 2003 und 2006 ganze Klassen gezogen werden, um klassenbezogene Unterrichts- und Kontexteffekte abschätzen zu können (bei PISA 2000 handelte es sich um Individualstichproben von Schülern).

In der neunten Klassenstufe werden die Erhebungen zeitlich und instrumententechnisch parallel zu den Erhebungen des dritten Zyklus von PISA durchgeführt (PISA 2006), um die angestrebte Aufklärung der Variation von PISA-Leistungen in Mathematik zu ermöglichen. Über diese Zielsetzung hinaus wird ein spezifisches Augenmerk auf die systematische Weiterverfolgung auch von Klassenwiederholern gelegt, um Daten zu dem in Deutschland bisher unzureichend analysierten Problem der Klassenwiederholung zu gewinnen (vgl. Bellenberg 1999).

Variablen und Instrumente

Die Variablenkonfiguration des Längsschnitts folgt der Zielsetzung des Projekts, Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von mathematischem Kompetenzerwerb zu analysieren. Das Instrumentarium orientiert sich deshalb konzeptionell und in der Verwendung einzelner Skalen an den Planungen für die nächsten Zyklen von PISA (erleichtert wird dies durch die Mitgliedschaft von R. Pekrun und W. Blum im nationalen Konsortium und R. vom Hofe in der nationalen Expertengruppe Mathematik von PISA 2003).

Neu entwickelt wurden neben dem Regensburger Mathematikleistungstest die *Münchener Skalen zu Mathematikemotionen*, die zur Erfassung von sieben mathematikbezogenen Emotionen bei Schülern dienen (Freude, Stolz, Ärger, Angst, Scham, Hoffnungslosigkeit und Langeweile in Unterrichts-, Prüfungs- und Hausaufgabensituationen in Mathematik; durchschnittliches $\alpha = .91$). Ferner werden weitere Skalen zu Schülervoraussetzungen (Intelligenztest und Selbstberichtskaalen, u.a. zu selbstbezogenen Kognitionen und Motivation in Mathematik) sowie Skalen zu Unterrichts-, Klassen- und Elternhauskontexten im Längsschnitt zum Einsatz kommen. Bei den Kontextskalen handelt es sich um mathematikbezogene Befragungsinstrumente, die jeweils nicht nur aus der Schülerperspektive, sondern auch aus der Eltern- bzw. Lehrerperspektive zu beantworten sind. Der gesamte Satz von Verfahren ist in der Voruntersuchung erprobt und revidiert worden.

3.2 Konstruktion des Regensburger Mathematikleistungstests für 5.–10. Klassen

Zeitliche Konzeption der Testkonstruktion und Stichproben

Für die Entwicklung des Mathematikleistungstests dienten die Voruntersuchungen I und II des Projekts (s.o.). In der bereits abgeschlossenen ersten Voruntersuchung wurde in Koordination mit der PISA-Expertengruppe Mathematik die erste Fassung des Mathematikleistungstests analysiert, wobei unterschiedliche Aufgaben und Aufgabenformate erprobt wurden. Auf der Grundlage dieser ersten Voruntersuchung erfolgt eine Revision und Weiterentwicklung der Instrumente, wobei Aufgabenformate fixiert und Skalenendformen entwickelt werden. Diese werden in der gegenwärtig laufenden zweiten Voruntersuchung erprobt.

Inhaltliche Strategie der Testkonstruktion

(1) *Ebenen der Leistungsmessung.* Entsprechend der Zielsetzung der geplanten Längsschnittstudie, Leistungsentwicklungen über die Zeit hinweg zu verfolgen, Ursachen dieser Entwicklungen zu analysieren und Handlungsmöglichkeiten abzuleiten, wurde das Instrumentarium so konzipiert, dass mathematische Fähigkeiten auf drei Ebenen erfasst werden können.

- A: Zur Erfassung von globalen Leistungsverläufen im Bereich Mathematik wird eine (nach dem dichotomen Raschmodell skalierbare) Serie von Aufgaben konstruiert, die mit der Methodik der Mathematikleistungstests bei PISA abgestimmt ist und zu einem *Gesamtscore* für die Mathematikleistung führt. Die inhaltliche und methodische Parallelität der Instrumente soll eine Übertragung der Befunde aus PISA-Erhebungen auf die Erhebungen dieses Projekts und vice versa möglich machen.
- B: Um ein differenzierteres Leistungsbild zu mathematischen Fähigkeiten in einzelnen Kompetenzbereichen zu erhalten, werden *Subskalen* entwickelt.
- C: Auf einer dritten Ebene werden Fähigkeitsentwicklungen bezüglich einzelner Items erfasst, was u.a. Rückschlüsse auf die Ausprägung entsprechender Grundvorstellungen erlaubt.

(2) *Formate der Aufgaben.* Analog zum Vorgehen bei PISA wird bei den Aufgabenformaten zwischen „multiple choice“, „kurze freie Antwort“ und „ausführliche freie Antwort“ unterschieden. Das Verteilungsverhältnis der einzelnen Formate orientiert sich an den Bedingungen von PISA 2003. Um den deutschen curricularen Gewohnheiten zu entsprechen, werden in Abstimmung mit dem nationalen PISA-Instrumentarium Itemserien, die von einem gemeinsamen „stimulus material“ abhängig sind, vermieden. Dadurch wird eine unabhängige Bewertung der einzelnen Items möglich und ihre statistische Unabhängigkeit erhöht (Neubrand u.a. 2001).

(3) *Inhaltliche Strukturierung der Aufgaben.* Die in der ersten Fassung des Verfahrens getesteten Stoffgebiete umfassen die Bereiche Arithmetik sowie Algebra und Funktionenlehre. Inhaltliche Schwerpunkte sind Bruch- und Prozentrechnung, Anteils- und Ver-

hältnisbegriff im Bereich Arithmetik sowie Umgang mit Skalen, Proportionalität, Zuordnungen und Linearität im Bereich Algebra/Funktionen. In der zweiten Voruntersuchung werden die Stoffgebiete unseres Tests so ergänzt, dass bis zur Implementierung des Längsschnitts stoffliche Äquivalenz zu den Erhebungen von PISA 2003 erreicht wird.

(4) *Grundvorstellungskomplexität als Aufgabenmerkmal.* Das Instrumentarium ist so angelegt, dass jedem inhaltlichen Teilbereich eine Itemserie entspricht, deren Items sich charakteristischen Grundvorstellungen bzw. Grundvorstellungskombinationen zuordnen lassen. Weiterhin wurde zu jedem Teilbereich eine Itemgruppe entwickelt, deren Items keine Grundvorstellungen erfordern, sondern lediglich durch Kalkülanwendung gelöst werden können. Die Ausprägungen anderer Aufgabenparameter (z.B. sprachlogische Komplexität) wurden kontrolliert niedrig gehalten.

Die Items werden sowohl hinsichtlich ihrer Grundvorstellungs- als auch ihrer Kalkülkomplexität klassifiziert. Hierzu dienen mit dem Vorgehen bei PISA abgestimmte Aufgabenklassifikationen durch Experten. Bei der Klassifizierung der Items nach ihrer Grundvorstellungsintensität wird ein Klassifikationssystem eingesetzt, das in Kooperation mit der Arbeitsgruppe um W. Blum entwickelt wurde (Blum/vom Hofe im Druck) und gegenwärtig auch für Detailanalysen von PISA 2000 verwendet wird. Dabei erfolgt eine Gewichtung der Items nach der Ausprägung der jeweils für eine Lösung der Items erforderlichen Grundvorstellungen.

(5) *Grundvorstellungskompetenz als Fähigkeitsmerkmal.* Um die für mathematische Modellierungsprozesse zentrale Grundvorstellungskompetenz – und im Vergleich dazu die Kalkülkompetenz – einzelner Individuen zu erfassen und in ihren Entwicklungsverläufen zu dokumentieren, werden zur Zeit Subskalen entwickelt. Neben einer globalen individuellen Grundvorstellungskompetenz soll über weitere Subskalen auch die Grundvorstellungskompetenz in spezifischen Inhaltsbereichen (z.B. Proportionalität) erhoben werden. Diese ergibt sich aus den Kompetenzwerten für diejenigen Itemgruppen, die sich den jeweiligen mathematischen Inhaltsbereichen zuordnen lassen. Die auf diese Weise beschriebene Kompetenz wird in der didaktischen Literatur häufig als *Grundverständnis* eines Stoffbereichs (Ausprägungsgrad von Grundvorstellungen und ihrer Vernetzung; vgl. Oehl 1970; Bender 1991; vom Hofe 1995) bezeichnet.

(6) *Itemserien und Testversionen.* Insgesamt wurden für die erste Voruntersuchung 69 Aufgaben mit 116 Items entwickelt. Diese 69 Aufgaben wurden auf vier alterstufenabhängige Testversionen verteilt (zwei Testversionen für die 5. Klassenstufe, zwei Testversionen für die 6.–10. Klassenstufe). Die Testversionen für die 5. Klassenstufe unterscheiden sich nur durch Rotation der Itemanordnungen. Bei den Testversionen für die 6.–10. Klassenstufe gibt es einen gemeinsamen Ankeritempool und jeweils zusätzliche Items zu jedem Themengebiet und jeder Grundvorstellungsklasse. Der Grund für die asymmetrische Aufteilung der Testversionen auf die Klassenstufen ist in der curricularen Strukturierung des Lehrstoffes zu sehen: Während in der 5. Klassenstufe vorwiegend der

Grundschulstoff vertieft wird, werden ab der 6. Klassenstufe neue Konzepte (z.B. Proportionalität) eingeführt, die bis zur 10. Klassenstufe durchgängig eine Rolle spielen.

Die Verankerung folgt dem Instrumentendesign der nationalen PISA-Erhebungen, das ausgehend von einem gemeinsamen Ankeritemsatz die zusätzlichen Items auf mehrere Testhefte verteilt.

(7) *Skalierung und Itemanalysen.* Zur Skalierung des Mathematiktests wird ebenso wie bei den Mathematiktests von PISA das eindimensionale Rasch-Modell verwendet (vgl. Rost 1996). Ein hier besonders wesentlicher Vorteil gegenüber Verfahren der klassischen Testkonstruktion liegt in der Möglichkeit, Leistungen verschiedener Personen auch dann auf einer gemeinsamen Skala abzubilden, wenn diese Personen unterschiedliche Aufgaben bearbeitet haben. Dies eröffnet Möglichkeiten, unterschiedliche Testversionen, die sich auf verschiedene Altersstufen und Themenbereiche beziehen, einer gemeinsamen Skalierung zu unterziehen, um auf diese Weise u.a. Entwicklungsverläufe über Klassenstufen hinweg angemessen modellieren zu können.

Die Grenzen einer eindimensionalen Modellierung werden allerdings deutlich, wenn man die zugrunde liegenden Modellannahmen betrachtet: Die Annahme, dass eine einzige Fähigkeitsdimension die Testleistung determiniert, spiegelt eher einen gewissen Skalierungspragmatismus wieder. Für die Ebene A unserer Leistungsmessung (Ermittlung eines Gesamtwerts für die Mathematikleistung) scheint ein solches Vorgehen – analog zum Vorgehen bei PISA – angemessen zu sein. Auf der Ebene B sind jedoch Subskalen für differenziertere Analysen erforderlich. Auf der Ebene C sind für noch spezifischere Untersuchungen differenzielle Itemanalysen geplant. Für einige Subskalen existieren bereits Vorformen; sie werden in der zweiten Voruntersuchung weiterentwickelt.

4. Ergebnisse und Diskussion

4.1 Raschskalierung des Mathematikleistungstests

Ziel der in der ersten Voruntersuchung vorgenommenen Raschskalierung war es, die Skalierbarkeit der vorläufigen Testversionen zu überprüfen und Hinweise zur Weiterentwicklung zu gewinnen. Erfreulicherweise erweist sich der Test bereits in seiner ersten Fassung als ausgewogen schwierigkeitsgestaffelt über ein Spektrum von etwa 8 Logit-Einheiten. Dabei ist insbesondere der Bereich zwischen -2.5 und $+2.5$ Logit-Einheiten mit 102 der insgesamt 116 Items dicht besetzt.

Betrachtet man die mittleren Fähigkeiten $\bar{\theta}_v$ der Probanden über die einzelnen Klassenstufen hinweg (vgl. Tabelle 1), so erkennt man eine Zunahme der mittleren Fähigkeit von etwa einer halben Standardabweichung gegenüber der jeweils vorangegangenen Klassenstufe. Dieses Ergebnis entspricht den Befunden vergleichbarer Untersuchungen (z.B. BIJU, Bildungslebensläufe und psychosoziale Entwicklung im Jugendalter; vgl. Baumert/Gruehn/Heyn/Köller/Schnabel 1997) und spricht damit für die Validität des Instrumentariums.

Tab. 1: **Mittlere Fähigkeit $\bar{\theta}_v$ der Probanden nach Klassenstufe**

Jahrgangsstufe	$\bar{\theta}_v$	N	s	D(-1)[s]
5. Klasse	-1.48	176	1.25	
6. Klasse	-.77	103	.81	.56
7. Klasse	-.29	140	.89	.59
8. Klasse	.26	150	1.15	.63
9. Klasse	.99	109	.93	.63
10. Klasse	1.47	102	1.09	.51

Anmerkung: D(-1)[s]: Fähigkeitsdifferenz der Schulformen in Einheiten der Standardabweichung gegenüber dem Vorjahr.

Tab. 2: **Mittlere Fähigkeit $\bar{\theta}_{rv}$ der Probanden nach Schulform**

Schulform	$\bar{\theta}_{rv}$	N	s	D(-1)[s]
Hauptschule	-.87	204	1.32	
Realschule	-.14	270	1.48	.49
Gymnasium	.43	306	1.28	.29

Anmerkung: D(-1)[s]: Fähigkeitsdifferenz der Schulformen in Einheiten der Standardabweichung gegenüber dem Vorjahr

Bezüglich der Schulformen ergibt sich ein differenziertes Bild (vgl. Tabelle 2): Der Fähigkeitsschritt von der Hauptschule zur Realschule ist etwas größer als der von der Realschule zum Gymnasium. Ein Grund ist möglicherweise darin zu sehen, dass die sechsstufige Realschule sich in Bayern gegenwärtig noch im Aufbau befindet. Dies hatte zur Folge, dass in unserer Realschülerstichprobe die fünfte Klassenstufe bereits vertreten ist, die sechste Klassenstufe hingegen noch nicht, so dass die Realschülerstichprobe überwiegend aus Schülern höherer Klassenstufen und damit höherer Leistungsfähigkeit bestand, während Schüler niedrigerer Fähigkeitsbereiche eher fehlten.

4.2 Itemanalysen und Testrevision

Eine genauere Analyse der Schwierigkeitsstaffelung innerhalb der Testversionen zeigt, dass die vorläufigen Testversionen eine differenzierte Erfassung des Fähigkeitsspektrums der unteren Klassenstufen erlauben, während die Erfassung in den höheren Klassen- und Leistungsstufen vergleichsweise weniger differenziert ausfällt. Hieraus ergeben sich Konsequenzen für die weitere Testentwicklung: Die Differenzierung der vorläufigen Testversionen scheint noch nicht optimal zu sein. Dies gilt insbesondere für die Testversionen für die 6.–10. Klassenstufe, da diese fünf Klassenstufen ein erhebliches Leistungsspektrum umfassen. Allerdings scheint es nicht notwendig zu sein, für jede Klassenstufe getrennte Testversionen zu entwickeln, da sich die Leistungsbandbreiten auf-

einander folgender Klassenstufen stark überschneiden (siehe Tabelle 1); zudem würde dies die Analyse von Entwicklungsverläufen eher erschweren.

Als sinnvoller Mittelweg wurde daher eine Aufteilung in drei schwierigkeitsgestaffelte Tests gewählt. Unter Beibehaltung des Prinzips der Paralleltestversionen wurden sechs Testversionen erstellt, die zur Zeit im Rahmen der zweiten Voruntersuchung erprobt werden. Anhand von gemeinsamen Ankeritempools kann dabei weiterhin eine gemeinsame eindimensionale Skalierung nach dem Rasch-Modell vorgenommen werden.

Zur Überprüfung, Selektion und Revision einzelner Items dienten neben den Itemschwierigkeiten auch die Fit-Werte und Trennschärfekoeffizienten der Items. Es zeigt sich, dass mehr als 80% der Items sowohl den für die Raschskalierung bei PISA üblichen Fitkriterien als auch den Trennschärfekriterien der klassischen Testtheorie genügen. Nur 14 der 116 Items weisen eine Trennschärfe von $r_{pbis} < .30$ auf. Für die revidierte Fassung des Mathematiktests wurden entsprechende Itemrevisionen vorgenommen.

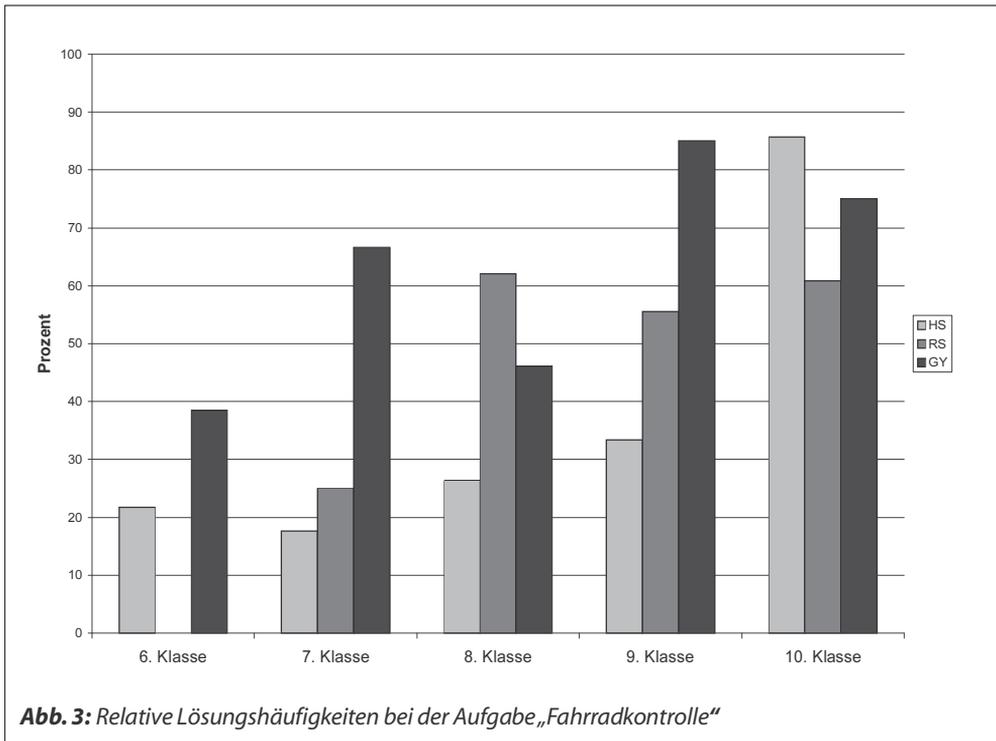
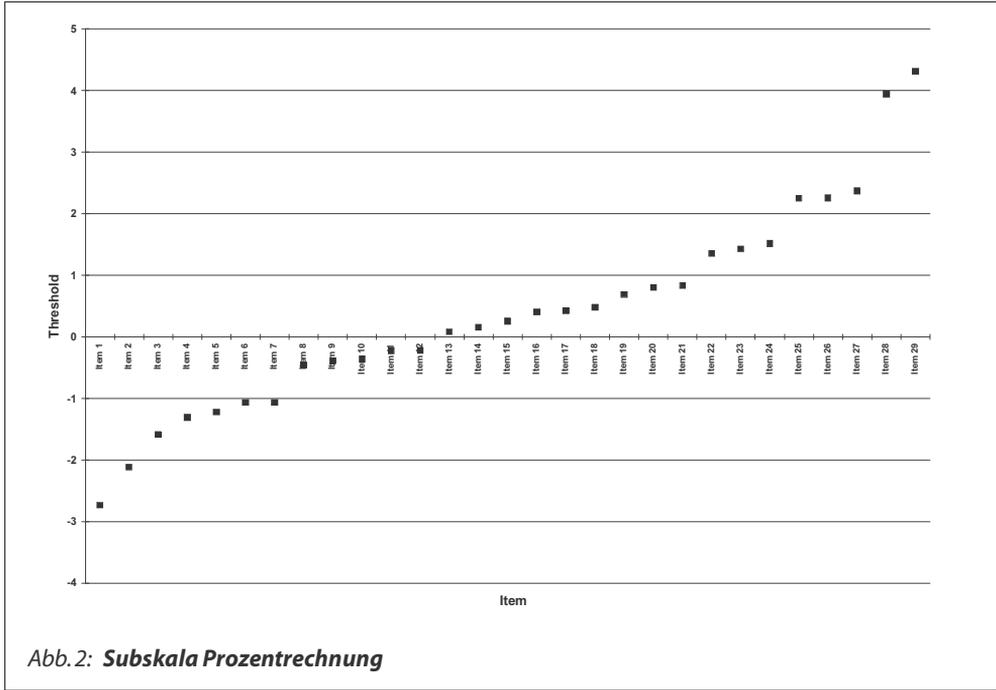
4.3 *Subskalen und differenziertere Analysen*

Mit der Konstruktion von Subskalen zur Analyse stoffgebietsspezifischer Fähigkeiten und Entwicklungsverläufe stehen wir gegenwärtig noch am Anfang. Exemplarisch kann aber an dieser Stelle am Beispiel des Bereichs Prozentrechnung ein Ausblick gegeben werden. Die Subskala „Prozentrechnung“ (Abbildung 2) beruht auf einer Rasch-Skalierung der Items zu diesem mathematischen Stoffgebiet. Theoriekonform entsprechen die Reihungen der Item-Thresholds den Reihungen, die sich für die Items im Gesamttest ergeben. Die gebietsspezifische Modellierung ermöglicht es, auf dieses Stoffgebiet bezogene Fähigkeiten der Schüler zu untersuchen und mit ihren Fähigkeiten in anderen Stoffgebieten sowie dem Gesamtwert im Mathematikleistungstest zu vergleichen.

Geplant ist die Entwicklung weiterer Subskalen, u.a. zu den curricularen Bereichen Arithmetik, Algebra und Geometrie, zur Grundvorstellungs- bzw. Kalkülkompetenz sowie – getrennt hiervon – zu den fünf mathematischen Kompetenzklassen, die für die Dimensionierung des nationalen Mathematikleistungstests bei PISA vorgesehen sind. Ferner sind differenziertere Analysen zu Bereichen wie Prozentrechnung, Verhältnisbegriff, Anteilsbegriff, Proportionalität, Linearität und Funktionsbegriff geplant.

4.4 *Entwicklungsverläufe und didaktische Handlungsperspektiven*

Bereits die – hier noch querschnittlichen – Befunde der ersten Voruntersuchung zu Entwicklungsverläufen zeigen interessante und z.T. sehr überraschende Ergebnisse. Unterhalb der Ebene einer allgemeinen Fähigkeitszunahme (s.o. 2.3) ist keineswegs von monoton klassenstufenabhängigen Fortschritten der Mathematikleistung in einzelnen Kompetenzbereichen auszugehen; vielmehr zeigen sich u.a. auch paradoxe, z.B. stagnierende Verläufe bzw. Phasen eines negativen Lernzuwachses. Betrachten wir hier – mit aller aufgrund der geringen Stichprobengröße der ersten Voruntersuchung gebotenen Vorsicht – zwei charakteristische Beispiele.



(1) Stagnierende Verläufe beim Verständnis des Anteilsbegriffs

Ein interessanter Effekt, der längsschnittlich näher zu untersuchen sein wird, deutet sich bei der Aufgabe „Fahrradkontrolle“ aus dem Bereich des Anteilsbegriffs an. Bei dieser Aufgabe wird der Proband aufgefordert, einen gegebenen relativen Anteil (2 von 5) aus einer Zeitungsmeldung in den entsprechenden prozentualen Anteil umzuwandeln. In Abbildung 3 sind die Lösungshäufigkeiten für diese Aufgabe wiedergegeben, differenziert nach Schulformen und Klassenstufen.

Betrachtet man die Lösungshäufigkeiten der Realschüler, so lässt sich von der 7. zur 8. Klassenstufe ein deutlicher Fähigkeitszuwachs erkennen, der in den nachfolgenden Klassenstufen zu stagnieren scheint. Eine ähnliche Entwicklung findet sich – um eine Klassenstufe vorgezogen – bei den Gymnasiasten. Zurückgehen könnte dieser Effekt auf mangelnde unterrichtliche Vernetzung des Anteilsbegriffs mit den Themen der höheren Klassenstufen, die zu Vergesseneffekten (vgl. Frank 1996) und damit zu Leistungsabfällen führen könnte. Sollte sich dieser Effekt längsschnittlich erhärten lassen, wären entsprechende curriculare Konsequenzen zu ziehen.

Erstaunlich mag hier auf den ersten Blick das hohe Fähigkeitsniveau der Hauptschüler der 10. Klassenstufe erscheinen. Dies wird jedoch verständlich, wenn man bedenkt, dass zum einen die zehnte Hauptschulklasse in Bayern nur noch von einer relativ kleinen und hoch motivierten Gruppe von Lernenden besucht wird, welche die Mittlere Reife anstreben, und dass zum anderen Hauptschüler aufgrund des bayerischen Curriculums in Mathematik mehr Erfahrungen im Bereich des Sachrechnens haben als ihre Altersgenossen in den anderen Schulformen.

Insgesamt ist bemerkenswert, dass ein erheblicher Teil der Untersuchungsteilnehmer aus allen Schularten die hier geforderte Fähigkeit, die etwa für das Verständnis von relativen Zahlenangaben in Zeitungsartikeln grundlegend ist, nicht beherrscht. Selbst in der zehnten Klassenstufe scheint die *Grundvorstellung des relativen Anteils* in vielen Fällen nicht hinreichend gefestigt zu sein.

(2) Einbrucheffekte beim Verständnis von Proportionalität

Einen anderen Effekt zeigt der Fähigkeitsverlauf beim Item „Lotto“ aus dem Bereich Proportionalität (vgl. Abbildung 4). Bei dieser Aufgabe geht es inhaltlich um die Frage der Aufteilung eines Jackpots bei einem Lottospiel in Abhängigkeit von der Anzahl der Gewinner (indirekt proportionale Zuordnung).

Interessant ist der Abfall in der Lösungshäufigkeit von der 5. zur 6. Klassenstufe. Bei Schülern des Gymnasiums ist ein Einbruch auf ein Drittel der in der 5. Klassenstufe erzielten Lösungshäufigkeiten zu beobachten. Dieses ist um so bemerkenswerter, als dieser Inhalt im bayerischen Lehrplan für die 6. Klassenstufe des Gymnasiums einen zentralen Punkt darstellt und damit ein wesentliches Gebiet des Pflichtunterrichts ist. Es sieht hier so aus, als würde die Fähigkeit in einem spezifischen Gebiet erheblich abnehmen, obwohl dieses Gebiet zum ersten Mal systematisch unterrichtet wird.

Eine Erklärung kann darin gesehen werden, dass das intuitive, erfahrungsweltliche Verständnis von indirekter Proportionalität in der 5. Klassenstufe zunächst relativ erfolgreich ist. So ist es etwa möglich, die Aufgabe „Lotto“ mittels der aus der Grundschu-



le vertrauten Grundvorstellungen des *Vervielfachens* und *Aufteilens* zu lösen. Zu vermuten ist, dass die schulischen, strukturorientierten Betrachtungen der 6. Klassenstufe diese frühen, erfolgreichen Konzepte nicht angemessen aufgreifen und weiterentwickeln, sondern eher behindern oder verschütten, ohne adäquate neue Konzepte aufzubauen. Bis zur 8. Klassenstufe können diese Defizite offenbar nicht mehr behoben werden, und erst mit einem erweiterten Funktionsverständnis in der 9. Klassenstufe scheint ein sicherer Umgang mit Aufgaben dieses Stoffgebiets gewährleistet zu sein.

Im Bereich der Verhältnisrechnung lieferte Wiegand (2000) in einer explorativen Studie ähnliche Befunde, die sich aus einer Kurzintervention mit negativem Lerneffekten ergaben. Beobachtungen dieser Art lassen darauf schließen, dass es in der Leistungsentwicklung Phasen gibt, in denen bestehendes intuitives Wissen nicht angemessen berücksichtigt wird, sondern durch inadäquate, meist schematische Verfahren verschüttet wird, was zu Leistungsabfällen und negativen Wirkungen im motivationalen Bereich führen kann (vgl. Fischbein 1990; Wittmann 1991).

4.5 Schlussfolgerungen

Die Befunde der ersten Voruntersuchung des Projekts zeigen, dass die entwickelten Instrumentarien zur Erfassung von mathematischer Grundbildung, selbstreguliertem Lernen in Mathematik sowie Unterrichts- und Kontextbedingungen erfolgreich eingesetzt werden können. Exemplarisch gezeigt wurde dies hier für den Regensburger Mathematikleistungstest. Dieses Verfahren erweist sich bereits in der ersten Fassung als raschskalierbar und differenzierungsfähig, und zwar über den gesamten Entwicklungsbereich der Sekundarstufe I und alle Schularten des gegliederten Schulwesens hinweg (mit Aus-

nahme von Sonderschulen). Die Kompetenzzunahme von einer Klassenstufe zur nächsten entspricht etwa einer halben Standardabweichung der Kompetenzvariation innerhalb der Klassenstufen; dies spricht für die Validität des Instruments.

Dabei ergeben sich überraschende Befunde, wenn man Kompetenzverläufe unterhalb der Ebene eines globalen Summenscores für die Mathematikleistung analysiert. In einzelnen curricularen Bereichen scheinen Phasen mit positivem Lernzuwachs solchen mit stagnierender Leistungsentwicklung oder sogar negativem Lernzuwachs gegenüber zu stehen. Bei der Interpretation ist allerdings zu berücksichtigen, dass die hier berichteten Befunde dieser Art zunächst querschnittlich angelegt sind und auf relativ kleinen Stichproben basieren. Zentrales Ziel der Längsschnittstudie wird es sein, sowohl gelungene Kompetenzentwicklungen wie auch problematische Leistungsverläufe in methodisch angemessener Weise zu identifizieren und ihre Bedingungen in Unterricht und sozialen Kontexten zu analysieren, um die sich ergebenden Ansatzpunkte für unterrichtliche und didaktische Handlungskonsequenzen realistisch zu erhärten.

Literatur

- Aspinwall, L.G. (1998): Rethinking the role of positive affect in self-regulation. In: *Motivation and Emotion*, 22, S. 1–32.
- Baumert, J./Gruehn, S./Heyn, S./Köller, O./Schnabel, K.U. (1997): *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU)*. Dokumentation. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiß, M. (Hrsg.) (2001): *PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich.
- Baumert, J./Lehmann, R. (Hrsg.) (1997): *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich*. Opladen: Leske & Budrich.
- Bellenberg, G. (1999): *Individuelle Schullaufbahnen. Eine empirische Untersuchung über Bildungsverläufe von der Einschulung bis zum Abschluss*. Weinheim: Juventa.
- Bender, P. (1991): *Ausbildung von Grundvorstellungen und Grundverständnissen – ein tragendes didaktisches Konzept für den Mathematikunterricht*. In: Postel, H./Kirsch, A./Blum, W. (Hrsg.): *Mathematik lehren und lernen*. Hannover: Schroedel, S. 48–60.
- Blum, W./Törner, G. (1983): *Didaktik der Analysis*. Göttingen: Vandenhoeck & Ruprecht.
- Blum, W./vom Hofe, R. (im Druck): *Welche Grundvorstellungen sind hier erforderlich? – Analysen zur Beurteilung des Anspruchsniveaus von Aufgaben*. In: *Mathematik lehren*, Heft 112.
- Blum, W. (1996). *Anwendungsbezüge im Mathematikunterricht – Trends und Perspektiven*. In: Kadunz, G./Kautschitsch, H./G. Ossimitz, G./Schneider, E. (Hrsg.): *Trends und Perspektiven. Beiträge zum 7. Internationalen Symposium zur Didaktik der Mathematik*, S. 15–38. Wien: Hölder.
- Brousseau, G. (1983): *Les obstacles épistémologique et les problèmes en mathématiques*. In: *Revue Recherches en Didactique des mathématiques*, 4 (2), 165–198.
- Cohen, J. (1988): *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- De Lange, J. (1996): *Real problems with the real world mathematics*. In: Alsina, C./Alvarez, J. M./Niss, M./Perez, A./Rico, L./Sfard, A. (Hrsg.): *Proceedings of the 8th International Congress on Mathematical Education, Sevilla July 1996*. Sevilla: S.A.E.M. Thales, S. 83–110.
- Fischbein, E. (1983): *Intuition and analytical thinking in mathematics education*. In: *Zentralblatt für Didaktik der Mathematik*, 2, S. 68–74.

- Fischbein, E. (1989): Tacit models and mathematical reasoning. In: *For the Learning of Mathematics*, 9, S. 9–14.
- Fischbein, E. (1990): The autonomy of mental models. In: *For the Learning of Mathematics*, 10, 23–30.
- Freudenthal, H. (1977): *Mathematik als pädagogische Aufgabe* (2 Bde.). Stuttgart: Klett.
- Freudenthal, H. (1983): *Didactical phenomenology of mathematical structures*. Dordrecht: Reidel.
- Frank, H. (1996): *Bildungskybernetik – Eine Kurzeinführung in die kybernetisch-pädagogischen Modellgrundlagen der Bildungstechnologie*. München: NITRA & KoPäd.
- Götz, T. (2002): *Emotionales Erleben und selbstreguliertes Lernen bei Schülern im Fach Mathematik*. Universität München: Fakultät für Psychologie und Pädagogik.
- Helmke, A./Weinert, F.E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Weinert, F.E. (Hrsg.): *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Serie Pädagogische Psychologie, Bd. 3, S. 71–76). Göttingen: Hogrefe.
- Hofe, R. vom (1992): Grundvorstellungen mathematischer Inhalte als didaktisches Modell. In: *Journal für Mathematik-Didaktik*, 13, S. 345–364.
- Hofe, R. vom (1995): *Grundvorstellungen mathematischer Inhalte*. Heidelberg: Spektrum.
- Hofe, R. vom (1998a): On the generation of basic ideas and individual images: Normative, descriptive and constructive aspects. Kilpatrick, J./Sierpinska, A. (Eds.), *Mathematics as a research domain*. Dordrecht, The Netherlands: Kluwer.
- Hofe, R. vom (1998b): Probleme mit dem Grenzwert – genetische Begriffsbildung und geistige Hindernisse. In: *Journal für Mathematik-Didaktik*, 19, S. 257–291.
- Hofe, R. vom (1999): Explorativer Umgang mit Funktionen – Interaktion und Kommunikation in selbstorganisierten Arbeitsphasen. In: *Journal für Mathematik-Didaktik*, 20, S. 186–221.
- Hofe, R. vom (im Druck): Investigations into students' learning of applications. Computer-based learning environments. In: *Proceedings of ICME 9*.
- Klieme E./Neubrand, M./Lüdtke, O. (2001): *Mathematische Grundbildung: Testkonzeption und Ergebnisse*. In: Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiß, M. (Hrsg.) (2001): *PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich.
- Ma, X. (1999): A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. In: *Journal for Research in Mathematics Education*, 30, S. 520–540.
- Maier, H. (1990): *Didaktik des Zahlbegriffs*. Hannover: Schroedel.
- Malle, G. (im Druck): *Grundvorstellungen und Grundverständnis*. In: *Mathematik lehren*.
- Neubrand, M./Bieler, R./Blum W./Cohors-Fresenborg, E./Flade, L./Knoche, N./Lind, D./Löding, W./Möller, G./Wynands, A. (2001): *Grundlagen der Ergänzung des internationalen PISA-Mathematik-Tests in der deutschen Zusatzerhebung*. Zentralblatt für Didaktik der Mathematik, 33, S. 1–15.
- Oehl, W. (1970): *Der Rechenunterricht in der Hauptschule*. Hannover: Schroedel.
- Pekrun, R. (1992): The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. In: *Applied Psychology: An International Review*, 41, S. 359–376.
- Pekrun, R. (1997): Kooperation zwischen Elternhaus und Schule. In: Vaskovics, L.A./Lipinski, H. (Hrsg.), *Familiale Lebenswelten und Bildungsarbeit* (Bd. 2). Opladen: Leske & Budrich, S. 51–79.
- Pekrun, R. (1998): Schüleremotionen und ihre Förderung: Ein blinder Fleck der Unterrichtsforschung. In: *Psychologie in Erziehung und Unterricht*, 45, S. 230–248.
- Pekrun, R. (2000): A social cognitive, control-value theory of achievement emotions. In: Heckhausen, J. (Ed.), *Motivational psychology of human development*. Oxford, UK: Elsevier Science.
- Pekrun, R. (2001): Familie, Schule und Entwicklung. In: Walper, S./Pekrun, R. (Hrsg.), *Familie und Entwicklung: Perspektiven der Familienpsychologie*. Göttingen: Hogrefe, S. 84–105.
- Pekrun, R. (2002): Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für die Bildungsforschung. In: *Zeitschrift für Pädagogik*, 48, S. 111–128.
- Pekrun, R./Götz, T./Titz, W./Perry, R.P. (im Druck): Emotions in students' self-regulated learning and achievement. In: *Educational Psychologist*.
- Rost, J. (1996): *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.

- Schubring, G. (1978): Das genetische Prinzip in der Mathematikdidaktik. Stuttgart: Klett-Cotta.
- Schupp, H. (1988): Anwendungsorientierter Mathematikunterricht in der Sekundarstufe zwischen Theorie und neuen Impulsen. In: Der Mathematikunterricht, 34 (6), S. 5–16.
- Sierpinska, A. (1992): On understanding the notion of function. In: Harel, G./Dubinsky, E. (Hrsg.): The concept of function – Aspects of epistemology and pedagogy (MAA Notes, Vol. 25). Washington, DC: Mathematical Association of America, S. 25–58.
- Usiskin, Z. (1991): Building mathematics curricula with applications and modeling. In: Niss, M./Blum, W./Huntley, I. (Hrsg.): Teaching of mathematical modelling and applications. Chichester: Harwood, S. 30–45.
- Wagenschein, M. (1983): Erinnerungen für Morgen. Weinheim, Basel: Beltz.
- Wiegand, B. (1998): Stoffdidaktische Analysen von TIMSS-Aufgaben. In: Mathematik Lehren, Heft 90, S. 18–22.
- Wiegand, B. (2000): Mathematische Anwendungsfähigkeiten – Detailanalysen von TIMSS und Kassel-Exeter-Studie. Hildesheim: Franzbecker.
- Wild, E. (1999): Elterliche Erziehung und schulische Lernmotivation. Unveröffentlichte Habilitationsschrift, Universität Mannheim.
- Wild, E./Remy, K. (2002). Quantität und Qualität der elterlichen Hausaufgabenbetreuung von Drittklässlern in Mathematik. In: Zeitschrift für Pädagogik (in diesem Band).
- Wittmann, E. (1991): Mathematikunterricht zwischen Skylla und Charybdis. In: Mathematische Gesellschaft Hamburg (Hrsg.), Mitteilungen der mathematischen Gesellschaft in Hamburg, Band XII, Heft 3 (Festschrift zum 300-jährigen Bestehen der Gesellschaft). Stuttgart: Klett, S. 663–679.

Anschriften der Autoren:

- Prof. Dr. Rudolf vom Hofe, Universität Regensburg, Didaktik der Mathematik, 93040 Regensburg.
- Prof. Dr. Reinhard Pekrun, Universität München, Department für Psychologie, Leopoldstraße 13, 80802 München.
- Michael Kleine, Universität Regensburg, Didaktik der Mathematik, 93040 Regensburg.
- Dipl.-Psych. Thomas Götz, Universität München, Department für Psychologie, Leopoldstraße 13, 80802 München.

Teil II:

Lehrerexpertise und Unterrichtsmuster in Mathematik und Physik

Videografie von Unterrichtssequenzen in Mathematik
und Physik: Diagnose, Analyse und Training erfolgreicher
Unterrichtsskripts

Eckhard Klieme

Einleitung

Die Ergebnisse der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS) waren, wie im einleitenden Beitrag zu diesem Band dokumentiert, Anlass zur Etablierung des DFG-Schwerpunktprogramms Bildungsqualität. Eine wesentliche Komponente von TIMSS war die Videoaufzeichnung von Unterrichtsstunden im Fach Mathematik, die vergleichend in den USA, Japan und Deutschland durchgeführt wurde (Stigler u.a. 1999). Diese Videostudie war für die öffentliche Rezeption von TIMSS, aber auch für die wissenschaftliche Debatte von zentraler Bedeutung, weil sie den Fokus von der reinen Leistungserhebung auf Fragen des unterrichtlichen Handelns von Lehrern, der Interaktionsprozesse im Unterricht, der typischen Ablaufmuster von Unterrichtsstunden im Fach Mathematik und andere methodisch-didaktische Aspekte der Schulpraxis verlagerte. Stigler und Mitarbeiter haben detailliert beschrieben, welche Art von Aufgaben im Mathematikunterricht der drei Länder behandelt werden, wie aktiv Lehrer und Schüler an der Erarbeitung neuer Inhalte und der Lösung von Übungsaufgaben beteiligt sind, welche Sozialformen und Medien zum Einsatz kommen und anderes mehr. Aufgrund der Videoaufzeichnungen und Transkriptionen war es möglich, die ausgewählten Unterrichtsstunden nach unterschiedlichsten Kategorien auszuwerten, durch Experten aus verschiedenen Blickwinkeln bewerten zu lassen und im Detail bis hin zu qualitativen Fallstudien auszuwerten (vgl. etwa die Ergänzungen zur TIMS-Videostudie in Deutschland, dargestellt bei Klieme/Schümer/Knoll 2002).

Populär geworden ist vor allem Begriff der „Unterrichtsskripts“, den Stigler und Mitarbeiter prägten, um typische Verlaufsmuster der Mathematikstunden darzustellen. Die Botschaft lautete: Trotz aller beachtlichen Unterschiede in Thematik, Methodik und Qualität des Unterrichts ähneln sich doch die beobachteten Verlaufsformen innerhalb der einzelnen Länder so sehr, dass man sie unter einem kulturspezifischen Muster subsumieren und gegen die pädagogische Praxis in anderen Kulturen deutlich abgrenzen kann. So wurde für den deutschen Unterricht das kleinschrittige, engführende fragend-entwickelnde Vorgehen bei der Erarbeitung neuer Inhalte wie auch in den Übungsphasen als typisch beschrieben, während etwa ein Drittel der japanischen Stunden einer Form des problemorientierten Unterrichts folgte, bei der Schüler in Einzelarbeit und anschließender Gruppendiskussion verschiedene Lösungswege für komplexe, offene Aufgaben explorieren. Diese Art des Unterrichts ist aus der japanischen Fachliteratur als Reformkonzept bekannt („open ended approach“). Auf dem Hintergrund aktueller pädagogischer Diskussion lässt sich ein solcher Unterricht als eine Form des „konstruktivistisch orientierten“ Lehrens ansehen, d.h. als ein Unterricht, der dem Prinzip folgt, dass Schüler sich ihr Verständnis mathematischer Konzepte durch aktive Auseinandersetzung mit komplexen Problemsituationen, anknüpfend an ihrem Vorwissen, und

nicht zuletzt im argumentativen Austausch innerhalb der Lerngruppe erarbeiten. Nimmt man den Befund hinzu, dass japanische Schüler vor allem bei komplexen Problemaufgaben deutlich besser abschneiden als gleichaltrige deutsche (Klieme/Bos 2000), so mag ein Plädoyer für die Nutzung solcher konstruktivistisch orientierten Unterrichtsformen in der Schulpraxis nahe liegen. (Der Begründungszusammenhang bleibt allerdings hypothetisch, weil Stigler seine Videoaufzeichnungen nicht mit klassenbezogenen Leistungsdaten verknüpfte.)

Die im folgenden Abschnitt vorgestellten vier Projekte des DFG-Schwerpunktprogramms Bildungsqualität arbeiten nicht nur mit der Methode der Videoaufzeichnung, um die Qualität von Lehr-Lernprozessen systematisch zu untersuchen, sondern sie nehmen auch explizit auf das Konzept der Unterrichtsskripts aus TIMSS Bezug. Die Projekte bearbeiten Fragestellungen, die bei Stigler u.a. und insbesondere durch die längsschnittliche Erweiterung der TIMS-Video studie in Deutschland vorbereitet wurden. Sie gehen jedoch in wesentlichen Punkten darüber hinaus und unterziehen damit die populär gewordene Option für konstruktivistisch orientiertes Lehren einer kritischen Prüfung.

1. Ein wesentlicher Erkenntnisgewinn des DFG-Schwerpunktprogramms liegt bereits darin, dass es die videogestützte Unterrichtsanalyse über das Fach Mathematik hinaus auf die *Naturwissenschaften* ausdehnt. Fischer u.a. sowie Prenzel u.a. beschreiben erstmals systematisch den Physikunterricht der Mittelstufe, wie er in Deutschland praktiziert wird. Insbesondere durch Schüler- und Demonstrationsexperimente kommen neue naturwissenschaftsspezifische Elemente in den Blick.

2. Die Beiträge aus dem DFG-Schwerpunktprogramm nehmen explizit und mit klarer theoretischer Orientierung Bezug auf *Konzepte der pädagogisch-psychologischen Forschung* und insbesondere der Fachdidaktik. So wird Unterricht systematisch unter dem Gesichtspunkt des Aufbaus von *literacy* (Fischer u.a.) behandelt. Durch alle vier Beiträge zieht sich eine klassische Grundfrage der „Unterrichtslehre“, die stark durch die Tradition der Reformpädagogik geprägt ist und durch die Kontroversen über konstruktivistisch orientierten Unterricht an Aktualität gewonnen hat: Die Frage nach dem optimalen Verhältnis zwischen Offenheit von Lernsituationen einerseits, Strukturiertheit der Lernumgebung und des Lehrerverhaltens andererseits.

3. Während Stigler u.a. (1999) Unterricht rein deskriptiv erfassten, sodass ein Bezug zu *Lernergebnissen* nur hypothetisch auf der Ebene des internationalen Vergleichs hergestellt werden konnte, untersuchen die DFG-Projekte sehr differenziert die *Wirkungen des Unterrichts*. Dabei wird schulisches Lernen multikriterial verstanden, d.h. neben den kognitiven Lernergebnissen stehen durchweg auch motivationale Veränderungen und Auswirkungen auf die Selbstregulationsfähigkeit der Schüler im Blickpunkt (vgl. insbesondere den Beitrag von Prenzel u.a.). Zudem wird in allen Projekten Wert darauf gelegt, kognitive und motivationale Prozesse über einen längeren Zeitraum, etwa ein ganzes Schuljahr hinweg zu beobachten, um nicht nur Statusvergleiche (die kausal kaum interpretierbar sind) ziehen, sondern unterschiedliche Zuwächse in Wissen, Verständnis, Interesse usw. erklären zu können.

4. Merkmale des Unterrichts selbst werden ebenfalls differenziert, aus unterschiedlichen Perspektiven und mit verschiedenen Methoden erfasst. In Anlehnung an Stigler u.a. wurden in mehreren Projekten „Basiskodierungen“ vorgenommen, bei denen beispielsweise zwischen Klassengespräch und Formen der Gruppenarbeit unterschieden wird; hinzu kommen hochinferente Ratings zu Unterrichtsqualität und Schülerbefragungen zu Unterrichtsmerkmalen, wie sie ähnlich in der deutschen Erweiterung der TIMS-Videostudie verwendet wurden (Clausen 2000). Insbesondere die Beiträge von Fischer u.a. sowie Prenzel u.a. geben einen guten Einblick in das Methodeninventar der empirischen Unterrichtsforschung.

5. Durch die systematische Verknüpfung von Ergebnisvariablen (s. Punkt 3) und Unterrichtsvariablen (Punkt 4) finden diese Arbeiten *Anschluss an die empirische Schulforschung*, die unter Unterrichtsqualität „jedes stabile Muster von Instruktionsverhalten (versteht), das als Ganzes oder durch einzelne Komponenten die substanzielle Vorhersage und/oder Erklärung von Schulleistung erlaubt“ (Weinert/Schrader/Helmke 1989). Neben klassischen Dimensionen wie Strukturiertheit und Klarheit des Lehrerhandelns, effizienter Klassenführung, Klassenklima (Qualität der Lehrer-Schüler-Beziehung) und Adaptivität des Lehrerverhaltens betonen die DFG-Projekte eine Dimension von Unterrichtsqualität, die sich in den Zusatzuntersuchungen zu TIMS-Video (Klieme/Schümer/Knoll 2002) als für die Lernentwicklung zentral erwiesen hat: die so genannte *kognitive Aktivierung* von Schülern, die sowohl in selbstgesteuerten, kooperativen Lernprozessen als auch im Unterrichtsgespräch realisiert werden kann. Ein verbindendes Ziel der DFG-Projekte ist es, das Konzept eines kognitiv aktivierenden Unterrichts theoretisch und empirisch klar herauszuarbeiten und von scheinbar ähnlichen, aber nicht bedeutungsgleichen Konzepten wie „aktivierenden Lernformen“ abzugrenzen.

6. Hinter dieser Abgrenzung verbirgt sich ein Grundproblem der Unterrichtsforschung: Die *Unterscheidung zwischen Oberflächen- und Tiefenstruktur* des Lehr-Lernprozesses. Während für Stigler u.a. noch die unmittelbar beobachtbaren Instruktionsmethoden, Sozialformen usw. im Vordergrund standen, zielen Fischer u.a. auf „Lehrzieltypen“ und „Inhaltshandlungen“, Fischler u.a. auf „umfassende Situationsklassen“, Prenzel u.a. auf „Gelegenheitsstrukturen“. Diese unterschiedlichen Begriffe spiegeln den Versuch, hinter der Oberfläche von Mustern der Unterrichtsorganisation die Tiefenstruktur der Instruktions- und Lernhandlungen zu identifizieren, von der die Qualität der Auseinandersetzung mit dem Lerngegenstand – eben die kognitive Aktivierung der Schülerinnen und Schüler – maßgeblich abhängt. Die Autoren berufen sich dabei u.a. auf die Unterscheidung zwischen Oberflächenstrukturen und sog. Basismodellen des Unterrichts, die Oser/Patry (1990) eingeführt haben, deren empirische Validierung jedoch bislang aussteht.

7. Auf der Ebene der Tiefenstrukturen lässt sich auch das *Konzept der Unterrichtsskripts* neu und vielleicht bedeutungshaltiger fassen, als es bei Stigler u.a. angelegt ist. Die dort identifizierten kulturspezifischen Unterrichtsskripts sind nicht wirklich qualitativ als Idealtypen rekonstruiert und auch nicht quantitativ-statistisch definiert worden; Stigler und Kollegen fassten hingegen Einzelmerkmale ihrer Unterrichtsbeobachtungen zu einer narrativen Beschreibung typischer Muster zusammen. Die Projekte des

Schwerpunktprogramms stehen vor der Aufgabe, den Skript-Begriff systematischer und gehaltvoller zu fassen.

8. Nimmt man den Begriff der Skripts ernst, wie er in der kognitiven Psychologie gebraucht wird, so muss man hierunter Wissen über typische Handlungsverläufe verstehen, das von Personen, die häufig in entsprechende Situationen involviert sind, geteilt wird. Auf diesem Hintergrund würde man Unterrichtsskripts rekonstruieren, indem man das Wissen von Lehrern und Schülern über typische Unterrichtsverläufe erhebt. In der Tat spielt die theoretische Beschreibung und empirische Erfassung von *unterrichtsbezogenen Kognitionen*, insbesondere bei Lehrern, in den hier vorgestellten Forschungsprojekten eine besondere Rolle (vgl. insbesondere Diedrich/Thußbas/Klieme sowie Fischler u.a.). Dabei geht es nicht nur um die Identifizierung von Unterrichtsskripts, sondern auch um den Versuch, Unterrichtsqualität zumindest partiell durch die Qualität des professionellen Wissens von Lehrern zu erklären.

9. Mit dem Wissen und den Einstellungen von Lehrern untersuchen die Projekte einen wichtigen Ansatzpunkt für *Verbesserungen der Schulpraxis*. Das Projekt Fischler u.a. hat durch Coaching von Lehrpersonen bereits begonnen, konkrete Interventionsmaßnahmen umzusetzen und zu evaluieren.

In den nachfolgenden Beiträgen können nur erste Ergebnisse des auf insgesamt sechs Jahre angelegten Forschungsprogramms wiedergegeben werden. Gerade die Erhebung, Dokumentation und Auswertung von Unterrichtsvideos ist ein äußerst arbeitsaufwändiger Prozess, für den ein DFG-Schwerpunktprogramm die optimale Arbeitsumgebung darstellt. Durch Verknüpfung mit längsschnittlichen und kontextbezogenen Daten werden sehr komplexe Datenstrukturen erzeugt, die in den meisten Fällen noch ausgewertet werden müssen. Eine wesentliche Fragestellung wird etwa darin bestehen, Zusammenhänge zwischen Qualitätsmerkmalen und Wirkungen des Unterrichts nicht nur global für ganze Lerngruppen bzw. Schulklassen zu untersuchen, sondern differenziell für unterschiedliche Schüler mit je individuellen Lernvoraussetzungen. Eine weitere Herausforderung wird darin bestehen, über erste, relativ globale Wirkungsmaße hinaus differenzierende Wirkungsprofile zu untersuchen, also beispielsweise zwischen dem Zuwachs an Faktenwissen und einem tiefen, qualitativen Verständnis als kognitiven Resultaten zu unterscheiden. Eine derartig differenzierende Forschung hätte eine gute Chance, die häufig eher ideologisch geführten Debatten über „konstruktivistisch orientierten“ Unterricht empirisch zu fundieren.

Literatur

- Clausen, M. (2000). Wahrnehmung von Unterricht – Übereinstimmung, Konstruktvalidität und Kriteriumsvalidität in der Forschung zur Unterrichtsqualität. Phil. Diss, FU Berlin, Fachbereich Erziehungswissenschaft und Psychologie.
- Klieme, E./Bos, W. (2000): Mathematikleistung und mathematischer Unterricht in Deutschland und Japan: Triangulation quantitativer und qualitativer Forschungsansätze im Rahmen der TIMS-Studie. In: Zeitschrift für Erziehungswissenschaft, 3(3), S. 359–379.

- Klieme, E./Schümer, G./Knoll, S. (2001): Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung im internationalen Vergleich. In: Klieme, E. & Baumert, J. (Hrsg). TIMSS – Impulse für Schule und Unterricht, Bonn: Bundesministerium für Bildung und Forschung, S. 43–57.
- Oser, F./Patry, J.-L. (1990): Choreographien unterrichtlichen Lernens, Basismodelle des Unterrichts. In: Berichte zur Erziehungswissenschaft Nr. 89, Pädagogisches Institut der Universität Fribourg, Schweiz.
- Stigler, J.W./Gonzales, P./Kawanaka, T./Knoll, S./Serrano, A. (1999): The TIMSS Videotape Classroom Study: Methods and Preliminary Findings. Report prepared for the National Center for Education Statistics, U.S. Department of Education. Los Angeles.
- Weinert, F.E./Schrader, F.W./Helmke, A. (1989): Quality of instruction and achievement outcomes. In: International Journal of Educational Research 13, S. 895–914.

Anschrift des Autors:

Prof. Dr. Eckhard Klieme, Deutsches Institut für Internationale Pädagogische Forschung. Arbeitseinheit Bildungsqualität und Evaluation, Schloßstr. 29, 60486 Frankfurt a.M.

Martina Diedrich/Claudia Thußbas/Eckhard Klieme

Professionelles Lehrerwissen und selbstberichtete Unterrichtspraxis im Fach Mathematik¹

1. Einleitung

Die PISA-Studie (Baumert u.a. 2001) hat unlängst den aus TIMSS (Baumert u.a. 1997) bekannten Befund bestätigt, dass deutsche Schüler im internationalen Vergleich lediglich unterdurchschnittliche bis mittelmäßige Mathematikleistungen erbringen. Für die Frage nach möglichen Ursachen und pädagogischen Konsequenzen sind sicherlich differenzielle Analysen noch wichtiger, wonach die deutschen Schüler bei komplexeren Problemlöseleistungen besonders schwach abschneiden, während ihre relative Stärke bei Routineaufgaben liegt (vgl. Klieme/Bos 2001 zur TIMS-Mittelstufenstudie, Klieme/Baumert zur TIMS-Oberstufenstudie und Klieme/Neubrand/Lüdtke 2001 zu PISA). Dieses ungünstige Leistungsprofil legt die Vermutung nahe, dass die „Unterrichtskultur“ in Deutschland das problemlösende Denken zu wenig fördert. In der Tat konnte innerhalb der TIMS-Videostudie die Dominanz eines eng geführten, kleinschrittigen fragend-entwickelnden Mathematikunterrichts nachgewiesen werden (vgl. Klieme/Bos 2001). Zugleich wurde belegt, dass kognitive Aktivierung der Schüler, beispielsweise über anspruchsvolle Übungsaufgaben und Fehlerdiskussionen, die Leistungsentwicklung fördert (Klieme/Schümer/Knoll 2001).

Das DFG-Projekt, aus dem hier berichtet wird, schließt daran an. Es untersucht, wie sich das Verständnis geometrischer Konzepte und die Fähigkeit der Schülerinnen und Schüler, algebraische Anwendungsaufgaben zu lösen, in unterschiedlichen Unterrichtskontexten entwickeln (Klieme 1999). Die zentrale Hypothese lautet, dass Intensität und Tiefe des Austauschs von Argumenten – sei es im Unterrichtsgespräch oder in der kooperativen Schülerarbeit – eine zentrale Rolle spielen für das Ausmaß der kognitiven Aktivierung und somit für die Entwicklung eines tiefen Verständnisses mathematischer Begriffe. Allerdings wird auch angenommen, dass diese sog. Diskursivität der Unterrichtsführung nicht für alle Schüler gleich hilfreich ist. Aufgrund der Forschungen zur Aptitude-Treatment-Interaktion lässt sich vermuten, dass vom diskursiven Unterricht eher die Schüler mit höheren Eingangsfähigkeiten profitieren.

Als eine weitere Determinante des Unterrichtsgeschehens werden die kognitiven Voraussetzungen der Lehrpersonen in den Blick genommen. Im ersten Schritt wurden

1 Die Studie wurde gefördert durch Mittel der DFG (KL 1057/3-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

Mathematiklehrer in Deutschland und in der Schweiz schriftlich befragt mit dem Ziel, verschiedene Facetten des *professionellen Lehrerwissens* in Anlehnung an Bromme (1992, 1997) und Shulman (1987) differenziert zu diagnostizieren, ihre Binnenstruktur sowie ihre Zusammenhänge mit der unterrichtlichen Praxis zu untersuchen. Der vorliegende Aufsatz stellt Hintergrund und Ergebnisse dieser Befragung dar.²

2. Theoretischer Hintergrund

2.1 Unterrichtsqualität und Lehrerwissen

Die Fokussierung des Instruktionsverhaltens von Lehrern als maßgeblichem Bestimmungstück von Unterrichtsqualität ist bereits in der Begriffsdefinition angelegt, wie sie etwa Weinert, Schrader und Helmke (1989, S. 899) vorgelegt haben. Unterrichtsqualität wird demnach verstanden als „jedes stabile Muster von Instruktionsverhalten, das als Ganzes oder durch einzelne Komponenten die substantielle Vorhersage und/oder Erklärung von Schulleistung erlaubt“. Der von uns gewählte Zugang zum unterrichtlichen Verhalten von Lehrern versucht, dieses Verhalten wiederum auf das Wissen zurückzuführen, das im Laufe der beruflichen Entwicklung erworben wird (sog. Novizen-Experten-Paradigma; vgl. z.B. Fennema/Carpenter/Peterson 1989; Bromme 1992, 1997). Der Begriff des Wissens wird in dieser Forschungstradition in einem weiten Sinne verstanden; er umfasst verschiedene Arten von subjektiven Kognitionen wie beispielsweise auch Handlungsroutrinen (prozedurales Wissen).

Mit dem Wissen sind insbesondere auch affektive und evaluativ-normative Aspekte eng verbunden. In jüngerer Zeit wird hierfür verstärkt der Begriff der *Beliefs* (wörtlich übersetzt: Überzeugungen) verwendet (vgl. dazu Schoenfeld 1983; Nespor 1987; Pajares 1992; Fischler 2000). Er entstammt ursprünglich der Kognitionsforschung, die *Beliefs* neben Wissen und Werthaltungen als Komponenten von Einstellungen begreift. *Beliefs* sind danach immer evaluativ geprägt, lassen sich jedoch durch faktische Hinweise untermauern (Aiken 1996). Die Unterscheidbarkeit zwischen kognitiven Aspekten im engeren Sinne von evaluativen Aspekte (*Beliefs*) wird bis heute in der Literatur kontrovers diskutiert (vgl. Abelson 1979; Fennema/Loef Franke 1992; Pejares 1992). Ein weiterer, auch praktisch sehr wichtiger Gegenstand der wissenschaftlichen Debatte ist die Frage, wie stark und eindeutig das tatsächlich beobachtbare Verhalten durch *Beliefs* und/oder Wissen gesteuert wird. Sowohl in theoretischen Arbeiten (z.B. zu Lehrer-*Beliefs* bei Schoenfeld 1983 oder zum verwandten Thema *Subjektive Theorien* bei Groeben/Wahl/Schlee/Scheele 1988) als auch in praktischen Interventionsansätzen (z.B. *Cognitively Guided Instruction*, Fennema/Carpenter/Peterson 1989; Fischler u.a., in diesem Band) wird die Handlungsrelevanz von subjektiven Kognitionen unterstellt.

2 Das Forschungsprojekt wird gemeinsam mit Prof. Dr. Kurt Reusser und Dr. Christine Pauli (Universität Zürich) durchgeführt. Wir danken den Züricher Kollegen für die Zusammenarbeit bei der Konzipierung der Studie und für die Erhebung der Lehrerdaten in der Schweiz.

Die Arbeitsgruppe um Fennema vertritt die Auffassung, dass dem Wissen der Lehrer die zentrale Rolle bei der Aufklärung von Unterschieden im Instruktionsverhalten zukommt (Peterson/Carpenter/Fennema 1989; Fennema/Loef Franke 1992). Für die empirische Forschung wurde eine Skala entwickelt, die es erlaubt, ein konstruktivistisches, d.h. selbstregulative Prozesse betonendes Verständnis des mathematischen Lernens (vgl. unten Abschnitt 2.4) von einer eher rezeptiven Auffassung zu unterscheiden (Fennema/Carpenter/Loef 1990). Staub und Stern (in press) haben diese Skala für eine Studie zum mathematischen Unterricht an deutschen Grundschulen adaptiert. Sie konnten etwa ein Drittel der Varianz im Leistungszuwachs der Schüler durch die handlungsleitenden Vorstellungen der Lehrer aufklären.

2.2 Modelle professionellen Lehrerwissens

Auch aus Gründen sprachlicher Klarheit wurde in vorliegender Studie in Anlehnung an Shulman (1987) und Bromme (1992, 1997) der Begriff des *professionellen Lehrerwissens* für die kognitiven Voraussetzungen unterrichtlichen Handelns verwendet. Dabei wird betont, dass nicht nur rein kognitive Aspekte, sondern auch eine evaluativ-normative Komponente, die ursprünglich stärker mit dem Beliefbegriff assoziiert ist, eingeschlossen wird. Professionelles Lehrerwissen im Sinne der hier vorzustellenden Untersuchung umfasst somit die Gesamtheit der kognitiven und evaluativ-normativen, damit auch bedingt affektiven Voraussetzungen, welche die individuelle Instruktionspraxis beeinflussen und erklären können.

Neben der Frage des Erklärungspotenzials von professionellem Lehrerwissen für die unterrichtliche Praxis interessiert insbesondere, welche Facetten dieses Konstrukts identifiziert werden können. Bei einer solchen Explikation kann auf verschiedene Ansätze unterschiedlichen Abstraktionsgrads zurückgegriffen werden. Die am stärksten differenzierende Taxonomie findet sich bei Shulman (1987), die in der Folge von Bromme (1992, 1997) weiter ausgeführt wurde. Bromme unterscheidet fünf inhaltliche Kategorien:

- *Allgemeines Pädagogisches Wissen* betrifft u.a. disziplinarische Fertigkeiten, Classroom Management sowie Vorstellungen über Bedingungsfaktoren des schulischen Lernerfolgs.
- *Fachliches Wissen* bezieht sich auf Methoden und Erkenntnisse der Referenzdisziplin des Schulfaches wie z.B. der Mathematik.
- *Curriculares Wissen* beinhaltet vor allem Vorstellungen über Ziele des Unterrichts, Auswahl und Sequenzierung von Lerninhalten.
- *Philosophie des Schulfaches*: Damit bezieht Bromme sich auf die Vorstellungen darüber, „wofür der Fachinhalt nützlich ist und in welcher Beziehung er zu anderen Bereichen menschlichen Lebens und Wissens steht“ (ebd., S. 197).
- *Fachspezifisch-pädagogisches Wissen*: Hiermit ist vor allem fachdidaktisches Wissen gemeint, also das Wissen um die angemessene Darstellung und Vermittlung der jeweiligen Unterrichtsinhalte.

Bromme (1992) hebt die kognitive Verknüpfung dieser unterschiedlichen Wissensbereiche hervor, die nicht unverbunden nebeneinander stehen, sondern im Verlaufe der Erfahrung und in Bezug auf die Praxis integriert werden.

Neben dieser stark differenzierenden Taxonomie wird häufig eine Dreiteilung der Domäne des professionellen Lehrerwissens vorgeschlagen. Es wird unterschieden zwischen dem (1) Wissen über das *Wesen* der Wissenschaft bzw. des Faches, das gelehrt wird, dem (2) Wissen über das *Lehren* und dem (3) Wissen über das *Lernen* des jeweiligen Faches. Eine solche Position findet sich u.a. bei Aguirre, Haggerty und Linder (1990), bei Fischler (2000) oder bei Ernest (1988). Eine noch weitergehende Reduktion der Dimensionalisierung professionellen Lehrerwissens findet sich bei Köller, Baumert und Neubrand (2000). Sie fassen die beiden Bereiche des Lehrens und Lernens zusammen und unterscheiden fortan Auffassungen über die „Struktur des Wissens“ einerseits und die „Struktur der Wissenserzeugung“ andererseits (ebd. S. 232).

Bei der Operationalisierung professionellen Lehrerwissens empfiehlt sich die Verwendung der Brommeschen Konzeption, da in ihr relevante Wissenskategorien am deutlichsten differenziert werden. Nichtsdestoweniger ist zu berücksichtigen, dass sich die fünf Kategorien – aufgrund ihrer engen kognitiven Verknüpfung – vermutlich den drei bzw. zwei Dimensionen anderer Modelle unterordnen lassen, und es wird Gegenstand der empirischen Analyse sein aufzuzeigen, inwiefern eine empirische Trennung tatsächlich möglich und sinnvoll ist. Im Folgenden sollen die beiden zuletzt genannten Komponenten der Brommeschen Taxonomie näher diskutiert werden.

2.3 „Philosophie des Schulfaches Mathematik“

Köller u.a. (2000) wie auch Grigutsch, Raatz und Törner (1998) haben verschiedene „mathematische Weltbilder“ beschrieben und in Fragebogenskalen umgesetzt, die letztlich das beschreiben, was nach Bromme die „Philosophie des Schulfaches“ ausmacht.

Köller u.a. (2000) bezeichnen mathematische Weltbilder in Anlehnung an Schoenfeld (1983) als „individuelle intuitive Theorien über das Wesen der Mathematik“ (S. 229). Sie unterscheiden (a) *Mathematik als kreatives Sprachspiel*, (b) *Mathematik als Entdecken eines finiten Kosmos von Ideen*, (c) *Schemaorientierung*, und (d) *Instrumentelle Relevanz von Mathematik*. Aufgrund der korrelativen Zusammenhänge, die bei Oberstufenschülern im Rahmen der TIMS-Studie ermittelt wurden, fassen die Autoren diese vier Konzeptionen zu Faktoren zweiter Ordnung zusammen, die als statisches (b und c) bzw. dynamisches Bild (a und d) von Mathematik interpretiert werden.

Grigutsch u.a. (1998) konnten in einer empirischen Untersuchung bei Lehrern faktorenanalytisch vier Skalen von jeweils zwölf Items bilden, die folgende Aspekte der Mathematik repräsentieren:

- *Formalismusaspekt*: Mathematik ist gekennzeichnet durch begriffliche und sprachliche Strenge und Exaktheit.

- *Anwendungsaspekt*: Mathematik weist einen hohen praktischen Nutzen auf, sie ist für das spätere Leben der Schüler wichtig.
- *Prozessaspekt*: Mathematik wird als Tätigkeit des Nachdenkens über Probleme begriffen. Betont werden Erschaffung, Erfindung und Nacherfindung, gleichzeitig aber auch das Verstehen und Einsehen von Zusammenhängen.
- *Schemaaspekt*: Mathematik ist eine Sammlung von Verfahren und Regeln (Mathematik als „Werkzeugkasten und Formelpaket“; Grigutsch u.a., S. 19).

Analog zur Analyse von Köller u.a. (2000) spiegeln Schema- und Formalismusaspekt eine statische Sicht von Mathematik als in sich geschlossenem System, während Prozess- und Anwendungsaspekt eine dynamische Sicht von Mathematik formen.

2.4 *Fachspezifisches pädagogisches Wissen: konstruktivistische vs. rezeptive Konzeption des mathematischen Lernens*

Fachspezifisches pädagogisches Wissen beinhaltet Vorstellungen über die „Erzeugung“ von mathematischem Verständnis durch den Schulunterricht. Angesichts der hohen Popularität so genannter *konstruktivistischer* Unterrichtskonzeptionen (Duffy/Jonassen 1992; Dubs 1995) stellt sich aktuell die Frage, welche Rolle solche Positionen im Denken der Lehrpersonen spielen. Zu einer konstruktivistischen Sicht des Mathematik-Lernens gehört, dass Wissen nicht einfach auf dem Wege der Instruktion in die Köpfe der Lernenden „übertragen“ wird, sondern das Ergebnis eines aktiven Konstruktionsprozesses darstellt, der kulturell und sozial vermittelt ist (Cobb 1994; Ernest 2000). „Konstruktivistisch orientierte“ Lernumgebungen zielen – bei aller Unterschiedlichkeit der epistemologischen und lerntheoretischen Grundlagen – darauf ab, die kooperative und selbstregulative Auseinandersetzung der Schüler mit den Lerninhalten in authentischen Anwendungskontexten zu fördern (Dubs 1995). Eine der offenen Forschungsfragen ist, inwieweit solche Lernumgebungen für verschiedene Schülergruppen unterschiedlich gut geeignet sind.

Zu vermuten ist, dass die Vorstellungen, die Lehrer zum mathematischen Lernen entwickeln, mit ihrem Bild vom Fach selbst zusammenhängen. Wer mathematisches Wissen nicht mehr als „platonisches“ Objekt betrachtet, sondern als kulturelles Produkt, wird vermutlich auch der Re-Konstruktion dieses Wissens durch die Schüler hohe Bedeutung einräumen.

2.5 *Forschungsfragen*

Mithilfe einer standardisierten Befragung wollten wir die Struktur professionellen Lehrerwissens sowie dessen systematische Zusammenhänge mit der unterrichtlichen Praxis im Fach Mathematik analysieren. Folgende Forschungsfragen waren dabei handlungsleitend:

- 1) Ist es möglich, das von Bromme (1992, 1997) entwickelte Modell professionellen Lehrerwissens zu operationalisieren und jedem seiner theoretisch spezifizierten Komponenten ein empirisches Korrelat zuzuordnen?
- 2) Können bei Mathematiklehrern komplexe Überzeugungsstrukturen identifiziert werden, die sich in Übereinstimmungen zwischen den Auffassungen über das Wesen sowie über das Lehren und Lernen von Mathematik ausdrücken? Konkret interessierte uns, ob zwei Syndrome mathematischer Grundüberzeugungen gefunden werden können: zum einen eine dynamische Sicht auf die Mathematik, die einhergeht mit einer eher konstruktivistischen Lerntheorie und der Überzeugung, dass Unterricht die Schüler vor allem zu selbstständigem Lernen aktivieren muss, und zum Zweiten einer eher statischen Sicht, die formalistische und schematisierte Aspekte der Mathematik sowohl in ihrer Theorie als auch in ihrer Didaktik betont.
- 3) Falls solche Einstellungssysteme aufgezeigt werden können: In welchem Zusammenhang stehen sie mit der konkreten, von den Lehrern im Fragebogen berichteten Unterrichtspraxis? Ist eine dynamische Sicht von Mathematik verbunden mit einem hohen Grad der Aktivierung von Schülern, etwa durch Diskursivität und anspruchsvolle Aufgaben? Diese Forschungsfrage stellt gleichzeitig einen Ausgangspunkt für Hypothesen einer geplanten Videostudie dar.

3. Methode

Zur Erfassung der theoretisch explizierten Facetten professionellen Lehrerwissens wurde ein Fragebogen entwickelt, der teilweise auf standardisierte, bereits erprobte Verfahren zurückgreift. Anhand zweier Inhaltsbereiche (Satzgruppe des Pythagoras und mathematische Textaufgaben) wurden auch Angaben zur Unterrichtspraxis in der 9. Jahrgangsstufe erfragt. Schließlich wurden schulische Rahmenbedingungen, Selbstwirksamkeitsüberzeugungen sowie Angaben zur Berufsbiographie erhoben. Im Folgenden wird lediglich über die für die Fragestellung relevanten Verfahren berichtet.

3.1 Erfassung professionellen Lehrerwissens

Unter Bezug auf das bereits dargestellte Modell von Bromme (1992, 1997) wurden fünf Facetten professionellen Lehrerwissens operationalisiert. Die Erfassung *fachlichen Wissens* von Mathematiklehrern in einer postalischen Befragung ist sicherlich nicht einfach. Als proximalen Indikator haben wir erhoben, wie viele unterschiedliche Informationsquellen bei der Unterrichtsvorbereitung zum Thema „Pythagoras“ genutzt werden.

Zur Erfassung des *curricularen Wissens* wurden 16 mögliche Ziele für den Einsatz mathematischer Textaufgaben in der 9. Klasse vorgegeben. Als Indikator wurde die mittlere Zustimmung (auf einer vierstufigen Likert-Skala) zu jenen Items ermittelt, die deutlich auf anspruchsvollen prozess- und problemorientierten Unterricht abheben (Itembeispiel: „Mathematische Textaufgaben setze ich im Wesentlichen mit dem Ziel ein, neue Regeln bzw. Konzepte eigenständig entwickeln zu lassen“).

Zur Operationalisierung der *Philosophie des Schulfaches* wurden aus dem Instrument von Grigutsch u.a. (1998) die jeweils fünf ladungsstärksten Items der Facetten Anwendungsorientierung, Prozess-, Schema- und Formalismusorientierung zusammengestellt. Auf der Grundlage der Voruntersuchung wurden weitere vier Items ausgeschlossen, so dass in der endgültigen Fassung eine Skala von 16 Items (Format Zustimmung, vierstufig) resultierte.

Mit dem *allgemeinen pädagogischen Wissen* werden vom Fach weitgehend unabhängige Aspekte von Lehrerkognitionen i.S. von Kausalattributionen erfasst. Dabei wurde unterschieden, ob die Befragten ein Ge- bzw. Misslingen ihres Unterrichts eher auf äußere Rand- und Rahmenbedingungen, auf schülerbezogene Variablen oder auf die eigene Person zurückführen.

Das *fachspezifisch-pädagogischen Wissen* wurde durch die in Abschnitt 2.1 dargestellten Fragebogenskalen von Fennema, Carpenter und Loef (1990) in der deutschsprachigen Adaptation von Staub und Stern (in press) nach einer weiteren Bearbeitung durch Reusser, Pauli und Staub erhoben. Die verbliebenen 18 Items wurden mit der Besonderheit vorgegeben, dass zwischen eher leistungsstarken und eher leistungsschwachen Schülern differenziert werden konnte. Damit sollte der Erwartung Rechnung getragen werden, dass auch aus Lehrersicht konstruktivistische Lernprinzipien eher für leistungsstarke Schüler geeignet sind.

3.2 Indikatoren für die Unterrichtspraxis

Am Beispiel des Themas „Satzgruppe des Pythagoras“ wurden drei Indizes zur eigenen, *selbstberichteten* Unterrichtspraxis gebildet:

- *Lehrerzentrierte Behandlung*: Es wurden mehrere Aussagen zu einem Index zusammengefasst, die eine Tendenz zum lehrerzentrierten Vorgehen, z.B. im Lehrervortrag, anzeigen.
- *Diskursivität*: Andere Aussagen des Fragebogens beschreiben Vorgehensweisen beim Satz des Pythagoras, die den wechselseitigen Austausch der Schüler und eine weitgehend selbst gesteuerte Form der Auseinandersetzung unterstützen. Derartige Items wurden in einem Summenindex zusammengefasst.
- *Intensität der Aufgabennutzung*: Den Befragten wurde eine Liste unterschiedlicher Aufgaben zur Behandlung der Satzgruppe des Pythagoras vorgegeben, bei welcher die Häufigkeit ihrer Verwendung im Unterricht interessierte. Dahinter stand die Überlegung, dass ein intensiver Einsatz von unterschiedlichen, in ihrem Anforderungsgehalt variierenden Aufgaben insgesamt zu einer stärkeren Beteiligung bzw. Aktivierung der Schüler führt.

3.3 Stichprobe und Durchführung

Die postalische Befragung erfolgte in zwei repräsentativen Stichproben Deutschlands und der Schweiz³. Sie wurde im Herbst des Jahres 2001 durchgeführt. Die Grundgesamtheit in der Schweiz umfasste dabei alle öffentlichen bzw. staatlich subventionierten Schulen in den Kantonen Bern und Zürich, in Deutschland staatliche und private Schulen aus den Bundesländern Brandenburg, Berlin und Baden-Württemberg. An die Schulleitungen der zufällig ausgewählten Schulen wurden Schreiben versandt mit der Bitte, den Fragebogen an eine Lehrperson in der Schule weiterzugeben, die im aktuellen Schuljahr in der 8. (Schweiz) bzw. 9. Jahrgangsstufe (Deutschland) Mathematik unterrichtet. Für den Fall, dass nach diesem Kriterium mehrere Lehrpersonen an einer Schule in Frage kämen, wurde um die Einhaltung eines zufälligen Auswahlverfahrens gebeten. Aus der deutschen Stichprobe wurden 146 Fragebögen ausgefüllt zurückgesandt, was einer Rücklaufquote von 58,4% entspricht. Die durchschnittliche Bearbeitung betrug 2.5 Stunden. In der Schweiz antworteten 116 (46,4%) der Befragten mit einer durchschnittlichen Bearbeitungszeit von 2.3 Stunden.

Hinsichtlich demographischer bzw. berufsbiographischer Merkmale bestanden kaum Unterschiede zwischen den schweizerischen und deutschen Lehrpersonen, die sich an der Untersuchung beteiligt haben. Alter, Dauer der Unterrichtstätigkeit, aber auch die Erfahrung im Unterrichten von Mathematik sind in den beiden Stichproben annähernd gleich. Von den antwortenden deutschen Lehrern unterrichteten 34% am Gymnasium, 30% an einer Realschule, 19% an einer Hauptschule und 17% an Gesamtschulen. 37,2% der Befragten waren Frauen, 62,8% Männer; das durchschnittliche Alter lag bei 46,9 Jahren. Das Durchschnittsalter in der schweizerischen Stichprobe betrug 48,6 Jahre; Frauen waren zu 14,7% vertreten, Männer zu 84,5%. 14% der Schweizer Lehrer unterrichteten an einem Untergymnasium, 64% an einer Sekundarschule (die in etwa der deutschen Realschule entspricht) und 20% an einer Realschule (die der deutschen Hauptschule vergleichbar ist). Da sich auch im Hinblick auf die Binnenstruktur der Lehrerkognitionen keine signifikanten Unterschiede gezeigt haben, wird in der weiteren Darstellung auf eine Trennung der beiden Stichproben verzichtet.

4. Ergebnisse

Die Ergebnisdarstellung folgt einer dreiteiligen Gliederung: Zunächst werden deskriptive Analysen für die beschriebenen Skalen und Indizes professionellen Lehrerwissens und der selbstberichteten Unterrichtspraxis berichtet. Danach wird deren Interdependenz dargestellt. Für die abschließende Analyse der Beziehungen zwischen Kognitionen und Unterrichtsvariablen wird eine Auswahl der beschriebenen Skalen getroffen.

3 Wir danken Erich Ramseier (Bern) und Heiko Sibberns (Hamburg) für die Ziehung der Stichproben sowie den Behörden in Berlin, Brandenburg und Baden-Württemberg für ihre Unterstützung.

4.1 Skalen zum professionellen Lehrerwissen

Mathematische Weltbilder

Zur Überprüfung der Dimensionalität der Items, die mathematische Weltbilder erfassen, wurde eine explorative Hauptkomponentenanalyse durchgeführt. Als Kriterium zur Bestimmung der angemessenen Faktorzahl diente der Scree-Test, der erwartungsgemäß vier Faktoren anzeigte. Die wiederholte Durchführung der Analyse mit der Vorgabe, vier Faktoren zu extrahieren, zeigte eine klare Bestätigung der in der Arbeit von Grigutsch u.a. (1998) gefundenen Faktorstruktur. Somit ließen sich Prozess-, Anwendungs-, Schema- und Formalismusaspekt unterscheiden (die vier Faktoren klärten 56,7% der Gesamtvarianz auf). Die Reliabilitätsanalyse der jeweils zu einem Aspekt gehörenden Items erlaubte ihre Aggregation zu vier Skalen (s. Tabelle 1). Da unsere Studie einen Schwerpunkt auf ein konstruktivistisches Verständnis von Lehr- und Lernprozessen legt, wird in den nachfolgenden Analysen nur noch ein prozessorientiertes mathematisches Weltbild berücksichtigt.

Tab. 1: Skalenkennwerte mathematischer Weltbilder

	Anwendung	Prozess	Formalismus	Schema
N	261	261	261	261
Mittelwert*	3.21	2.92	3.05	2.67
Standardabweichung	.45	.53	.59	.64
Reliabilität (Cronbachs α)	.67 (5 Items)	.72 (5 Items)	.79 (4 Items)	.60 (2 Items)
* Antwortformat vierstufig (1–4)				

Konstruktivistisches vs. rezeptives Mathematikverständnis

Den Befragten war bei dieser Skala die Möglichkeit gegeben, zwischen leistungsstarken und -schwachen Schülern zu differenzieren, was von 225 Personen genutzt wurde (das entspricht etwa 86% der Stichprobe). Die 37 Personen, welche beide Leistungsgruppen in ihren Antworten gleich behandelten, unterschieden sich in demographischer Hinsicht nicht vom Rest; auch konnten keine systematischen Länderunterschiede festgestellt werden. Sie wurden bei den weiteren Analysen ausgeschlossen.

Die Ergebnisse einer explorativen Faktorenanalyse, in die sämtliche zu dieser Skala gehörenden Items mit jeweils beiden Antwortmöglichkeiten eingingen, sprachen deutlich für eine Zwei-Faktorlösung (35,2% der Gesamtvarianz wurden durch diese Faktoren aufgeklärt). Sie unterscheidet Aussagen, die einem konstruktivistischen Mathematikverständnis entsprechen, von jenen Items, die ein rezeptives Mathematikverständnis indizieren. Bis auf zwei Items, die ausgeschlossen wurden, konnte die entsprechende Zuordnung auf der Grundlage der Voruntersuchung beibehalten werden. Es wurden zwei Skalen gebildet, deren Kennwerte Tabelle 2 entnommen werden können.

Tab. 2: **Skalenwerte für rezeptives und konstruktivistisches Verständnis vom Mathematiklernen und -lehren.**

	Konstrukt. MA-Verständnis	Rezeptives MA-Verständnis
N	261	261
Mittelwert*	3.01	2.55
Standardabweichung	.48	.47
Reliabilität (Cronbachs α)	.82 (16 Items)	.89 (20 Items)
* Antwortformat vierstufig (1–4)		

Darüber hinaus wurde untersucht, inwieweit bei gleicher Faktorstruktur Niveauunterschiede zwischen den Angaben zu leistungsstarken und -schwachen Schülern bestehen. Wie Abbildung 1 aufzeigt, scheinen die Befragten deutlich zwischen den beiden Leistungsgruppen zu differenzieren. Während die Befragten für die Gruppe der Leistungsschwächeren eher ein rezeptives Verständnis vom Lernen und Lehren zugrunde legen, beschreiben sie das Lernen und Lehren für leistungsstarke Schüler eher im Sinne einer konstruktivistischen Auffassung.

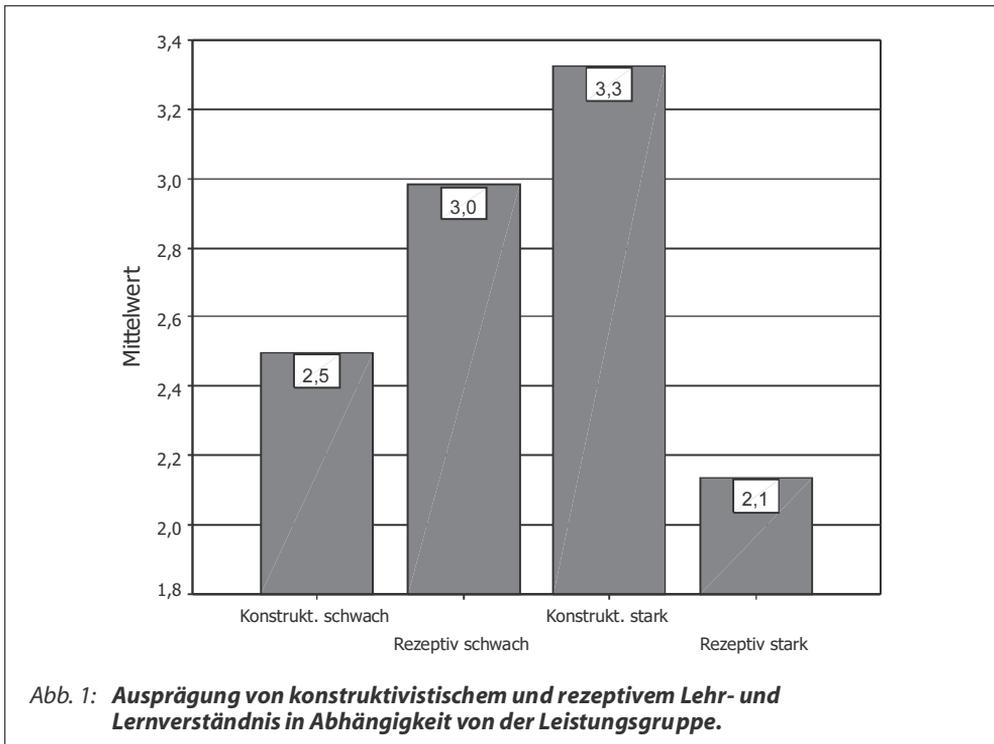


Abb. 1: **Ausprägung von konstruktivistischem und rezeptivem Lehr- und Lernverständnis in Abhängigkeit von der Leistungsgruppe.**

Ziele mathematischer Textaufgaben

Die 16 Items, welche Ziele beim Einsatz von Textaufgaben beschreiben, lassen sich den Dimensionen „Übung“, „Anwendung“, „Aktivierung der Schüler“, „Förderung bestimmter Schülergruppen“ und „Nicht-mathematische Zielsetzungen“ zuordnen. Diese fünf Dimensionen klären gemeinsam 61,1% der Varianz auf. Es wurden entsprechend fünf Skalen gebildet, die jeweils eine Dimension abbilden. Für unsere Fragestellung interessiert von diesen Faktoren in besonderer Weise der Aspekt der Aktivierung, da in ihr eigenständiges, selbstreguliertes Lernen betont wird, was mit der von uns gewählten Definition konstruktivistischen Unterrichts übereinstimmt. Im Folgenden wird deshalb lediglich dieser Zielaspekt weiter berücksichtigt (die Skala weist bei einem vierstufigen Antwortformat einen Mittelwert von $M = 2,89$ und eine Standardabweichung von $s = ,52$ auf).

Fachliches und allgemeines pädagogisches Wissen

Der Schwierigkeit der direkten Erfassung von fachlichem Wissen wurde begegnet, indem ein approximativer Indikator gebildet wurde, der die Anzahl verwendeter Materialien bei der Vorbereitung der Satzgruppe des Pythagoras berücksichtigt. Von sieben möglichen Arbeitsgrundlagen kreuzten die Befragten durchschnittlich 3,7 Möglichkeiten an; die Standardabweichung lag bei $s = 1,36$.

Für den Bereich allgemeinen pädagogischen Wissens wurden die subjektiven Erfolgsthemen über das Gelingen von Unterricht zugrunde gelegt. Es interessierten dabei insbesondere solche potenziellen Ursachen, die in der Person des Lehrers begründet sind. Berücksichtigt wurde deshalb die Nennung von guter Unterrichtsplanung, Zusammenarbeit mit Fachkollegen und der Fähigkeit, auch leistungsschwache Schüler zu motivieren, als ausschlaggebende Gründe für das Gelingen von Unterricht. Die daraus resultierende Skala hat bei einem vierstufigen Antwortformat einen Mittelwert von $M = 3,2$ und eine Standardabweichung von $s = ,50$.

Indikatoren für die Unterrichtspraxis

Zur approximativen Beschreibung der Unterrichtsführung wurden aus den Angaben der Befragten drei Indikatoren gebildet, welche die theoretisch explizierten Aspekte eines prozessorientierten, diskursiven (konstruktivistischen) Lehr- und Lernverständnisses umfassten. Für den ersten Indikator, der spezifisch die diskursiven Elemente der selbstberichteten Unterrichtspraxis beschreiben sollte, wurden die Durchführung von Kleingruppenarbeit bei der Einführung und dem Beweisen des Satzes des Pythagoras sowie deren Erarbeitung im Unterrichtsgespräch berücksichtigt. Die lehrerzentrierte Behandlung der Satzgruppe des Pythagoras (Bevorzugung des Lehrervortrags) wurde als einer diskursiven Auffassung entgegengesetzt verstanden; sie bildete den zweiten Indikator. Des Weiteren wurde die Intensität der Aufgabennutzung als Ausdruck für die Aktivierung der Schüler im Unterricht betrachtet. Es wurde erfasst, wie häufig Textaufgaben unterschiedlichen Anspruchsniveaus im Unterricht eingesetzt werden. Die Verteilungen können Tabelle 3 entnommen werden.

Tab. 3: Kennwerte (oberer Tabellenteil) und Korrelationen (unterer Teil) der Indikatoren für die Unterrichtspraxis

	Diskursive Unterrichts-führung	Lehrerzentriertheit (Satz d. Pyth.)	Intensität der Aufgabennutzung
N	258	252	257
Mittelwert	3.1	.6	16.0
Standardabweichung	1.82	.74	3.46
Minimum	.00	.00	5.00
Maximum	8.00	2.00	24.00
Unterrichtsführung	–		
Lehrerzentriertheit	–.23**	–	
Aufgabennutzung	.23**	–.12	–

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Bei diesen Ergebnissen fällt auf, dass die Befragten stark in ihren Antworten variieren, was sich in vergleichsweise hohen Standardabweichungen der Indikatoren ausdrückt. Während die Lehrer seltener eine diskursive Unterrichtsführung berichten (der Mittelwert liegt mit 3,1 eher im unteren Bereich der Verteilung), drückt sich in den mittleren Antworten zur Lehrerzentriertheit ($M = ,6$) und zur Aufgabennutzung ($M = 16,0$) deutlich ein die Schüler aktivierendes Instruktionsverhalten aus.

4.2 Zusammenhänge zwischen den Indikatoren professionellen Lehrerwissens

Da die Studie neben der Spezifizierung von relevanten Bestimmungsstücken professionellen Lehrerwissens insbesondere deren Binnenrelationen klären wollte, wurden Korrelationen zwischen den einzelnen Skalen berechnet (vgl. Forschungsfrage 2). Die Ergebnisse können Tabelle 4 entnommen werden.

Die Wissensindikatoren weisen deutliche Zusammenhänge in erwartungsgemäßer Richtung auf. So geht die Betonung des Prozessaspekts in der Mathematik mit einer konstruktivistischen Lehr- und Lerntheorie und mit einer hohen Bewertung der Aktivierung der Schüler bei den Unterrichtszielen einher. Auch wird der Erfolg von Unterricht eher der eigenen Person zugeschrieben. Die Daten geben Hinweise darauf, dass ein Syndrom pädagogischer Überzeugungen besteht, das eine dynamische Sicht der Mathematik mit einem konstruktivistischen, die Schüler aktivierenden Lehr- und Lernverständnis vereint und dabei die Eigenverantwortlichkeit für den Unterricht betont. Dagegen bestehen negative Zusammenhänge zwischen einer prozessorientierten, die Schüler aktivierenden Haltung und einem rezeptiven Verständnis vom Mathematiklernen und -lehren.

Tab. 4: **Zusammenhänge zwischen den Indikatoren professionellen Lehrerwissens (N > 250).**

	(1)	(2)	(3)	(4)	(5)	(6)
1) Prozessaspekt	—					
2) Konstruktivistisches MA-Verständnis	.52**	–				
3) Rezeptives MA-Verständnis	–.32**	–.43**	–			
4) Ziele math. Textaufgaben: Aktivierung	.43**	.41**	–.14*	–		
5) N Materialien (S.d.Pyth.)	.11	.15*	–.18**	.13*	–	
6) Kausalattribution: Lehrer	.23**	.34**	.04	.30**	.05	–

* Die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant.
** Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

4.3 Zusammenhänge zwischen professionellem Lehrerwissen und der Unterrichtspraxis

In einem weiteren Schritt wurden Zusammenhänge zwischen den Komponenten des Lehrerwissens und den Indikatoren für die Unterrichtspraxis ermittelt (vgl. Forschungsfrage 3). Damit sollte ansatzweise überprüft werden, inwieweit unterrichtsbezogene Kognitionen auf Unterrichts- bzw. Instruktionsverhalten Einfluss nehmen.

Tab. 5: **Einfluss unterrichtsbezogener Kognitionen auf die selbstberichtete Unterrichtspraxis (angegeben sind die standardisierten β -Koeffizienten).**

	diskursive Unterrichtsführung	Lehrerzentriertheit (Satz d. Pyth.)	Intensität der Aufgabennutzung
Prozessaspekt	0.03	–0.01	0.03
Konstruktivistisches MA-Verständnis	0.14	–0.21*	0.10
Rezeptives MA-Verständnis	–0.09	0.17*	–0.01
Ziele math. Textaufgaben: Aktivierung	0.13*	–0.09	0.25**
N Materialien (S.d.Pyth.)	0.14*	0.05	0.16*
Kausalattribution: Lehrer	0.07	–0.06	0.01
Multiple R	.37	.36	.38

* Der Koeffizient ist auf dem Niveau von 0.05 (2-seitig) signifikant.
** Der Koeffizient ist auf dem Niveau von 0.01 (2-seitig) signifikant.

Zur Überprüfung dieses Zusammenhangs wurden multiple Regressionen berechnet, die den Einfluss der Wissensfacetten auf die berichtete Unterrichtspraxis spezifizieren soll-

ten. Durch die in Tabelle 5 aufgeführten multiplen Regressionskoeffizienten von $R = .37$, $R = .36$ und $R = .38$ werden 14% bzw. 13% der Varianz im jeweiligen Unterrichtsindikator aufgeklärt. Somit bestehen Zusammenhänge zwischen unterrichtsbezogenen Kognitionen und dem selbstberichteten Instruktionsverhalten.

Die Qualität der selbstberichteten Unterrichtspraxis hat vor allem mit jenen Facetten des professionellen Wissens zu tun, die explizit in der Fachdidaktik vermittelt werden. Statistisch signifikante Zusammenhänge finden sich nämlich beim curricularen und beim fachlichen Wissen sowie beim fachspezifischen pädagogischen Wissen, nicht jedoch beim allgemeinpädagogischen Wissen und bei der „Philosophie des Schulfaches“. Je stärker ein Lehrer aktivierende Ziele befürwortet, und je vielfältiger er sich vorbereitet, desto intensiver und vielfältiger nutzt er mathematische Aufgaben, und desto stärker unterstützt er den diskursiven Austausch von Argumenten im Mathematikunterricht. Wer sich mathematisches Lernen eher als rezeptiven Prozess vorstellt, nutzt auch häufiger lehrerzentrierte Methoden, während Lehrer mit konstruktivistischen Ansichten dies seltener tun.

5. Diskussion

Unsere Studie hat in einem ersten Schritt aufgezeigt, dass die Facetten professionellen Lehrerwissens i.S. der von Bromme (1992, 1997) spezifizierten Taxonomie unterscheidbar sind. So wurden seine fünf inhaltlichen Bereiche durch teils standardisierte, teils neue Skalen erfasst. Darüber hinaus hat sich gezeigt, dass diese Facetten interdependent sind, sodass zwei komplexe Syndrome unterrichtsbezogener Kognitionen unterschieden werden können: eine eher dynamische Sichtweise, welche die Mathematik, ihr Lehren und Lernen als kreativen Prozess begreift, und eine eher statische, wonach mathematisches Wissen aus regelhaften Zusammenhängen besteht, die in fest stehenden Formeln beschrieben und eher rezeptiv erlernt werden. Aufbauend auf diesen Ergebnissen wurde versucht, Unterschiede der selbstberichteten Unterrichtspraxis auf der Basis zugrunde liegender unterrichtlicher Kognitionen zu erklären. Es fanden sich deutliche Hinweise darauf, dass die Unterrichtspraxis von den kognitiven Voraussetzungen, dem Wissen und Beliefs der Lehrer, beeinflusst wird.

Zu diskutieren ist die Validität dieser Ergebnisse. Die gewählte Operationalisierung aufseiten der Prädiktoren wäre zum Teil verbesserungsbedürftig, insbesondere im Bereich des fachlichen Wissens (hier nur distal als Anzahl zur Unterrichtsvorbereitung eingesetzter Materialien operationalisiert) sowie des allgemeinen pädagogischen Wissens (hier: Kausalattributionen über das Gelingen von Unterricht). Aufseiten der abhängigen Variablen ist zu fragen, ob die hier verwendeten Indikatoren die Unterrichtspraxis angemessen abbilden. Ihre Auswahl kann jedoch damit begründet werden, dass die berichtete Unterrichtspraxis lediglich auf die Aspekte eingeschränkt wurde, die für (bzw. explizit gegen) ein konstruktivistisches Unterrichtsverständnis bzw. eine diskursive Unterrichtsführung sprechen. Somit wurden bewusst solche Unterrichtsmerkmale gewählt, die sinnvoll in Beziehung zu den Lehrerkognitionen gesetzt werden können.

Des Weiteren wäre zu fragen, inwieweit die schriftliche Form der Lehrerbefragung anfällig für Antworttendenzen ist. So ist es denkbar, dass aufgrund der öffentlichen Forderung nach verändertem, stärker konstruktivistisch orientiertem Mathematikunterricht die Befragten im Sinne sozialer Erwünschtheit antworteten. Diesem Argument kann jedoch entgegengetreten werden, dass wir empirisch ein Erklärungsmodell identifizieren konnten, das interindividuelle Unterschiede erklärt. Eventuelle Antworttendenzen könnten also die mittlere Ausprägung der berichteten Einstellungen beeinflusst haben, nicht jedoch die hier untersuchten Zusammenhänge.

Abschließend ist zu diskutieren, welche Aussagekraft unsere Ergebnisse für die Erklärung und Vorhersage der tatsächlichen Unterrichtspraxis haben. Einen Beleg für die Validität der Selbstauskünfte in der schriftlichen Befragung kann erst die geplante Videostudie geben. Erste Zusammenhänge zwischen Lehrerkognitionen und beobachtetem Instruktionsverhalten in der TIMS-Videostudie (Research Report 1998–2000, S. 95) lassen erwarten, dass unsere Videostudie den Einfluss unterrichtsbezogener Kognitionen auf die Unterrichtspraxis bestätigen wird.

Literatur

- Abelson, R.P. (1979). Differences between belief and knowledge systems. In: *Cognitive Science* 3, S. 355–366.
- Aguirre, J.M./Haggerty, S.M./Linder, C.J. (1990). Student-teachers' conceptions of science, teaching and learning: a case study in preservice science education. In: *International Journal of Science Education* 12, H. 4, S. 381–390.
- Aiken, L.R. (1996). Rating scales and checklists. Evaluating behavior, personality, and attitudes. New York: John Wiley & Sons.
- Baumert, J./Lehmann, R./Lehrke, M./Schmitz, B./Clausen, M./Hosenfeld, I./Köller, O./Neubrand, J. (1997). TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiß, M. (Hrsg.). (2001). PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Bromme, R. (1992). Der Lehrer als Experte. Zur Psychologie des professionellen Wissens. Bern: Hans Huber.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In: Weinert, F. E. (Hrsg.): *Enzyklopädie der Psychologie: Themenbereich D Praxisgebiete. Serie I Pädagogische Psychologie, Band 3 Psychologie des Unterrichts und der Schule*. Göttingen: Hogrefe, S. 177–212.
- Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. In: *Educational Researcher* 23, H. 7, S. 13–20.
- Diedrich, M. (2001). Professionelles Lehrerwissen von Mathematiklehrern. Entwicklung und Erprobung eines Fragebogens. Unveröffentlichte Diplomarbeit, Universität Mannheim.
- Dubs, R. (1995). Konstruktivismus: Einige Überlegungen aus der Sicht der Unterrichtsgestaltung. In: *Zeitschrift für Pädagogik* 41, H. 6, S. 889–903.
- Duffy, T.M./Jonassen, D.H. (1992). Constructivism: New implications for instructional technology. In: Duffy, T.M./Jonassen, D.H. (Eds.): *Constructivism and the technology of instruction: A conversation*. Hillsdale, NJ: Lawrence Erlbaum, S. 1–16.

- Ernest, P. (1988, August). The impact of beliefs on the teaching of mathematics. Paper presented at the 6th International Congress of Mathematical Education, Budapest. Verfügbar unter: <http://www.ex.ac.uk/~PERnest/impact.htm> [31.03.2001].
- Ernest, P. (2000). Social constructivism as a philosophy of mathematics: Radical constructivism rehabilitated? [Online Paper]. Verfügbar unter: <http://www.ex.ac.uk/~PERnest/soccon.html> [28.06.2000].
- Fennema, E./Carpenter, T.P./Loef, M. (1990). Teacher belief scale: Cognitively guided instruction project. Madison, WI: University of Wisconsin.
- Fennema, E./Carpenter, T.P./Peterson, P.L. (1989). Learning mathematics with understanding: Cognitively guided instruction. In: *Advances in Research on Teaching* 1, S. 195–221.
- Fennema, E./Loef Franke, M. (1992). Teachers' knowledge and its impact. In: Grouws, D.A. (Ed.): *Handbook of Research on Mathematics Teaching and Learning*. New York: Macmillan, S. 147–164.
- Fischler, H. (2000). Über den Einfluss von Unterrichtserfahrungen auf die Vorstellungen vom Lehren und Lernen bei Lehrerstudenten der Physik. Teil 1: Stand der Forschung sowie Ziele und Methoden einer Untersuchung. In: *Zeitschrift für Didaktik der Naturwissenschaften* 6, S. 27–36.
- Grigutsch, S./Ratz, U./Törner, G. (1998). Einstellungen gegenüber Mathematik bei Mathematiklehrern. In: *Journal für Mathematikdidaktik* 19, H. 1, S. 3–45.
- Groeben, N./Wahl, D./Schlee, J./Scheele, B. (1988). *Das Forschungsprogramm Subjektive Theorien*. Tübingen: Francke.
- Klieme, E. (1999). Unterricht und mathematisches Verständnis in verschiedenen Unterrichtskulturen. Antrag auf Gewährung einer Sachbeihilfe im Schwerpunktprogramm „Bildungsqualität von Schule“. Berlin: Max-Planck-Institut für Bildungsforschung.
- Klieme, E./Baumert, J. (2001). Identifying national cultures of mathematics education: Analysis of cognitive demands and differential item functioning in TIMMS. In: *European Journal of Psychology of Education* 16, H. 3, S. 383–400.
- Klieme, E./Bos, W. (2000). Mathematikleistung und mathematischer Unterricht in Deutschland und Japan: Triangulation quantitativer und qualitativer Forschungsansätze im Rahmen der TIMS-Studie. In: *Zeitschrift für Erziehungswissenschaft* 3, H. 3, S. 359–379.
- Klieme, E./Neubrand, M./Lüdtke, O. (2001). Mathematische Grundbildung: Konzepte und Befunde. In Baumert, J. u.a. (Hrsg.). *Programme for International Student Assessment*. Nationaler Bericht. Opladen: Leske & Budrich, S. 141–190.
- Klieme, E./Schümer, G./Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In: Klieme, E./Baumert, J. (Hrsg.): *TIMSS – Befunde für Schule und Forschung*. Bonn: Bundesministerium für Bildung und Forschung, S. 43–58.
- Köller, O./Baumert, J./Neubrand, J. (2000). Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Band 2 Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske + Budrich, S. 229–270.
- Nespor, J. (1987). The role of beliefs in the practice of teaching. In: *Journal of Curriculum Studies* 19, H. 4, S. 317–328.
- Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. In: *Review of Educational Research* 62, H. 3, S. 307–332.
- Peterson, P.L./Carpenter, T.P./Fennema, E. (1989). Teachers' knowledge of students' knowledge in mathematics problem solving: Correlational and case analysis. In: *Journal of Educational Psychology* 81, H. 4, S. 558–569.
- Research Report 1998 – 2000. Max-Planck-Institut für Bildungsforschung. Berlin 2001.
- Schoenfeld, A. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. In: *Cognitive Science* 7, S. 329–363.
- Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. In: *Harvard Educational Research* 57, H. 1, S. 1–22.

Staub, F./Stern, E. (in press). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology*.

Weinert, F.E./Schrader, F.W./Helmke, A. (1989). Quality of instruction and achievement outcomes. In: *International Journal of Educational Research* 13, S. 895–914.

Anschrift der Autoren:

Martina Diedrich, Deutsches Institut für Internationale Pädagogische Forschung. Arbeitseinheit Bildungsqualität und Evaluation, Schloßstr. 29, 60486 Frankfurt a.M.

Dr. Claudia Thußbas, Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Franklinstr. 5–7, 10587 Berlin.

Prof. Dr. Eckhard Klieme, Deutsches Institut für Internationale Pädagogische Forschung. Arbeitseinheit Bildungsqualität und Evaluation, Schloßstr. 29, 60486 Frankfurt a.M.

Hans E. Fischer/Thomas Reyer/Christina Wirz/Wilfried Bos/Nicole Höllrich

Unterrichtsgestaltung und Lernerfolg im Physikunterricht¹

1. Theoretische Grundlagen

Bereits TIMSS hat gezeigt, dass die Physikleistungen der Schülerinnen und Schüler in der deutschen Sekundarstufe I nicht den gesellschaftlichen Erwartungen entsprechen. Sie bleiben weit hinter den Lehrplananforderungen der einzelnen Bundesländer zurück: Zum Beispiel verfügen 20% der Schülerinnen und Schüler am Ende des achten Jahrgangs noch nicht einmal über physikalisches Wissen auf Grundschulniveau, nur 25% zeigen ein beginnendes Verständnis von naturwissenschaftlichen Konzepten und Verfahren (Baumert u.a. 1997).

Im Kontext internationaler Vergleiche spielen die Begriffe „Scientific-Literacy“ und „naturwissenschaftliches Arbeiten“ eine besondere Rolle (Fischer 1998). Über rein fachliche Kompetenz hinaus sollen Schülerinnen und Schüler in die Lage versetzt werden, gesellschaftlich relevante Diskussionen über fachtypische Themen zu verstehen und eventuell sogar mitzugestalten (DeBoer/Bybee 1995), um als autonome Persönlichkeit am gesellschaftlichen Leben teilnehmen zu können. Roth (1997) postuliert sogar, dass nur Diskurs ermöglichende Lernumgebungen sich überhaupt positiv auf Physiklernen auswirken.

Scientific-Literacy ist, trotz seiner grundlegenden Schwäche bezüglich der Evaluation, in den USA zu einem zentralen Erziehungsziel in den naturwissenschaftlichen Fächern geworden. Unter anderem die „American Association for the Advancement of Science“ (1989), die „Biological Science Curriculum Study“ (1990) und die „National Science Teacher Association“ (1990) geben Scientific-Literacy sogar explizit als notwendig an, um sich in der zukünftigen Gesellschaft zurechtzufinden und diese weiterzuentwickeln (siehe auch Gräber/Bolte 1997; Millar 1996; Miller 1997).

Die Befunde von TIMSS und die skizzierten Ansprüchen an Physikunterricht legen deshalb die Forschungsfrage nahe: Wie muss Unterricht organisiert sein, damit gemeinsames Handeln, Sprechen über naturwissenschaftliche Inhalte in sozialen Kontexten und das Lösen von Problemen auf der Basis der hierzu notwendigen physikalischen Modelle und Arbeitsweisen gelernt werden kann? Wie muss Unterricht strukturiert sein, damit Schülerlernen im Gesamttablauf optimal ermöglicht und zudem das beabsichtigte Lernergebnis effizient erreicht wird? Lehrerzentrierter Unterricht kann Schülerinnen und Schülern nur sehr begrenzt die Möglichkeit geben, die Tragweite eigener

1 Die Studie wurde gefördert durch Mittel der DFG (FI 477/12-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

Ideen auszuloten – aber auch die bisherige Praxis des schülerorientierten Experimentalunterrichts scheint das skizzierte Problem nicht lösen zu können. Nach Lunetta (1998) ist er ähnlich unwirksam wie lehrerzentrierter Unterricht, da in ihm hauptsächlich Handlungsschritte nach Art eines Kochrezeptes abgearbeitet werden. Es muss daher geprüft werden, welche Aspekte von Unterricht als Operationalisierung von Unterrichtsqualität in Frage kommen und nachweisbar lernförderlich sind.

Einen Ansatz für eine zunächst fachunabhängige Beurteilung von Unterrichtsverläufen und -strategien bieten die so genannten Basismodelle von Oser/Patry (1990). Sie gehen davon aus, dass Lehrerinnen und Lehrer ihre Unterrichtshandlungen nach wenigen von den Autoren formulierten Basismodellen strukturieren können und sollen, um effizient bestimmte Typen von Lehrzielen zu verfolgen. Die TIMS-Videostudie (Baumert u.a. 1997) lässt vermuten, dass in der gegenwärtigen Praxis des Mathematikunterrichts nur wenige Lehrzieltypen stereotyp verfolgt werden. Es sind deshalb immer ähnliche Schemata zur Handlungsregulation oder, nach Oser und Patry, ähnliche Choreografien zu beobachten.

In diesem Projekt wird für die Analyse der Sichtstruktur dennoch vom TIMSS-Konzept (wie auch vom Biqua-Projekt vom IPN Kiel, vgl. Prenzel et al. in diesem Band) ausgegangen, aber entsprechend der Bedingungen für Physikunterricht adaptiert und ergänzt durch eine allgemeine Einschätzung der Unterrichtsqualität wie sie von Clausen (2000) zusammengestellt worden ist. Eine Kategorisierung der Unterrichtsaufnahmen auf einer oberflächlichen Ebene ermöglicht eine Unterrichtsbeschreibung vergleichbar mit der TIMS-Videostudie. Hiermit wird zunächst eine Arbeit nachgeholt, die TIMSS bisher nur für den Mathematikunterricht geleistet hat, nämlich eine Bestandsaufnahme von Physikunterricht (hier in der gymnasialen Sekundarstufe I). Die im Physikunterricht gebräuchlichen Skripte sind bisher noch nicht empirisch untersucht (Klime/Schecker 2001).

Aus dem Ansatz der Basismodelle lässt sich für das Wirken von Unterricht schlussfolgern, dass Sichtstrukturen zwar Merkmale liefern, nach denen sich Unterricht typisieren lässt – sie aber nicht in der Lage sind, die Schülerleistungen hinreichend zu erklären. Hingegen scheinen weniger sichtbare, hoch-inferente Merkmale des Unterrichts und seiner Abläufe für die Lernhandlungen der Schülerinnen und Schüler und für ihre Leistungen entscheidend zu sein. Die Tiefenstruktur wird in diesem Projekt auf zwei Ebenen analysiert, der Lehrerebene und der Schülerebene. Zur Einschätzung der Wirksamkeit der Lehrerhandlungen ist nicht nur eine Leistungsmessung bei den Schülerinnen und Schülern nötig, sondern, da es auch um naturwissenschaftliches Arbeiten und physikalische Konzepte als Lehrziel geht, zusätzlich die Einschätzung der konzeptuellen Entwicklung und der spezifischen Kompetenzen.

Für die Tiefenstruktur wurde das folgende Modell entwickelt: Nach Oser/Patry (1990) werden Lehrerkonzepte als Basiskonzepte theoretisch beschrieben und expliziert. Es wird erwartet, dass es entsprechende Basiskonzepte der Lehrerinnen und Lehrer gibt, die ihre Handlungen latent beeinflussen, nur ist es bisher unklar, wie sich die daraus zu entwickelnden Choreographien auf die Lernhandlungen und die Lernerfolge der Schülerinnen und Schüler auswirken.

Wir gehen davon aus, dass Unterricht, der entlang der Basismodelle geplant und durchgeführt wird, zu besseren Leistungen der Schülerinnen und Schüler führt. Für den Physikunterricht ist danach anzunehmen, dass inhaltlich verschiedene Unterrichtsabschnitte, z.B. die Erarbeitung eines neuen physikalischen Konzepts oder das Einüben eines mathematischen Verfahrens zur Darstellung von Messwerten, eine ähnliche Sichtstruktur ergeben können, obwohl ihnen unterschiedliche Tiefenstrukturen zuzuordnen sind. Von unterschiedlichen Lehrerinnen oder Lehrern erteilter Unterricht zur Erarbeitung eines physikalischen Konzepts kann deshalb bei oberflächlicher Betrachtung (Sichtstruktur) ähnlich organisiert sein, aber dennoch unterschiedliche Intentionen der Unterrichtenden zum Ausdruck bringen.

Werden Basismodelle nach Oser als kognitive Schemata der Handlungsregulation aufgefasst, kann ihre Wirkung auf die Unterrichtsgestaltung durch Operationalisierung und Analyse der Lehrer- und Schülerhandlungen mit Hilfe der Analyse geeigneter video-graphierter Unterrichtsszenen (Jacobs/Kawanaka/Stigler 1999) und die Auswirkung auf die Leistungen und das Interesse der Schülerinnen und Schüler durch entsprechende Tests untersucht werden. Um Aussagen über die Unterrichtsmechanismen zu erhalten, ist deshalb nicht nur zu erfassen, wie sich das Basiswissen der Lehrerinnen und Lehrer im Unterricht widerspiegelt, sondern auch, welche Reaktionen darauf bei den Schülerinnen und Schülern zu erkennen sind. So kann je nach den genutzten Basiskonzepten erklärt werden, dass trotz ähnlicher Sichtstruktur die Leistung und das Interesse der Schülerinnen und Schüler unterschiedlich beeinflusst werden.

Oser und Patry (Oser u.a. 1997) schlagen insgesamt zwölf Basismodelle vor, die für die Gestaltung schulischen Lernens relevant sind: 1) Lernen durch Eigenerfahrung/ entdeckendes Lernen, 2) Entwicklung als Ziel der Erziehung, 3) Problemlösendes Lernen, 4) Begriffs- und Konzeptbildung, 6) Lernen von Strategien, 7) Routinebildung und Training von Fertigkeiten, 8) Motilität, 9) Aufbau dynamischer Sozialbeziehungen, 10) Wert- und Identitätsaufbau, 11) Hypertextlernen, 12) Verhandeln Lernen.

Die den Basismodellen zugeordneten Kategorien sind in dieser Untersuchung unter Berücksichtigung der besonderen Bedingungen des Physikunterrichts entwickelt worden. Die Basismodellabhängigkeit des naturwissenschaftlichen Lernens sollte die unterschiedlichen Schülerleistungen nicht nur im generellen Leistungsvergleich (TIMSS-Scores), sondern auch bei einzelnen TIMSS-Aufgaben beeinflussen. Diese können mittels der Analyse differenzieller Itemfunktionen (DIF, vgl. Holland/Wainer 1993; Baumert/Klieme/Watermann 1999) – trotz einparametrisch raschskalierten Leistungswerte – auf tatsächlich auftretende Leistungsunterschiede zwischen bestimmten Personengruppen untersucht werden, sodass Hinweise auf ihre Basismodellabhängigkeit ermittelt werden können.

Nach Brouër (2001) deutet sich ein positiver Effekt von basismodellorientiertem Unterricht auf die Differenzierung der Unterrichtswahrnehmung durch die Schüler an. Eine entsprechende Interventionsstudie im Fach Deutsch macht deutlich, dass die Schülerinnen und Schüler in basismodellorientiertem Unterricht ihre eigenen Lernprozesse signifikant häufiger beschreiben können als Schülerinnen und Schüler der Kontrollgruppe ohne Intervention. Eine größere Übereinstimmung zwischen vom Lehrer ange-

strebten und von den Schülern beschriebenen Lernschritten in der Versuchsgruppe kann allerdings nicht festgestellt werden. Die Leistungen der Gruppen sind nicht verglichen und die Untersuchung ist nur mit je einem Lehrer für die Kontroll- und die Vergleichsgruppe durchgeführt worden, die vorher kein besonderes Training hierfür absolvieren mussten; es ist deshalb bezüglich der Leistung eher von geringen Effektstärken auszugehen. Dennoch stützt die zitierte Untersuchung die Annahme, dass Basismodell-Unterricht, oder allgemeiner: geeignet strukturierte Lernverläufe, ihre Spuren im Lernergebnis hinterlassen. Es ist aber noch zu untersuchen, ob an Basismodellen orientierte Lehrerbildung bei Lehrern und Schülern im Physikunterricht, sowohl konzeptuell als auch auf Leistung bezogen, Effekte erkennen lässt.

Um diese Verknüpfung von Lehrerkonzepten, Unterrichtsstruktur und Schülerleistung empirisch zu untersuchen, bedarf es eines Ansatzes, der gleichzeitig Lehrervorstellungen, ihre handlungsleitenden Kognitionen, ihre Ausprägung im Unterricht und die damit in Zusammenhang stehenden Schülerhandlungen erfasst.

2. Hypothesen

Das dargestellte Forschungsanliegen legt keine geschlossene Theorie zu Grunde, die es nun zu testen gelte. Stattdessen sollen Wirkweisen von Physikunterricht exploriert werden. Hier besteht insbesondere folgender Klärungsbedarf:

- Bestandsaufnahme grundlegender Parameter und Erscheinungsformen des Physikunterrichts (niedrig- und hoch-inferente Daten), mit dem Ziel einer Beschreibung des Physikunterrichts nach TIMSS-Video.
- Identifizierung der Lehrervorstellungen und Basiskonzepte zur Handlungsregulation, Organisation und Gestaltung des Physikunterrichts.
- Identifizierung von lehrerabhängigen Mustern im Unterrichtsverlauf (Oberflächen- und Tiefenstruktur).
- Korrelation spezifischer Cluster von Basismodellen mit Lernerfolg und Interesse der Schülerinnen und Schüler.

Ziel der theoretischen Arbeit ist es, die oserschen Handlungsketten – empirisch belegt – neu zu formulieren und ein entsprechendes Analyseinstrument als Grundlage einer späteren Intervention und Implementation vorzubereiten. Darüber hinaus sollen Gestaltungsaspekte erfolgreichen Physikunterrichts benannt und Inhalt einer zielgerichteten Lehrerbildung werden. Die folgenden Hypothesen müssen hierzu bestätigt werden:

- Die Basismodelle von Oser lassen sich für Physikunterricht modifizieren.
- Einzelne Basismodelle lassen sich in „normalem“ Physikunterricht als Gestaltungs- und Handlungsgrundlage beobachten.
- Einzelne Unterrichtsmerkmale (Muster in Oberflächen- und Tiefenstruktur) korrelieren positiv mit dem Lernerfolg.
- Das Lernergebnis der Schülerinnen und Schüler lässt sich nur hinreichend unter Berücksichtigung der Unterrichtsmerkmale in der Tiefenstruktur (nicht nur in der Sichtstruktur) erklären.

Mit dieser Studie soll geklärt werden, welche Unterrichtsebenen auf das Schülerlernen wirken. Es werden Oberflächen- und Tiefenstrukturmerkmale berücksichtigt, um die längsschnittlich gemessenen Schülerleistungen begründet auf den Unterricht zurückzuführen. Auf der Seite der Lehrer wird untersucht, welche Basiskonzepte sich in der Gestaltung des normalen Physikunterrichts finden lassen, ob es typische Choreografien gibt und wie am naturwissenschaftlichen Arbeiten orientierte Basismodelle der Unterrichtenden mit der Leistung und dem Interesse der Schülerinnen und Schüler verknüpft sind. Hierzu werden in unterschiedlichen Inhaltsbereichen des Physikunterrichts (nach Lehrplan) pro Lehrer etwa zehn Unterrichtsstunden aus verschiedenen Unterrichtsphasen gefilmt. Diese längsschnittliche Verteilung der Aufnahmen über einen Zeitraum von drei Halbjahren soll erreichen, dass die Unterrichtsqualität zwischen den unterschiedlichen Klassen reliabel beurteilt werden kann.

3. Untersuchungsmethoden und Design

Die Datenbasis bilden im Querschnitt sechs Schulklassen (ca. 170 Schüler) an zwei Dortmunder Gymnasien. Die Studie ist längsschnittlich angelegt, um die Auswirkungen des Physikunterrichts isolieren zu können. Der Beobachtungszeitraum umfasst die nach Lehrplan üblichen Inhaltsbereiche (Elektrizität, geometrische Optik und Kraft) der Klassen Acht und Neun in Folge, um die Konstanz der Choreografie und eventuelle Änderungen über drei Halbjahre zu kontrollieren. Begonnen wurde die Untersuchung in der Klassenstufe Acht. Insgesamt sechs Lehrer (keine Lehrerin) führen ihren Unterricht nach Lehrplan durch. Zu vier Messzeitpunkten wird die Leistung durch TIMSS-Testhefte und zusätzliche inhaltsbezogene Tests gemessen, und die Interessenentwicklung der Schülerinnen und Schüler, ihre Motivation und Weltbilder mit den entsprechenden Instrumenten aus der TIMSS-Pop-II-Untersuchung erhoben.

Für die Aufklärung der Wirkung von Lehrerkonzepten und Unterricht müssen drei Untersuchungsebenen betrachtet werden: Die vom Lehrer intendierten Konzepte zu Lehren und Lernen im Physikunterricht sowie ihr Professionswissen, die Konzepte, die im Physikunterricht verwirklicht werden, bzw. die Auswirkungen der Lehrerhandlungen auf mögliche Lernwege der Schülerinnen und Schüler und die Auswirkungen auf ihre Leistungen. Das Testinstrumentarium erfasst also Unterricht nach strukturellen Kriterien, Qualitätsmerkmalen und nach Leistungsdaten der Schülerinnen und Schüler.

Die laufende zweijährige Projektphase wurde im Juni 2000 begonnen; die empirische Phase begann im September 2000 und ist im Februar 2002 abgeschlossen worden. Ein Teil des hier aufgeführten Inventars wird zur Messung von Entwicklungen vier Mal angewendet, so dass die drei Unterrichtsreihen als mehrfaches Pre-Post-Design eingeklammert werden. Dies betrifft die Lehrerkonzeptinterviews und das Aufgaben- und Interessen-Inventar (vgl. Abbildung 2).

Im Folgenden wird ausgeführt, welche Erhebungsmethoden und Analyseverfahren verwendet bzw. entwickelt worden sind.

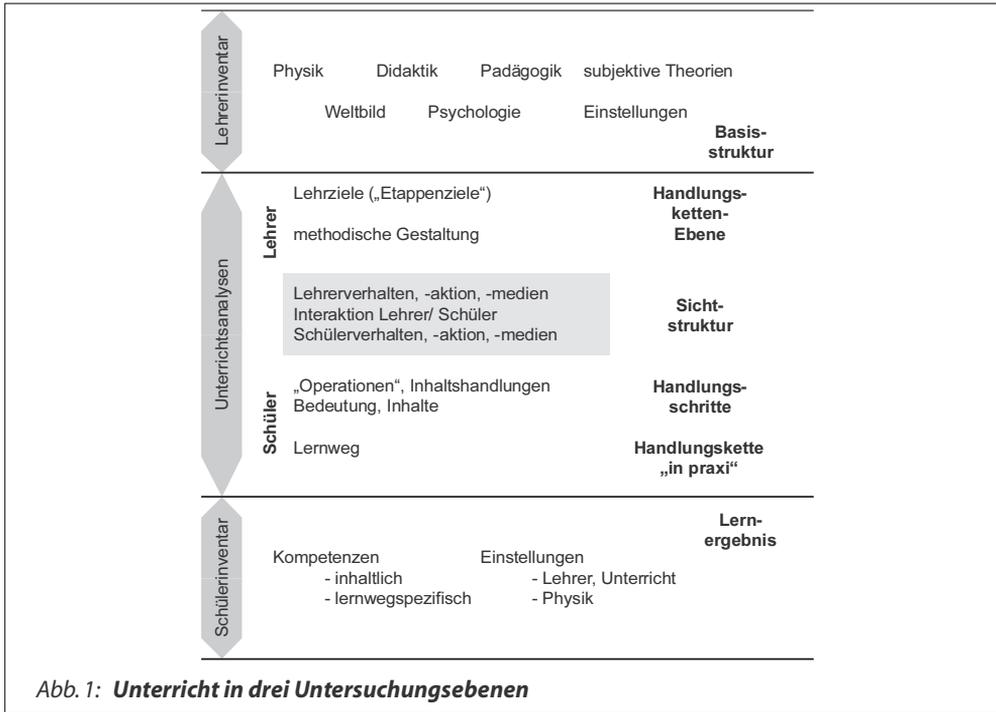


Abb. 1: **Unterricht in drei Untersuchungsebenen**

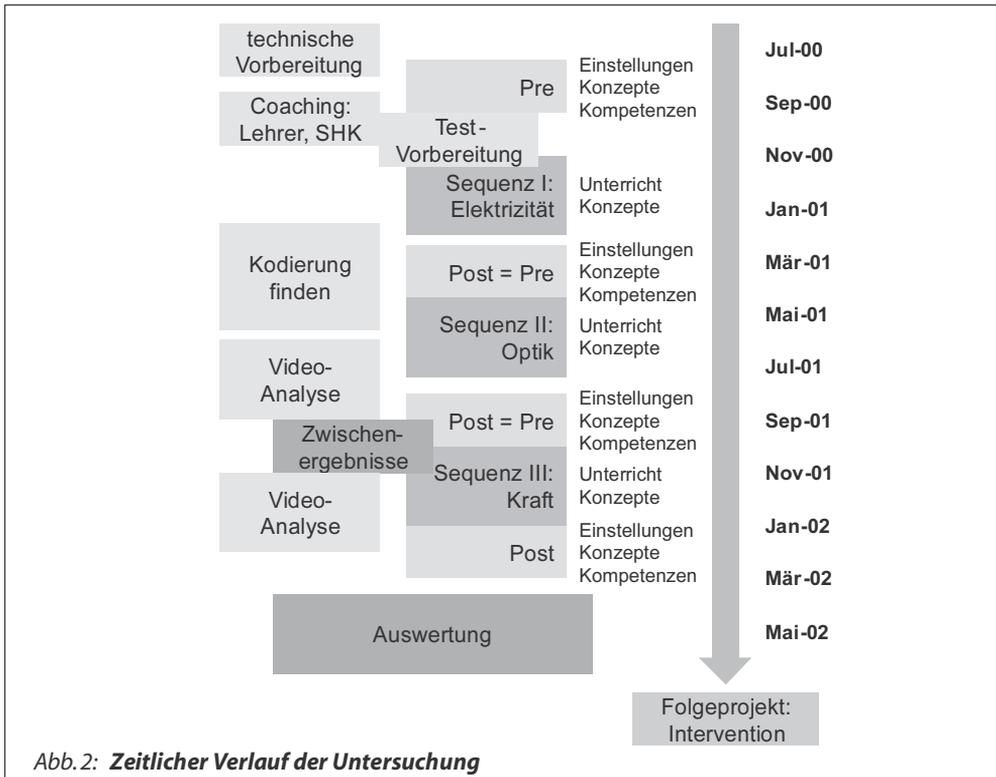


Abb. 2: **Zeitlicher Verlauf der Untersuchung**

TIMSS-Aufgabentests

Die TIMSS-Aufgabentests (für Population II: siebte und achte Jahrgangsstufe) dienen der Erfassung des Lernfortschritts der Schülerinnen und Schüler und darüber hinaus der Beschreibung der Kompetenzprofile der Klassen. Die hier verwendete Aufgabenauswahl entspricht nur einem halben TIMSS-Testheft, um die Testbelastung der Schülerinnen und Schüler zu minimieren. Die Aufgabentests liefern für jeden Probanden Leistungsscores in Rasch-Skalierung (in internationaler TIMSS-Metrik: Mittelwert 500 ± 100) für die Bereiche „Science“ und „Mathematics“.

TIMSS-Hintergrundfragebögen

Um lernrelevante Einstellungen und Interessen der Schüler zu erheben, wird eine Auswahl des deutschen TIMSS-Hintergrundfragebogens verwendet.

Eine modifizierte Form dieses Fragebogens wurde von den Lehrern zeitgleich beantwortet; mit diesen Daten sind Vergleiche zwischen den physikalischen und pädagogischen Konzepten der Lehrer untereinander und mit denen der Schüler möglich.

Spezifische Inhaltstests

Um den Lernerfolg an den tatsächlich behandelten Lehrinhalten zu messen und ein curriculares Maß für den querschnittlichen Vergleich der sechs Schulklassen zu ermöglichen, werden nach jedem Halbjahr, d.h. nach jedem Themenblock, unterrichtsspezifische Inhaltstests durchgeführt, die mit Hilfe der Unterrichtsdokumentation durch die unterrichtenden Lehrer und anhand der Lehrpläne konstruiert wurden.

Der Test zur Elektrizitätslehre (erstes Halbjahr) beinhaltet Aufgaben von Shipstone (Shipstone 1984; Übersetzung: von Rhöneck 1986) und eigene Aufgaben. Der Test zur Optik (zweites Halbjahr) besteht nur aus eigenen Aufgaben. Der Test zum Thema Kraft (drittes Halbjahr) besteht aus eigenen Aufgaben und einer Auswahl aus dem Force-Concept-Inventory (Hestenes/Wells/Sweckhammer 1992).

Kognitiver Fähigkeitstest (KFT)

Um die Schulklassen in ihren Lernvoraussetzungen miteinander vergleichen zu können, sind mit dem kognitiven Fähigkeitstest KFT (Heller/Perleth 2000) die kognitiven Fähigkeiten der Schüler auf drei Skalen abgebildet worden (verbal, nonverbal, quantitativ). Diese Skalen können in der abschließenden Auswertung als Einfluss auf die Leistungsdaten der TIMSS-Scores berücksichtigt werden.

Fragebögen zum Unterricht

Die Lehrinhalte des gesamten Unterrichts im Beobachtungszeitraum werden als Dokumentation der Unterrichtsverlaufspläne der Lehrer für jede Unterrichtsstunde erfasst. Direkt nach jeder aufgezeichneten Unterrichtsstunde erhalten Lehrer und Schülerinnen und Schüler zwei verschiedene Fragebögen, die bei den Schülern Begriffe für das Unterrichtsgeschehen und bei den Lehrern die Unterrichtsziele und die eigene Erfolgseinschätzung abfragen.

Lehrerinterviews

Das unterrichtsrelevante und handlungsleitende Wissen der Lehrer wird in einem Leitfaden-Interview (etwa 45 Minuten) abgefragt. Der Leitfaden basiert auf einem Leitfaden von Labudde (2000) und fragt darüber hinaus weitere Aspekte zur Unterrichtsplanung und der kollegialen Kooperation ab. Die Interviews wurden dreimal durchgeführt und als Transkripte textbasiert inhaltsanalytisch ausgewertet. Die Inhaltsanalyse zu den Lehrerinterviews ist mit Hilfe des Datenmaterials der ersten beiden Interviews textanalytisch entwickelt worden. Ein weiteres Interview befragt die Lehrer zu ihrem Professionswissen und zu ihrem persönlichen Verständnis ihrer Profession.

Unterrichtsaufzeichnungen

Die sechs Lehrer und ihre Klassen werden exemplarisch im Physikunterricht videografiert. Die Videoaufnahmen erfolgen lehrerzentriert nach Regeln, die im Wesentlichen den Anweisungen zur Kameraführung der TIMS-Videostudie entspricht (Stigler/Fernandez 1995); die Kamera folgt hier vor allem dem Lehrer und schließt seinen Interaktionsradius ein (zum Beispiel Schüler als Dialogpartner oder Impulsgeber). Es sind 60 Unterrichtsstunden in den insgesamt sechs Schulklassen aufgezeichnet worden.

Kodierung der Unterrichtsvideos

Das oben skizzierte Unterrichtsmodell legt es nahe, das umfangreiche Kodierverfahren zur Analyse der verschiedenen Oberflächen- und Tiefenaspekte in getrennten Kodierdurchgängen anzulegen. Diese Durchgänge werden im Folgenden dargestellt; wegen des Umfangs kann die Darstellung hier nur exemplarisch erfolgen.

Die ersten beiden Kodierdurchgänge dienen der Analyse von Oberflächenmerkmalen des Unterrichts als Sichtstruktur und Unterrichtsqualität. Diese beiden Datenquellen sollen den Bezug zwischen den von den Lehrern geäußerten methodischen Zielen herstellen und partiell die Lernleistung der Schülerinnen und Schüler erklären.

Die Kodierung der Sichtstruktur entspricht einer Erweiterung des TIMSS-Kodierschemas für den Physikunterricht, der insbesondere eine Berücksichtigung der Rolle des Experiments im Physikunterricht fordert. Hier wurde auch ein entsprechendes Verfahren von Tina Seidel (siehe Beitrag vom IPN Kiel in diesem Band) berücksichtigt, um die im Rahmen des DFG-Schwerpunktprogramms „Biqua“ gewonnenen Ergebnisse vergleichen zu können; dies betrifft die Kodierung der Interaktionsformen (Klassengespräch, Gruppenarbeit, Stillarbeit und Übergang) sowie der untergeordneten Schüleraktivitäten.

Beispiele für die zusätzlichen Verhaltenskategorien sind: Aktion, Medien und Verhalten von Schülern und Lehrer, insbesondere Umgang des Lehrers mit Schüleräußerungen (aufgreifen versus ignorieren) und on/off-Task-Rating der Schüler; ebenso werden die Gruppenarbeit (arbeitsteilig/nicht arbeitsteilig) und die Form der Aufgabenstellung (differenziert/nicht differenziert) bewertet.

Ein weiterer Durchgang dient der vor allem am Lehrerverhalten orientierten Bewertung der Unterrichtsqualität. Er setzt ebenfalls an der Sichtstrukturebene an. Hier wurde ein Fragebogen eingesetzt, wie er in der BRIDGE-Studie vom MPI Berlin (Clausen

2000) für den Mathematikunterricht zusammengestellt und angewandt worden ist. Diese Skalen können auf den Physikunterricht übertragen werden, wenn man berücksichtigt, dass die Bedeutung anspruchsvoller Aufgabenstellungen im Physikunterricht auf experimentelle Aufgaben ausgeweitet werden muss. Die Skalen heißen: Pacing, Monitoring, Disziplin, Sprunghaftigkeit, Interaktionstempo, anspruchsvolles Üben, genetisch-sokratisches Vorgehen, individuelle Bezugsnormierung, Individualisierung, Partizipation, Sozialorientierung, repetitives Üben, Zufriedenheit der Kodierer. Dieser Fragebogen wird nach Abschluss jedes anderen Kodierdurchgangs von jedem Kodierer ausgefüllt. Als „integriertes“ Maß lässt dieses Verfahren allerdings keine Einschätzung der Unterrichtsqualität an den Ansprüchen z.B. von Scientific-Literacy oder naturwissenschaftlicher Arbeitsweisen zu; diese Inhalts- und Verlaufsaspekte werden durch die weiteren Kodierdurchgänge berücksichtigt.

Als Tiefenmerkmale von Unterricht sind die Kodierdurchgänge für Lehrzieltypen und Inhaltshandlungen ausgewählt worden. Die Inhaltshandlungen entsprechen Lernschritten auf der Schülerseite, sie wurden in Bezug auf Osers Basismodelltheorie formuliert. Während die so genannten Handlungsketten bei Oser als Handlungsanleitung für Lehrerinnen und Lehrer zur Organisation entsprechender Lernmöglichkeiten gemeint sind, wird hier versucht, sie als durchgeführte Handlung zu kodieren. Die daraus resultierenden Lernverläufe werden mit den Handlungsketten verglichen, um spezifische Schülerkompetenzen vorherzusagen. Zur Analyse der Unterrichtsaufzeichnungen wurde auf der Basis der Daten und der Anforderungen an Physikunterricht (wie Scientific-Literacy oder naturwissenschaftliches Arbeiten; vgl. Fischer 1998) die folgende Auswahl aus Osers Basismodellen getroffen: Lernen durch Eigenerfahrung, Conceptual Change, Problemlösen, Theoriewissen, Routinebildung und Übersichtslernen.

Kodiert werden die Inhaltshandlungen in 30-Sekunden-Schritten; dabei wird zunächst die Gruppe der Inhaltshandlungen ausgewählt (z.B. „Formulieren/Fixieren des (nun) Bekannten“ oder „Anwenden/Übertragen/Konkretisieren“) und dann über die Inhaltshandlung entschieden (z.B. Festhalten: „Theorie/Merksatz/Formel/Begriff“ bzw. Anwenden: „(erfolgreicher) Lösungsweg auf anderes Problem“). Zusätzlich zu den Handlungsschritten wird kodiert, in welchem Kontext sich der Inhalt bewegt (Teilbereich, übergeordnetes Thema, Alltagswelt) und in welcher Repräsentationsform er benutzt wird (z.B. verbal-mündlich, verbal-schriftlich oder symbolisch).

Als Tiefenstruktur auf der Lehrerseite wird versucht, den „latenten Plan“ des Lehrers für den Unterricht im Verlauf zu interpretieren. Das Verfahren bezieht sich auf die oben genannte Auswahl von Osers Lehrzieltypen, die für den Physikunterricht relevant ist. Zusätzlich wurden neue Lehrzieltypen hinzugenommen, die notwendig sind, um die Unterrichtspraxis im vorliegenden Datenmaterial zu beschreiben, und die bislang nicht in Osers Modell berücksichtigt wurden: Reproduktion, Disziplin und Aktivierung/Lernkontrolle/Leistungskontrolle.

Die Kodierung der Lehrzieltypen erfolgt entlang einer Item-Liste – je Zieltyp sind fünf Items formuliert, wie z.B. für das Theoriewissen: „Der Lehrer möchte, dass die Schüler das allgemeine Gesetz im Spezialfall erkennen.“ – oder für die Reproduktion: „Die Schüler brauchen nach Ansicht des Lehrers das Gegebene nicht verstanden zu ha-

ben, bevor sie es nachmachen oder übernehmen.“ Der Kodierer urteilt in 60-Sekunden-Schritten über die Items, deren Zutreffen zu Zieltypskalenwerten aufaddiert werden.

Inhaltshandlungen und Lehrzieltypen beziehen sich auf die gleiche Auswahl an oser-schen Basismodellen, sodass hier (partiell) eine gegenseitige Validierung möglich ist, wenn beide Kodierungen übereinstimmende Aussagen beinhalten.

4. Erste Ergebnisse

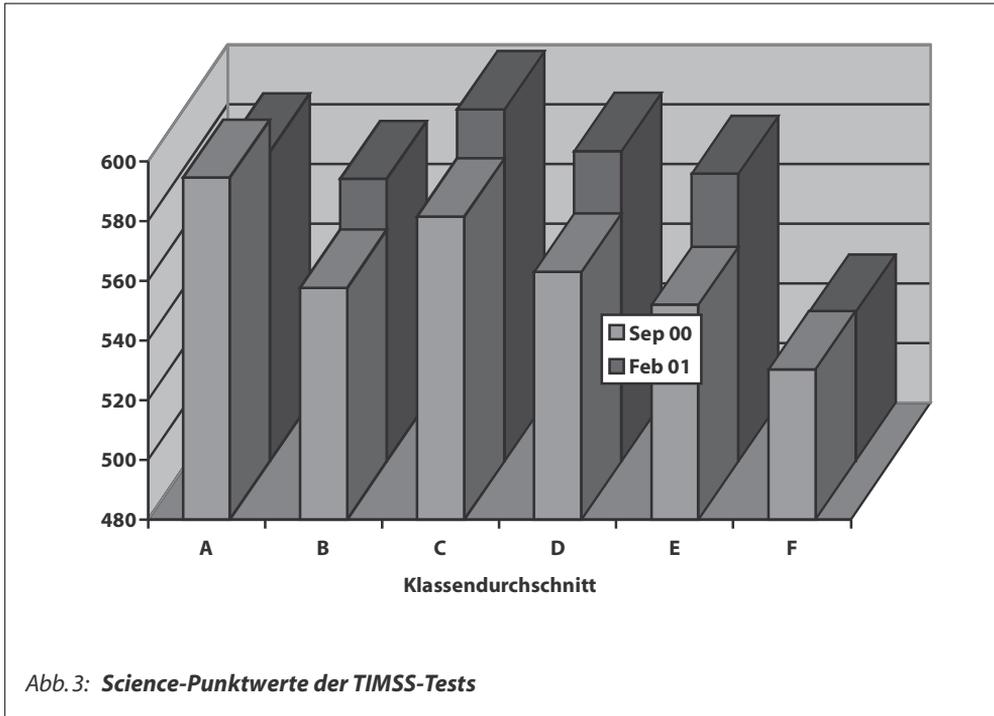
Erste Hinweise aus dem bereits vorliegenden Datenmaterial: Vorbehaltlich aller statistischer Einwände (sechs Lehrer wurden untersucht) und der fragmentarischen Ergebnisse, die zu diesem Zeitpunkt keine sichere Interpretation zulassen (die Datenerhebung dauerte bis Februar 2002), deutet sich an, dass Lehrer deutliche Schwerpunkte in der Wahl ihrer Lehrzieltypen zeigen; im Vergleich mit den Schülerleistungen heben sich erfolgreichere Lehrzieltypen ab.

Ein erster Blick auf die Häufigkeiten der Sichtstrukturanalyse für je eine Unterrichtsstunde je Lehrer zeigt vor allem in folgenden Kategorien Variation zwischen den einzelnen Lehrern: Durchführung von Experimenten (als Aktivität während der Schülerarbeitsphasen), Schülerorientierung des Lehrers (konstant oder variabel) Sozialformen (Verhältnis von Lehrerzentrierung zu Schülerzentrierung), Gruppeninteraktion des Lehrers (Verhältnis von „individuell“ zu „global“), Schülerdialogbeiträge, Schülerexperiment, Aktivität in der Klasse (Variationsbreite). Anhand dieser Merkmale bietet sich eine Gruppierung der Klassen bzw. Unterrichtsformen an in 1. „schülerorientierte Erarbeitung mit experimentellen Aufgaben“ (Klassen B, C, D), 2. „lehrerzentrierte Instruktion mit Demonstrationsexperimenten“ (Klassen A, E, F).

Mit den vorliegenden TIMSS-Daten der ersten beiden Messzeitpunkte ist ein endgültiger Vergleich mit den Schülerleistungen noch nicht möglich; die Lernzuwächse können vollständig erst mit allen vier Messzeitpunkten berechnet werden. Die bisherigen Ergebnisse können plausibel gemacht werden, bestätigen das entwickelte Kodierverfahren – und es lässt sich die beabsichtigte Auswertungsargumentation vorführen.

Das Diagramm (S. 134) der Science-Punktwerte aus den TIMSS-Aufgaben zeigt, dass nur die Klassen B, C, D und E Leistungszuwächse verzeichnen. Die Klassen B, C, D fallen andererseits in die oben genannte Gruppierung von Unterricht als „schülerorientierter Erarbeitung mit experimentellen Aufgaben“. Die Sichtstrukturkodierung kann aber nach bisherigem Stand noch keine Aussagen über die Unterrichtsqualität und die Wirkweisen von Unterricht fällen.

Einige Skalen des Kodiererfragebogens zur Unterrichtsqualität deuten Unterschiede zwischen den Lehrern an (ANOVA mit zwei Unterrichtsstunden pro Lehrer von zwei Kodierern, $N = 18$): Disziplin, Sprunghaftigkeit, genetisch-sokratisches Vorgehen, Individualisierung, Partizipation, globale Zufriedenheit der Kodierer mit dem Lehrer. Die Skalen mit den signifikanten Gruppenunterschieden zeigen, dass die Gruppe B, C, D ein geringes Interaktionstempo und wenig Sprunghaftigkeit zeigt; die Klassen C und D werden von den Kodierern deutlich genetisch-sokratisch argumentierend und individu-



alisierend wahrgenommen. Allerdings sind die Reliabilitäten nach Cronbachs α für zwei dieser Skalen zum Teil nicht ausreichend – mit $\alpha = 0.60$ (gültig: $N = 14$) für Sprunghaftigkeit und $\alpha = 0,39$ für Individualisierung (gültig: $n = 16$); hingegen zeigen die anderen Skalen Werte von $\alpha = 0.83$ bis $\alpha = 0.96$, sodass sich bislang folgendes Bild zeigt: Gerade auf den Skalen, die die stabileren Aussagen treffen, zeigen sich unsystematische Ausprägungen, die keine zusammenfassenden Aussagen über hohe oder geringe Unterrichtsqualität ermöglichen. Dies entspricht der anfänglichen Erwartung, dass diese Analyseebene nicht mit den Schülerleistungen korreliert.

Ein erster Blick auf die Ergebnisse aus der Kodierung der Lehrzieltypen als Mutmaßungen über den „latenten Plan“ des Lehrers verhält sich systematischer. Zunächst wurden die Häufigkeiten der Lehrzieltypen pro Unterrichtsstunde ausgezählt und dann über zwei Unterrichtsstunden pro Lehrer (A bis F) gemittelt. Clusteranalyse und Faktorenanalyse (84% Varianzerklärung) legen eine Lösung mit vier Zieltypengruppierungen nahe, deren Aufbau und Lehrerzuordnungen in Tabelle 1 gezeigt ist („+“ deutet die Stärke der Ausprägung an).

Wenn die Gruppe B, C, D mit der Sichtstrukturgestaltung als „schülerorientierte Erarbeitung mit experimentellen Aufgaben“ die erfolgreichere Gruppe ist, dann scheinen hier auch die Zieltypen Theoriebildung/Strukturtransformation zusammen mit dem Eigenerfahrungsansatz erfolgreich zu sein. Die anfängliche Hypothese unterstellt der Sichtstruktur die geringere Wirkung auf das unterrichtliche Lernen – welche Ebene hier die wirksame ist, kann noch nicht entschieden werden. Entweder ist also die Sichtstruk-

tur eine variationsarme Choreografie der analysierten Lehrzieltypen (was nicht für Lehrer A zutreffen würde) oder aber die Sichtstruktur könnte – entgegen der Eingangshypothese – als zentrale Unterrichtsebene gedeutet werden.

Tab. 1: Zuordnung der Lehrer zu Lehrzieltypen

	Zieltyp Eigenerfahrung (I)	Zieltyp Problemlösen (II)	Zieltypen Theoriebildung und Conceptual Change (III)	Zieltypen Routinebildung, Ü- bersichtslernen, Reproduktion (IV)
Lehrer A	(+)	+	+++	0
Lehrer B	++	0	++	0
Lehrer C	++	0	+++	0
Lehrer D	+++		+++	0
Lehrer E	0	0	0	+++
Lehrer F	0	0	0	+++

Die Kodierung der Inhaltshandlungen soll diese Frage beantworten. Eine mit den Lehrzieltypen stimmige Analyse validiert das Kodiervorgehen für die Tiefenstruktur, und beschreibt die Effizienz der Unterrichtsführung entlang des „Lehrzielfahrplans“. Außerdem sollen die identifizierten regelmäßigen Abfolgen mit Osers Formulierungen von Handlungsketten verglichen werden, die auf spezifische, den Lehrzieltypen entsprechende Schülerkompetenzen zielen (z.B. Theoriereproduktion oder Problemlösen).

Eine solche Verknüpfung von Schülerlernen mit den gemessenen spezifischen Schülerkompetenzen ist zum derzeitigen Stand der Studie noch nicht möglich.

Erste Ergebnisse, abgeleitet aus den Analysen der Professionsinterviews, lassen Hinweise zu einer Clusterbildung finden: Während die Lehrer B und C in einem Großteil der Kategorien ähnliche, wenn nicht sogar gleiche Werte annehmen (bis zu zwei Dritteln der Gesamtnennungen), divergieren die Lehrer A, D, E und F fast regelmäßig von diesen Werten. Die deutlichsten Übereinstimmungen der Gruppe B/C mit starker Abgrenzung zur Gruppe A/D/E/F finden sich in den Kategorien: Auswirkungen der Arbeitszufriedenheit, Personeneinfluss auf Arbeitszufriedenheit, Zufriedenheit als Lehrer, Schülereffekte und Selbstbewertung.

Der Vergleich mit der Gruppierung anhand der Unterrichtsdaten zeigt, dass dies die Aussagen für Lehrer B und C die Modellannahme stützen, dass Lehrer auf der konzeptionellen Ebene verglichen und übergreifende Konzepte isoliert werden können. Die

Lehrer werden wie folgt den zwei verschiedenen Schulen zugeordnet: Schule 1 – A, B, C, D und Schule 2 – E, F. Das Herausheben der Lehrergruppe B und C macht die Schulzugehörigkeit als Alternativerklärung unwahrscheinlicher.

5. Schlussfolgerungen und Ausblick

Mit den skizzierten Ergebnissen kann das Kodierverfahren zur Trennung von Oberflächen- und Tiefenstrukturen im Unterricht als arbeitsfähig angesehen werden. Damit bestätigt sich auch, dass sich das zu Grunde liegende deskriptive Unterrichtsmodell, das von Osers präskriptiver Basismodelltheorie ausgegangen ist, als sinnvoll erweist. Die eingangs aufgestellten Untersuchungshypothesen werden gestützt, nämlich dass sich einzelne Basismodelle als Handlungsgrundlage im Physikunterricht finden lassen, und dass einzelne vergleichsweise stabile Unterrichtsmerkmale mit Lernerfolg korrelieren. Weiterhin zeigen sich deutliche Hinweise, dass die Kenntnis der Sichtstrukturebene allein nicht ausreicht, um das Wirken von Physikunterricht hinreichend zu erklären.

Im Anschluss an diese Studie sollen die wirksamen und qualitativ hochwertigen Merkmale isoliert und zur Grundlage eines Lehrertrainings werden. Es ist zu erkennen, dass eine bestimmte Auswahl von Lehrzieltypen, ergänzt um physikdidaktische Forderungen (z.B. naturwissenschaftliche Arbeitsweisen, Problemlösen, Aufgabenkultur usw.) besonders lernwirksam ist. In einer Interventionsstudie sollen diese isolierten „Erfolgsmerkmale“ mit Lehrerinnen und Lehrern als Gestaltungs- und Handlungskompetenz trainiert und der Erfolg dieses Trainings evaluiert werden. Der Zusammenhang zwischen diesen Elementen des Professionswissens von Lehrerinnen und Lehrern und dem Lernerfolg der Schülerinnen und Schüler wäre damit partiell erklärbar. Da sich bereits abzeichnet, dass die hier entwickelten Untersuchungsinstrumente reliabel und valide sind, kann daraus ein mehrteiliges Testinstrumentarium zur Erfassung des Lehrerwissens über Lernprozesse entwickelt werden. Auf der Theorieebene ist abzufragen, welche Theorien Lehrerinnen und Lehrer kennen und welche Relevanz sie ihnen für die Unterrichtsgestaltung zuordnen. Auf der Handlungsebene kann geklärt werden, inwieweit diese Kenntnisse zur Bewältigung unterrichtlicher Anforderungen tatsächlich genutzt werden.

Literatur

- American Association for the Advancement of Science (1989): *Science for all Americans*. Washington D.C.
- Baumert, J./Klieme, E./Watermann, R. (1999): *Jenseits von Gesamttest- und Untertestwerten: Analyse differentieller Itemfunktionen am Beispiel der mathematischen Grundbildungstests der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie der IEA (TIMSS)*: In Herber, H.-J./Hofmann, F. (Hrsg.), *Schulpädagogik und Lehrerbildung*, S. 301–324. Innsbruck: Studienverlag.

- Baumert, J./Lehmann, R./Lehrke, M./Schmitz, B./Clausen, M./Hosenfeld, I./Köller, O./Neubrand, J. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen: Leske + Budrich.
- Biological Science Curriculum Study (1990): BSCS and educational reform. BSCS: The Natural Selection, S. 1.
- Brouër, B. (2001): Förderung der Wahrnehmung von Lernprozessen durch die Anwendung der Basismodelle des Lernens bei der Gestaltung von Unterricht. In: Unterrichtswissenschaft, Weinheim, S. 153–170.
- Clausen, M. (2000): Wahrnehmung von Unterricht – Übereinstimmung, Konstruktvalidität, und Kriteriumsvalidität in der Forschung zur Unterrichtsqualität. Dissertation am Fachbereich Erziehungswissenschaft und Psychologie, Freie Universität, Berlin.
- DeBoer, G.E./Bybee, R.W. (1995): The goals of science curriculum. In: Bybee, R.W./McInerney, J.D. (Hrsg.): Redesigning the science curriculum: A report on the implications of standards and benchmarks for science education. Colorado Springs, CO: Biological Sciences Curriculum Study.
- Fischer, H.E. (1998): Scientific Literacy und Physiklernen. IN: Zeitschrift für Didaktik der Naturwissenschaften 4, Heft 2, S. 41–52.
- Gräber, W./Bolte, C. (Hrsg.) (1997): Scientific literacy, an international Symposium. Kiel: IPN.
- Gunter, M.A./Estes, T.H./Schwab, J.H. (1990): Instruction. A models approach. Boston: Allyn and Bacon.
- Heller, K.A./Perleth, Ch. (2000): Kognitiver Fähigkeitstest für 4. bis 12. Klassen Revision (KFT 4–12+ R). Göttingen: Hogrefe-Verlag.
- Hestenes, D./Wells, M./Sweckhammer, G. (1992): Force Concept Inventory. In : The Physics Teacher 30 (3), S. 141–151.
- Holland, P.W./Wainer, H. (Hrsg.): (1993): Differential Item Functioning. Hilldale, N.J: Lawrence Erlbaum.
- Jacobs, J.K./Kawanaka, T./Stigler, J.W. (1999): Integrating Qualitative and Quantitative Approaches to the Analysis of Video Data on Classroom Teaching. In Bos, W./Tarnai, Ch. (Hrsg.): Computer Aided Content Analysis. Special Issue of International Journal of Educational Research.
- Klieme, E./Schecker, H. (2001): Mehr Denken, weniger Rechnen. Physikalische Blätter 57 (2001) Nr. 7/8, S. 113–117.
- Labudde, P. (2000): Konstruktivismus im Physikunterricht der Sekundarstufe II (Schulpädagogik – Fachdidaktik – Lehrerbildung). Wien: Haupt.
- Lunetta, V.N. (1998): The school science laboratory: historical perspectives and contexts for contemporary teaching. In Tobin, K./Fraser, B. (Hrsg.): International Handbook of Science Education. Amsterdam: Kluwer.
- Millar, R. (1996): Towards a science curriculum for public understanding. In: School science review 77 (280), S. 7–18.
- Miller, J.D. (1997): Civic scientific literacy in the United States: A developmental analysis from middle-school through adulthood. In: Gräber, W./Bolte, C. (Hrsg.): Scientific Literacy, S. 104–121. Kiel: IPN.
- National Science Teacher Association⁴ (1990): Science education initiatives for the 1990s. Washington, D.C.
- Oser, F./Patry, J.-L. (1990): Choreographien unterrichtlichen Lernens, Basismodelle des Unterrichts. In: Berichte zur Erziehungswissenschaft Nr. 89, Pädagogisches Institut der Universität Freiburg, Schweiz.
- Oser, F./Patry J.-L./Elsässer, T./Sarasin, S./Wagner, B. (1997): Choreographien unterrichtlichen Lernens – Schlußbericht an den Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung. Pädagogisches Institut der Universität Freiburg, Schweiz.
- Rhöneck, Ch. von (1986): Vorstellungen vom elektrischen Stromkreis und zu den Begriffen Strom, Spannung und Widerstand. In: Zeitschrift für Didaktik der Naturwissenschaften 13. Kiel: Schmidt und Klaunig.

- Roth, W.-M. (1997): The interaction of learning environment and student discourse about knowing, learning, and the nature of science: Two longitudinal case studies. In: *International Journal of Educational Research*, 27, S. 311–320.
- Shipstone, D.M. (1984): A study of children's understanding of electricity in simple DC circuits. In: *European Journal of Science Education* Vol. 6, No. 2, S. 185–198.
- Stigler, J.W./Fernandez, C. (1995): *TIMSS Videotape Classroom Study/Field Test Report*. Den Haag: IEA.

Anschrift der Autoren:

- Prof. Dr. Hans E. Fischer, Universität Dortmund, Fachbereich Physik, Arbeitsgruppe Didaktik der Physik, Otto-Hahn-Straße 4, 44221 Dortmund.
- Dipl. Phys. Thomas Reyer, Universität Dortmund, Fachbereich Physik, Arbeitsgruppe Didaktik der Physik, Otto-Hahn-Straße 4, 44221 Dortmund.
- Christina Wirz, Universität Dortmund, Fachbereich Physik, Arbeitsgruppe Didaktik der Physik, Otto-Hahn-Straße 4, 44221 Dortmund.
- Prof. Dr. Wilfried Bos, Universität Hamburg, Fachbereich Erziehungswissenschaft, Institut für International und Interkulturell Vergleichende Erziehungswissenschaft, Sedanstraße 19, 20146 Hamburg.
- Nicole Höllrich, Universität Dortmund, Fachbereich Physik, Arbeitsgruppe Didaktik der Physik, Otto-Hahn-Straße 4, 44221 Dortmund.

Manfred Prenzel/Tina Seidel/Manfred Lehrke/Rolf Rimmele/Reinders Duit/
Manfred Euler/Helmut Geiser/Lore Hoffmann/Christoph Müller/Ari Widodo

Lehr-Lernprozesse im Physikunterricht – eine Videostudie¹

1. Theorie und Fragestellung

Hintergrund für dieses Projekt sind, wie für das gesamte DFG-Schwerpunktprogramm, die Befunde aus internationalen Vergleichsstudien (TIMSS, PISA). Im Vergleich zu anderen Ländern erreichen deutsche Schülerinnen und Schüler im Mittel ein niedrigeres Kompetenzniveau. Sie scheitern insbesondere an Aufgaben, die ein tieferes Verständnis von physikalischen Begriffen und Prinzipien oder naturwissenschaftlichen Denk- und Arbeitsweisen voraussetzen (Baumert u.a. 1997; Prenzel u.a. 2001). Viele Schülerinnen und Schüler scheitern bei der Anwendung ihres Wissens, sie verlieren das Interesse an der Physik und am Physikunterricht und wenden sich im Verlauf der Schulzeit gänzlich von der Physik ab. Vor allem Mädchen interessieren sich kaum für dieses Fach (Hoffmann 2002; Hoffmann/Häußler/Lehrke 1998). Diese Befunde über eine problematische Qualität der Bildungsergebnisse signalisieren Handlungsbedarf. Aussichtsreiche Maßnahmen setzen allerdings empirisch fundiertes Erklärungswissen voraus. Auf entsprechendes Wissen zielt die hier dargestellte Studie. Sie untersucht den deutschen Physikunterricht und die Lernprozesse, die in diesem Unterricht ermöglicht werden.

Der Physikunterricht in Deutschland unterstützt zu wenig das gründliche Verstehen (Lehtinen 1994). Für die Qualität des Lernens sind dabei nicht nur Merkmale des lernenden Individuums entscheidend, sondern Lernen ist „auch externalen Faktoren (Ermöglichkeiten und Einschränkungen) unterworfen“ (Reusser 1998, S. 155). Dies gilt vor allem unter der theoretischen Vorstellung von Lernen als aktivem und konstruktivem Prozess. Eine Lehr-Lernumgebung ist „mehr als eine passive Material- und Angebotslieferantin“, sie hat vielmehr „die Rolle einer aktiven Entwicklungshelferin“ (Reusser 1998, S. 155) zu übernehmen. Die Frage ist, ob und inwieweit der deutsche Physikunterricht dieser Rolle gerecht wird.

Die im Rahmen von TIMSS durchgeführte Videostudie zielte auf einen systematischen transkulturellen Vergleich des Mathematikunterrichts in Japan, den USA und Deutschland (Stigler u.a. 1999). Der Ansatz erwies sich als geeignet, mögliche Bedingungsfaktoren für die Leistungsunterschiede zwischen den Schülern verschiedener Länder in der Gestaltung des Fachunterrichts aufzudecken. Im internationalen Vergleich traten die Gemeinsamkeiten beim Unterrichten *innerhalb* der Länder und die Unterschiede *zwischen* den Ländern besonders hervor (Stigler u.a. 1999; Baumert u.a. 1997).

1 Die Studie wurde gefördert durch Mittel der DFG (PR 473/2-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

Der für Deutschland charakteristische kleinschrittige und enggeführte fragend-entwickelnde Unterricht scheint zu bestimmten Problemen zu führen: Es bleibt wenig Spielraum für eigenständiges Planen und Denken, Lern- und Leistungssituationen werden vermischt, eine differenzielle Förderung lässt sich kaum verwirklichen.

1.1 *Unterrichtsmuster und „Skripts“*

In der Darstellung und Diskussion von TIMSS-Video werden die in einer Kultur vorherrschenden Unterrichtsmuster als „Skripts“ bezeichnet. Es wird angenommen, dass die Unterrichtsskripts oder „Drehbücher“ auf kulturell geteilten Vorstellungen von gelungenem Unterricht beruhen (Stigler/Gallimore/Hiebert 2000). Ein entscheidender Fortschritt dieser Betrachtung liegt darin, dass nun nicht isolierte Unterrichtsmerkmale, sondern Unterrichtsmuster analysiert werden. Im Sinne von kulturell geteilten Skripts repräsentieren die Unterrichtsmuster gewissermaßen kulturelle Kontexte. Damit wird zugleich eine neue theoretische Perspektive der Lehr-Lernforschung aufgegriffen, die den Blick auf Ablaufmuster oder „Choreographien“ (Oser/Patry 1990) richtet. „Choreographien“ des Unterrichts legen demnach das Interaktions- und Handlungsrepertoire für Lehrkräfte sowie für Schülerinnen und Schüler fest und beeinflussen interne Verarbeitungs- oder Steuerungsprozesse beim Lehren und Lernen. Sie beruhen auf einem Skript.

Wie die Lehr-Lernforschung der letzten Jahrzehnte zeigt, gibt es nicht den „Königsweg einer einzigen Unterrichtskonzeption, -strategie oder -methode“ (Baumert/Köller 2000, S. 271). Die Unterrichtsforschung sollte deshalb auch die Vorstellung einer erfolgreichen Gestaltung von Unterricht durch die Manipulation einzelner, isolierter Maßnahmen aufgeben. Die Qualität von Unterricht beruht vielmehr auf einer „Orchestrierung“ verschiedener didaktischer Zugänge und Grundformen (Aebli 1998; Oser/Patry 1990). Im Rahmen des hier vorgestellten Projekts wird deshalb bei der Untersuchung des Wechselspiels zwischen Lehrbedingungen und Lernprozessen versucht, Unterrichtsmuster oder Skripts zu identifizieren.

Hinsichtlich der Funktion von Unterrichtsskripts für individuelle Lernprozesse aufseiten der Schülerinnen und Schüler gibt es zwar eine Reihe von Vermutungen; im Detail belegt sind diese Annahmen allerdings nicht (Klieme 1999; Prenzel u.a. 1999). Oser und Patry (1990) unterscheiden bei der Betrachtung von Unterricht zwischen *Sichtstrukturen* und *Basismodellen* des Lernens. Demnach sind beobachtbare Unterrichtsaktivitäten von intern ablaufenden (lernpsychologisch erforderlichen) Lernprozessen zu unterscheiden. Offen bleibt bei Skriptansätzen weiterhin, welche Rolle subjektive Vorstellungen vom Lehren und Lernen bei der Ausgestaltung von Unterricht (bei Lehrkräften und Schülern) und im kognitiven und motivationalen Lerngeschehen spielen (vgl. dazu Staub/Stern 2002; Fischler 2002).

Diese Zusammenhänge werden im vorliegenden Projekt durch Rückgriff auf lernpsychologische und didaktische Theorien geklärt. Bei der Analyse des Lehr-Lerngeschehens wird die Vorstellung aufgegriffen, dass zur Aktivierung bestimmter kognitiver

bzw. motivationaler Prozesse Gelegenheiten oder Unterstützungen bereitgestellt werden müssen (Seidel 2002). Wir sprechen deshalb von „*Gelegenheitsstrukturen*“ der Lernumgebung und beziehen diese auf die unmittelbar beobachtbaren Unterrichtsmuster („*Sichtstrukturen*“). Auf diese Weise kann geklärt werden, inwieweit bestimmte Unterrichtsmuster bestimmte Gelegenheitsstrukturen bereit stellen. Für die detaillierte Analyse von Gelegenheitsstrukturen wird auf aktuelle theoretische Konzeptionen des Lehrens und Lernens sowie auf physikdidaktische Theorien zum Konzeptwechsel und zum Experimentieren zurück gegriffen.

1.2 Warum Videostudien?

Videostudien sind mit einem sehr großen Aufwand verbunden; besonders dann, wenn komplette Unterrichtsstunden in großer Zahl aufgenommen und analysiert werden sollen. Dennoch gibt es bei bestimmten Fragestellungen sehr gute Gründe, sich für ein solches Vorgehen zu entscheiden. Videostudien können dazu beitragen, bekannte Schwächen herkömmlicher Unterrichtsforschung zu überwinden, zum Beispiel die Tendenz vieler Pädagogischer Psychologen, mit einem sehr engen Blickwinkel einzelne Unterrichtsaspekte zu untersuchen und sie auf den ganzen Unterricht zu übertragen (Shuell 1996, S. 726). Ähnlich argumentieren Helmke und Weinert (1997, S. 125 f.). So gibt es Beispiele, dass einerseits unterschiedliche Unterrichtsverfahren gleich erfolgreich sein können und dass andererseits gleich erscheinender Unterricht je nach Bedingungskonstellation unterschiedliche Effekte erzielen kann. Ein von der Lehrkraft stark kontrollierter Unterricht kann je nach Art und Weise der Kontrolle entweder positive oder negative Auswirkungen auf das Verhalten und die Leistungen der Schülerinnen und Schüler haben (Helmke/Weinert 1997). Es ist also notwendig, theoretisch begründet hinreichend komplexe Segmente des Unterrichtsgeschehens zu analysieren.

Videoaufnahmen erfassen auf den ersten Blick den „sichtbaren“ Teil, also das Verhalten der beteiligten Personen. Sie enthalten aber darüber hinaus wichtige Informationen, die nur durch differenzierte, theoriegeleitete Analysen identifiziert und deutlich gemacht werden können. Die so aufgeschlüsselten Bereiche sind aus lehr-lerntheoretischer Perspektive besonders bedeutsam. Beispiele sind die realisierten sachlogischen Strukturen des dargebotenen bzw. im Unterricht entfalteten Stoffes, qualitative Aspekte des Umgangs und Aufgreifens von Schülerbeiträgen, die kognitive Aktivierung der Schülerinnen und Schüler oder die didaktische Einbettung und Zielsetzung von Experimenten.

Auch feinkörnige Analysen des aufgezeichneten Unterrichtsgeschehens erlauben jedoch nur in einem sehr eingeschränkten Maße die Rekonstruktion internal ablaufender Prozesse. Dazu zählen beispielsweise das erlebte kognitive Engagement, die Qualität der Lernmotivation oder die individuelle Wahrnehmung unterstützender Lehr-Lernbedingungen. Des Weiteren bleiben bei der Beschränkung auf das Videomaterial Aspekte unberücksichtigt, die im Bereich der Vorstellungen von Lehrenden und Lernenden anzusiedeln sind. Welche Vorstellungen leiten beispielsweise eine Lehrperson bei der Pla-

nung und Umsetzung des Unterrichts? Nur durch die Kombination und Integration verschiedener Perspektiven (durch die Nutzung von Beobachtungs-, Fragebogen- und Testdaten) wird es möglich, das komplexe Zusammenspiel unterrichtlicher Lehr-Lernprozesse zu durchdringen.

Wie verschiedentlich empirisch belegt werden konnte, sind Schülerinnen und Schüler in der Lage, bestimmte Merkmale von Unterricht, z.B. solche, die mit dem Lernerfolg oder dem Interesse kovariieren, valide einzuschätzen (Clausen 2001; Prenzel/Drechsel/Kramer 1998). Unter Rückgriff auf die Schülersicht können Zusammenhänge zwischen Unterrichtsmerkmalen und dem Lernerfolg festgestellt werden. Ohne Beobachtungsdaten bleibt jedoch offen, welche Situationsmerkmale und Verhaltensweisen oder Entscheidungen der Lehrperson sowohl dem Urteil der Lernenden als auch dem Ausmaß des Lernerfolges zu Grunde liegen. Hinweise zur Optimierung des Unterrichts lassen sich damit, wenn überhaupt, nur auf einer sehr allgemeinen Ebene ableiten.

Zusammenfassend lässt sich festhalten: Als Reaktion auf erkannte Defizite in der Unterrichtsforschung scheint sich eine Konvergenz einzustellen zwischen den eingeforderten komplexen Erklärungsansätzen, dem Zugang über videogestützte Unterrichtsbeobachtung mit Begleiterhebungen sowie dem Skriptbegriff.

1.3 Arbeitsschwerpunkte und Fragestellungen

Die vorliegenden Erkenntnisse über kulturell geteilte Muster der Unterrichtsorganisation (TIMSS-Video) beschränken sich auf den Mathematikunterricht und können nicht ohne weiteres auf die Physik übertragen werden. Bis heute fehlt eine systematische und zuverlässige Datenbasis über vorherrschende Merkmale und Besonderheiten des Physikunterrichts in Deutschland. Es gibt allerdings Hinweise, dass auch der Physikunterricht in Deutschland wie der Mathematikunterricht durch ein fragend-entwickelndes Vorgehen gekennzeichnet ist. Da im Physikunterricht naturwissenschaftliche Denk- und Arbeitsweisen im Vordergrund stehen, müsste das fragend-entwickelnde Klassengespräch jedoch durch Experimente durchbrochen oder angereichert werden. Dies trifft zumindest auf den Anspruch zu wie er in verbreiteten Didaktiken (Bleichroth u.a. 1999) formuliert ist und wie er auch in den Lehrplänen zum Ausdruck kommt (Pippig/Schneider 1995). Im Vergleich zum theoretischen Anspruch zeigen empirische Befunde zur Unterrichtsrealität jedoch ein ernüchterndes Bild: Aus Sicht der Schülerinnen und Schüler wird der Physikunterricht in Deutschland vorwiegend als kridelastiger „Demonstrationsunterricht“ erlebt (Baumert/Köller 2000). In der gymnasialen Oberstufe herrscht demnach das folgende Unterrichtsmuster vor: Eine lehrergeleitete Konzeptentwicklung, die als Vortrag oder als fragend-entwickelndes Unterrichtsgespräch angelegt ist, wird mit Versuchen bzw. Demonstrationsexperimenten verknüpft. Die Schüler machen Notizen, Schülerexperimente sind selten, unvorstellbar scheint es zu sein, Schüler selbst Experimente planen zu lassen. Man kann aufgrund bisheriger Befunde davon ausgehen, dass dem Experiment im deutschen Physikunterricht nicht die Funktion im Erkenntnis- und Lernprozess zukommt, die man ihm unterstellt. Vielmehr

hängt die Funktion von Experimenten entscheidend davon ab, wie diese in den Unterrichtszusammenhang eingebettet und integriert werden (Euler 2001; Harlen 1999; Lunetta 1998).

In der einschlägigen Forschungsliteratur lassen sich zahlreiche Hinweise auf einzelne Bedingungsfaktoren für intendierte Lernprozesse finden. Doch fehlen komplexere Modelle, aus denen man abgestimmte Maßnahmenbündel zur Qualitätsverbesserung ableiten könnte. Ziel ist es deshalb, eine erste Bestandsaufnahme der „didaktischen Orchestrierung“ von Physikunterricht vorzunehmen und insbesondere die Möglichkeiten für verständnisvolles Lernen zu untersuchen. Die Arbeit konzentriert sich auf drei Schwerpunkte:

- 1) Die Entwicklung, Erprobung und Ausdifferenzierung von Erhebungs- und Auswertungsinstrumenten für Videoanalysen,
- 2) die Beschreibung und didaktische Klassifikation von Mustern des Physikunterrichts bei zwei Themenbereichen des Anfangsunterrichts und
- 3) die Exploration der empirischen Tragfähigkeit von Annahmen über Relationen zwischen Unterrichtsmustern, Lehraktivitäten und Lernprozessen.

Mit theoriegeleiteten Analysen von Unterrichtsbeobachtungen und von ergänzenden Daten aus Befragungen und Tests soll geklärt werden, wie sich Bedingungen bzw. Muster des Physikunterrichts auf Lernprozesse und Bildungsergebnisse auswirken. Das Projekt fokussiert vier theoretisch bedeutsame Aspekte des Lehrens und Lernens im Physikunterricht, die mögliche Problembereiche des deutschen Physikunterrichts darstellen (vgl. Prenzel 2000):

- a) die Zielorientierung (Explikation und Transparenz von Lehrzielen als Basis für die Konstruktion von Lernzielen);
- b) die Begleitung des Lernens und die Unterstützung zielgerichteter mentaler Aktivitäten (Diagnose, Monitoring, Rückmeldung durch die Lehrkraft bzw. durch den Schüler/die Schülerin im Sinne selbstgesteuerten Lernens);
- c) die Berücksichtigung von Schülervorstellungen und das Umgehen mit Fehlern oder Fehlvorstellungen;
- d) den Einsatz und die Einbettung von Lehrer- oder Schülerexperimenten und die Thematisierung naturwissenschaftlicher Prozesse (Denk- und Arbeitsweisen bzw. Vorstellungen von den Besonderheiten der Naturwissenschaften).

Der vorliegende Beitrag konzentriert sich auf die beiden erstgenannten Bereiche, also die Zielorientierung und die Lernbegleitung. Für diese Bereiche werden Beziehungen zwischen beobachtbaren Unterrichtsmustern und ihren Wirkungen auf Lernprozesse und Bildungsergebnisse theoretisch postuliert und explorativ auf ihre empirische Tragfähigkeit überprüft. Zur Untersuchung von Unterrichtsmustern und ihren Wirkungen ist in der ersten Phase des Projekts der Anfangsunterricht in Physik gewählt worden. Routinen und Unterrichtsmuster als Wechselspiel zwischen Lehrenden und Lernenden

werden zu Beginn einer neuen fachlichen Ausbildung innerhalb einer Lerngemeinschaft entwickelt und etabliert. Auf der Ebene dieser ersten „Konfrontation“ der Schülerinnen und Schüler mit Physik und Physikunterricht können so mögliche Unterschiede in den Erfahrungen zwischen den Klassen auf den jeweiligen Anfangsunterricht zurückgeführt werden. Dieser Beitrag fasst Ergebnisse zu den folgenden Fragen zusammen:

- Liefern die gewählten Methoden bzw. die Kombination der Methoden (Beobachtung, Befragung, Tests) aussagekräftige Informationen über Lehr-Lernprozesse im Anfangsunterricht in Physik?
- Lassen sich die untersuchten Klassen hinsichtlich ihres „Erfolges“ (Lernzuwachs, Entwicklungen) differenzieren? Unterscheiden sich die Klassen in der Art des Unterrichts?
- Lassen sich Unterschiede bei Unterrichtseffekten (Lernerfolg, Interesse) mit theoretisch bedeutsamen Indikatoren auf der Beobachtungsebene in Zusammenhang bringen, nämlich mit Mustern der Unterrichtsorganisation und mit Gelegenheitsstrukturen für individuelle Lernprozesse?

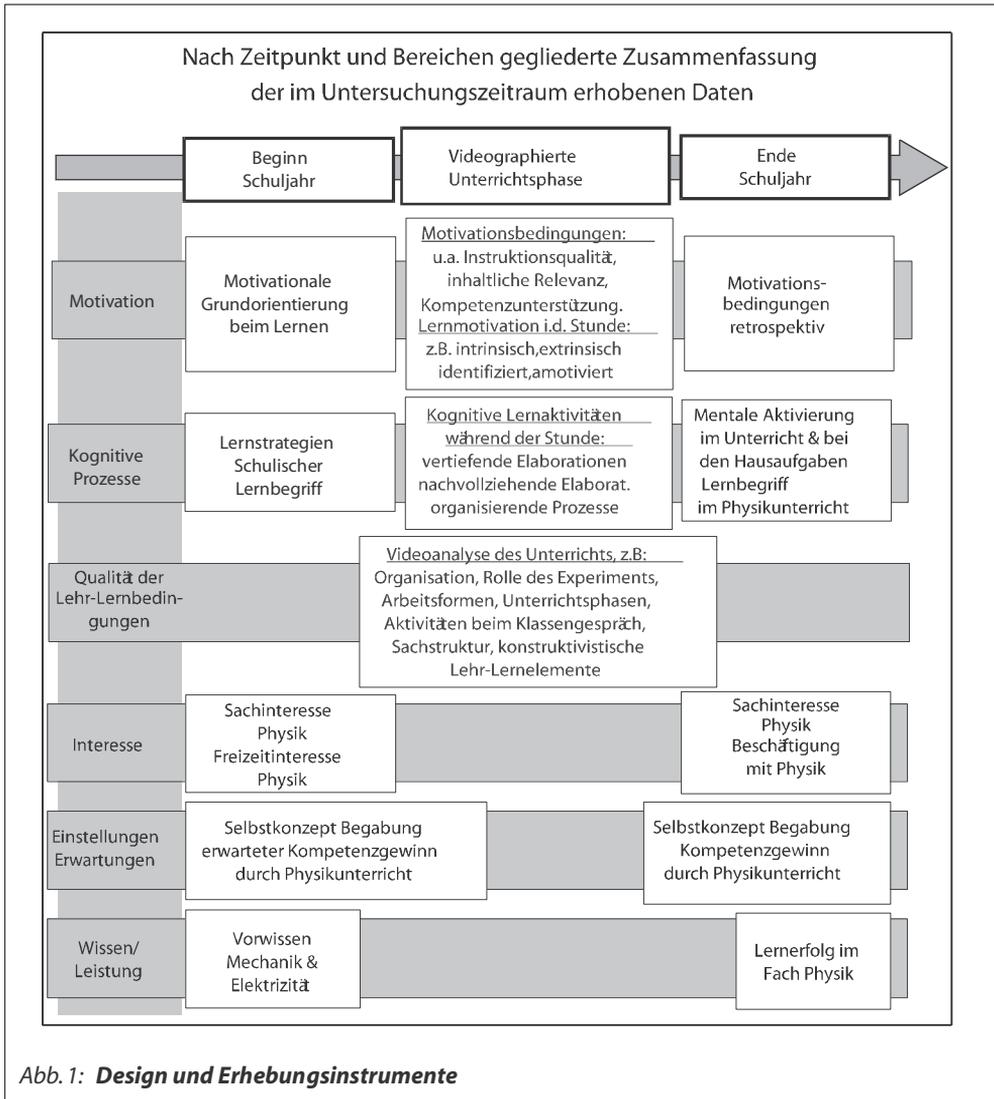
2. Methoden

Ein Ziel der ersten Projektphase lag in der Entwicklung und Anpassung methodischer Instrumentarien. Es wurde dabei eine Herangehensweise gewählt, in der Lehr-Lernprozesse konsekutiv über verschiedene methodische Zugänge analysiert werden. Neben der Darstellung des Designs und der Stichprobenbeschreibung wird im Folgenden ein Überblick über die im Projekt angewandten Erhebungsverfahren gegeben.

2.1 Design und Stichprobe

An der Untersuchung nahmen insgesamt 13 Schulklassen (9 Gymnasial- und 4 Realschulklassen) aus zwei Bundesländern teil. In einem längsschnittlich angelegten Design wurde über das Schuljahr 2000/01 hinweg der Anfangsunterricht in Physik (Jahrgangsstufe 7/8) untersucht (Abb. 1).

Im Abstand von etwa einem halben Jahr wurden zwei Unterrichtseinheiten zu je drei Stunden (Thema 1: Einführung in den Stromkreis, Thema 2: Einführung in den Kraftbegriff) aufgezeichnet. Bei 13 Klassen ergibt sich damit ein Stichprobenumfang von $N = 78$ Unterrichtsstunden (2 Themen \times 3 Stunden \times 13 Klassen). Die Aufzeichnung der Stunden erfolgte nach standardisierten Richtlinien. Daneben erfolgten bei den Schülerinnen und Schülern zu Beginn, direkt nach den Videoaufzeichnungen und am Ende des Schuljahres Begleiterhebungen. Nach Abschluss der Videoaufzeichnungen wurden die beteiligten Lehrkräfte interviewt. Insgesamt nahmen an der Studie 344 Schülerinnen und Schüler (48,4% weiblich; 51,6% männlich) und 13 Physiklehrkräfte (2 weiblich; 11 männlich) teil.



2.2 Erhebungsverfahren

Der folgende Abschnitt gibt einen Überblick über die eingesetzten Erhebungsinstrumente. Eine ausführliche Darstellung aller Verfahren findet sich bei Prenzel u.a. (2001).

Erhebungsinstrumente auf Schülerseite

(a) *Erhebungen zu Beginn und am Ende des Schuljahres:* Zur Erfassung der Qualität der Bildungsprozesse im Verlauf des Anfangsschuljahres in Physik wurden zu Beginn und am Ende des Schuljahres Fragebogen und Tests zu kognitiven und motivationalen bzw.

affektiven Schülermerkmalen eingesetzt. Im Rahmen dieses Beitrags wird auf zwei Aspekte eingegangen: Den Lernerfolg und die Interessenentwicklung der Schülerinnen und Schüler im Verlauf des Anfangsschuljahres.

- *Lernerfolg*: Die Schülertests zur Erfassung des Lernerfolgs enthalten 24 Items im Vortest sowie 36 im Nachtest. Alle Items wurden dichotom kodiert (0 = nicht richtig; 1 = richtig), und es wurden Prozentwerte gelöster Aufgaben pro Test (Vor- und Nachtest) berechnet. Als Maß für den Lernerfolg werden hier die Residualwerte der Regressionen von Nachtest-Score auf Vortest-Score herangezogen. Dieser Indikator für den Lernzuwachs ist unkorreliert mit der Vortestleistung.
- *Interesse an Physik*: Die Schülerinnen und Schüler wurden zu Beginn und am Ende des Schuljahres zu ihrem Interesse an physikbezogenen Themen befragt. Als Maß für die Interessenentwicklung wurden ebenfalls Residualwerte verwendet.

(b) *Unterrichtsbezogene Schülereinschätzungen*: Zur Rekonstruktion von kognitiven und motivational/affektiven Prozessen beim Lernen schätzten die Lernenden jeweils direkt im Anschluss an die Videoaufzeichnungen *unterrichtliche Lehr-Lernbedingungen, die Qualität der Lernmotivation, kognitive Lernaktivitäten* sowie ihre *emotionale Befindlichkeit* anhand von Skalen ein.

Beobachtungsverfahren

(a) *Organisation unterrichtlicher Aktivitäten („Sichtstrukturen“)*: Die Erfassung von unterrichtlichen Ablauf- und Interaktionsmustern erfolgte anhand differenzierter Beobachtungsverfahren (Seidel 2002; Seidel u.a. 2002). Mittels eines Zeitstichprobenplans wurde das gesamte Videomaterial in 10-Sekunden-Einheiten analysiert, die Werte pro Schulstunde aggregiert und in Minuten umgerechnet. Mit der Auswahl der Kategoriensysteme sollte an bisherige Forschungsarbeiten zu Lehr-Lernskripts im Unterricht (TIMSS-Video) angeknüpft werden. Dazu wurden vorhandene Verfahren an den Anfangsunterricht in Physik angepasst. Die verwendeten Kategoriensysteme erlauben es, Schwerpunktsetzungen in Ablauf- und Interaktionsmustern differenziert zu beschreiben. Insgesamt wurden folgende Kategoriensysteme unterschieden: Unterrichtszeit, unterrichtliche Arbeitsformen, Experimentalformen, Unterrichtsphasen, Aktivitäten während des Klassengesprächs und während Schülerarbeitsphasen. Die Durchführung der Videokodierungen erfolgte auf der Basis von Beobachterschulungen und der Überprüfung der Beobachterübereinstimmung (Cohens Kappa $>.80$; direkte Übereinstimmung $> 85\%$).

(b) *Qualität der Lehr-Lernbedingungen*: Wie oben ausgeführt stehen die folgenden vier Aspekte im Mittelpunkt der hier vorgestellten Studie: Zielorientierung, mentale Unterstützung und Begleitung des Lernens, Umgang mit Schülervorstellungen und Fehlern, Umsetzung und Einbettung von Experimenten. Zur Untersuchung dieser Aspekte wurden Beobachtungsverfahren entwickelt und angewandt, mit denen die genannten Bereiche abgedeckt werden können. Über konsekutive Auswertungen aus variierenden Per-

spektiven werden Anhaltspunkte für die Auswirkungen auf Lernprozesse identifiziert und schließlich an weiteren Daten (Tests und Fragebogen) überprüft. Eine ausführliche Beschreibung der Beobachtungsverfahren mit Kodierungsrichtlinien und Kennwerten findet sich bei Prenzel u.a. (2001).

3. Ergebnisse

3.1 Identifikation von Ablauf- und Interaktionsmustern

Bisherige Projektauswertungen bezogen acht Gymnasialklassen ein (Seidel u.a. 2002). Dabei wurden zwei Muster der Unterrichtsorganisation identifiziert: In Unterrichtsstunden mit Muster I wird der Schwerpunkt auf das Klassengespräch gelegt. Schülerarbeitsphasen haben einen geringeren Anteil. Muster II spiegelt dagegen einen Unterricht wider, bei dem der Schwerpunkt auf einer Kombination von Schülerarbeitsphasen und Klassengespräch liegt. Während das Unterrichtsmuster I eher lehrerzentriert zu sein scheint, würde man Unterrichtsmuster II auf den ersten Blick als stärker schülerzentriert einstufen. Bei beiden beobachteten Unterrichtsmustern spielte allerdings das Klassengespräch zeitlich betrachtet eine wichtige Rolle.

Diese Befunde werden durch die Auswertung der gesamten Stichprobe (N = 75 Stunden, 13 Klassen, Gymnasien und Realschulen) bekräftigt. Tabelle 1 stellt pro Klasse die gemittelten Werte (in Minuten) über alle sechs Unterrichtsstunden dar sowie die Einteilung in die beiden Muster der Unterrichtsorganisation.

Tab. 1: Durchschnittliche Dauer des Klassengesprächs (K) und der Schülerarbeitsphasen (S) in den einzelnen Klassen (in Minuten). Einteilung mittels Mediansplit und Typenzuordnung (Muster I / II)													
	Klassen												
	1	2	3	4	5	6	7	8	9	10	11	12	13
K	42,9	24,2	33,2	38,8	26,8	26,9	21,9	32,9	30,3	27,4	20,4	23,3	20,4
S	0,9	10,1	8,5	5,9	14,0	12,7	16,3	1,6	7,1	13,8	18,3	18,1	17,5
SMd	-	-	-	-	+	+	+	-	-	+	+	+	+
Typ	I	I	I	I	II	II	II	I	I	II	II	II	II
SMd = Mediansplit für Schülerarbeitsphasen; Md = 12,7Min; Typ: I = Schwerpunkt Klassengespräch, Schülerarbeitsphasen < Mediansplit; II = Schwerpunkt Klassengespräch & Schülerarbeitsphasen													

Insgesamt lassen sich bereits auf der unmittelbar beobachtbaren Ebene der Unterrichtsorganisation deutliche Unterschiede zwischen den Klassen feststellen. Vergleicht man die Organisation unterrichtlicher Arbeitsformen beider Schulformen (Gymnasial- und Realschulklassen) bestehen keine nennenswerten Unterschiede.

3.2 Entwicklungen im Anfangsschuljahr Physik

Zur Überprüfung der Unterrichtsqualität werden Schülerdaten zur Leistungs- und Interessenentwicklung im Verlauf des Anfangsschuljahres in Physik herangezogen. Die Erfassung des Lernerfolges erfolgte mittels Leistungstests zu den beiden im videographierten Unterricht behandelten Themen. Der Lernzuwachs der Schülerinnen und Schüler wurde regressionsanalytisch ermittelt. Abbildung 2 stellt die mittleren Lernzuwächse in den einzelnen Klassen dar. Dabei variiert der Lernzuwachs zwischen den Klassen nennenswert, $F_{(12, 272)} = 20,07$, $p < .001$.

Gemäß Abbildung 2 erreichen sieben Klassen der Stichprobe einen Lernzuwachs (positive Werte), während bei sechs Klassen aufgrund des Vorwissenstests höhere Leistungen zu erwarten waren (negative Werte). Betrachtet man die Interessenentwicklung in den einbezogenen Schulklassen, werden ebenfalls Differenzen zwischen den Klassen deutlich, $F_{(12, 266)} = 2,38$, $p < .01$ (Abb. 3).

Einige Schulklassen zeichnen sich durch Zuwachswerte im Bereich des Lernerfolgs und des physikbezogenen Interesses aus (Klassen 1, 10 und 11). Demgegenüber stehen Klassen, bei denen keine positiven Entwicklungen sowohl bzgl. des Lernerfolgs als auch des Interesses festzustellen sind (Klassen 6 und 13). Dazwischen liegen Schulklassen, die beim Lernerfolg im Vergleich zu den anderen Klassen positive Werte erreichen, bei denen aber eine Abnahme des Interesses zu verzeichnen ist (Klassen 3, 9 und 12), und Klassen mit einer positiven Interessenentwicklung aber einem negativen Lernzuwachs (Klassen 2, 5, 7 und 8). Dieser Vergleich zeigt, dass einzelne Entwicklungsbereiche (z.B. Interesse und Wissen) unterschiedlich gefördert werden können und dass sich diese Bereiche nicht gegenseitig ausschließen müssen (Gruehn 1995). So gelang es einigen Klassen, in beiden Bereichen Zuwächse zu erzielen.

3.3 Unterrichtsmuster und die Entwicklungen der Lernenden

Vor dem Hintergrund dieser Befunde stellt sich die Frage, wie der Unterricht in den aufgrund der Leistungs- und Interessenentwicklung als „erfolgreich“ zu bezeichnenden Klassen gestaltet ist. Zieht man die Typisierung der Klassen aufgrund der „sichtbaren“ Gestaltung des Unterrichts heran (Muster I: Schwerpunkt Klassengespräch; Muster II: Kombination Klassengespräch und Schülerarbeitsphasen), lassen sich *keine* nennenswerten Unterschiede in der Leistungs- und Interessenentwicklung feststellen (beide $F < 1$).

Es gibt damit keine Hinweise darauf, dass eine (auf den ersten Blick) stärker lehrer- vs. schülerzentrierte Organisation des Unterrichts (Muster I und II) eine bedeutsame Bedingung für die Entwicklung der Schülerinnen und Schüler auf Klassenebene darstellt. Die Ausrichtung des Unterrichts auf Schüleraktivitäten fördert nicht notwendigerweise die Leistungsentwicklung der Schülerinnen und Schüler. Vielmehr legen die Befunde die Vermutung nahe, dass die Lerngelegenheiten im Unterricht vor allem durch die Qualität der Gestaltung der Lehr-Lernbedingungen bestimmt werden.

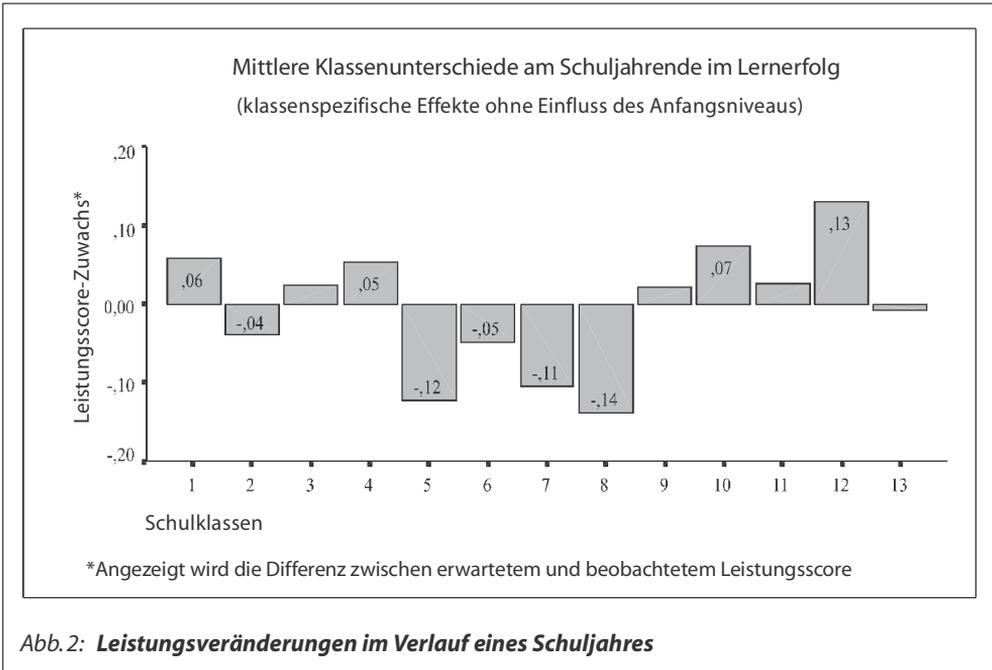


Abb. 2: **Leistungsveränderungen im Verlauf eines Schuljahres**

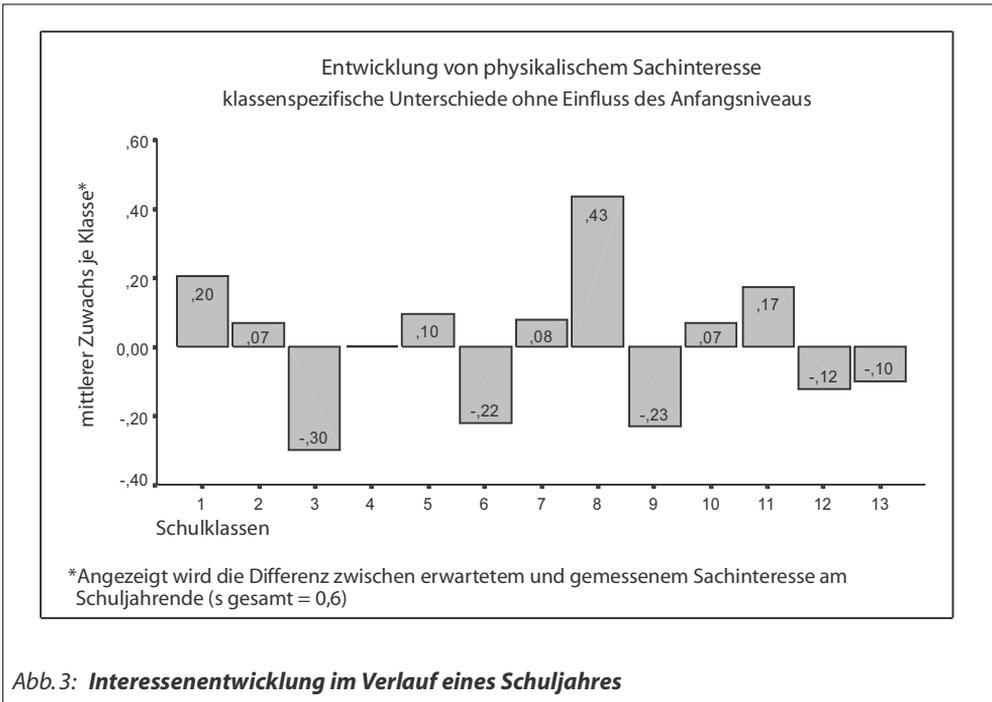


Abb. 3: **Interessenentwicklung im Verlauf eines Schuljahres**

3.4 Qualität der Lehr-Lernbedingungen als Kriterium für Unterrichtsqualität

Bereits veröffentlichte Projektbefunde für die acht Gymnasialklassen deuten auf gemeinsame Merkmale der Schulklassen in den Lehr-Lernbedingungen hin (Seidel u.a. 2002). Die in diesem Beitrag dargestellten Muster der Unterrichtsorganisation belegen bereits die dominante Rolle des Klassengesprächs. Die differenzierteren Auswertungen zu Aktivitäten während des Klassengesprächs und während der Schülerarbeitsphasen mit dem kompletten Datensatz unterstützen darüber hinaus die bisherigen Befunde (Gymnasialklassen) für die gesamte Stichprobe. Ergänzend zu den bei Seidel u.a. (2002) beschriebenen unterrichtlichen Schwerpunktsetzungen werden im Folgenden exemplarisch Befunde zu zwei Aspekten unterrichtlicher „Gelegenheitsstrukturen“ dargestellt: Die Frage nach der Zielorientierung im Unterrichtsverlauf sowie Aspekte der Begleitung des Lernens beim Klassengespräch.

Zielorientierung: Nachvollziehbarkeit und Schlüssigkeit

Die Zielorientierung im Unterrichtsverlauf wird als ein bedeutsamer Indikator für unterstützende Gelegenheitsstrukturen für individuelles Lernen im Unterricht betrachtet. Die Erfassung von Zielorientierung erfolgte im Projekt auf verschiedene Weise: Erstens wurden von den einzelnen Unterrichtsstunden Sachstrukturdiagramme (Duit/Martin/Wachsmuth 2001) erstellt, die u.a. hinsichtlich inhaltlicher Stringenz, Komplexität und Vernetztheit ausgewertet werden. Zweitens wurden die aufgezeichneten Unterrichtsstunden von geschulten Beobachtern hinsichtlich ihrer Nachvollziehbarkeit („Ziel und Ablauf der Stunde sind nachvollziehbar“) und Schlüssigkeit („Die Stunde ist in sich schlüssig und abgeschlossen“) auf einer vierstufigen Skala eingeschätzt (0 = „trifft nicht zu“; 3 = „trifft zu“). Tabelle 2 gibt die Ratings für die jeweils erste Unterrichtsstunde der beiden videographierten Themengebiete in den 13 Schulklassen wieder.

	Klassen												
	1	2	3	4	5	6	7	8	9	10	11	12	13
ZO	2,5	0,5	1,0	2,5	0,0	1,0	1,2	1,5	1,5	2,5	0,5	2,2	2,7
Md	+	-	-	+	-	-	-	-	-	+	-	+	+

Durchschnittliche Ratings für „Nachvollziehbarkeit“ und „Schlüssigkeit“. Skala: 0 = „trifft nicht zu“ bis 3 = „trifft zu“. Md = Mediansplit; Md = 1,5. Einteilung: Nachvollziehbarkeit und Schlüssigkeit „+“ = stärker ausgeprägt; Rating > Mediansplit; „-“ = schwach ausgeprägt, Rating ≤ Mediansplit

Vergleicht man die Schulklassen, die bezüglich der beiden Indikatoren für Zielorientierung über dem Medianwert liegen (+) mit denen unter dem Medianwert (-), dann zeigen sich nennenswerte Differenzen in Bezug auf die Schülereinschätzungen unterrichtlicher Lehr-Lernbedingungen (Seidel u.a. in Vorb.). In den Klassen mit hohen Ratings für Zielorientierung (+) schätzen die Lernenden unterrichtliche Lehr-Lernbedingungen

vergleichsweise positiv ein: In diesen Klassen erkennen die Schülerinnen und Schüler vermehrt inhaltliche Relevanz an den Lerninhalten, sie schätzen die Instruktionsqualität höher ein, erleben die Lehrenden als interessierter, fühlen sich stärker in die Lerngemeinschaft eingebunden sowie in ihrer Kompetenz und in ihrer Autonomie unterstützt. Gleichzeitig erleben sich diese Schülerinnen und Schüler vorwiegend selbstbestimmt motiviert und berichten häufiger darüber, die im Unterricht behandelten Inhalte nachvollzogen, vertieft und gedanklich organisiert zu haben. Vor allem aber erreichen sie im Verlauf des Anfangsschuljahres einen höheren Lernzuwachs als die Schülerinnen und Schüler aus Klassen mit einer weniger ausgeprägten Zielorientierung. Keine Differenzen ergeben sich für die Interessenentwicklung der Lernenden. Die dargestellten Befunde weisen damit darauf hin, dass der über die Videoanalysen ermittelte Aspekt von Zielorientierung einen nennenswerten Indikator für die Qualität unterrichtlicher Lehr-Lernbedingungen darstellt.

Aktive Beteiligung der Lernenden am Klassengespräch

Es wurde bereits deutlich, dass die Organisation des Unterrichts in den untersuchten Klassen erheblich (wenn auch in einem unterschiedlichen Umfang) durch das Klassengespräch geprägt ist. Im Folgenden werden exemplarisch Analysen der Unterrichtsaufzeichnungen in Bezug auf die aktive Beteiligung der Lernenden am Klassengespräch dargestellt (für Details vgl. Seidel 2002). Die intensive Beteiligung der Lernenden stellt im Rahmen dieser Studie ein grundlegendes Kriterium für entstehende Lerngelegenheiten beim Klassengespräch dar. Es wird vermutet, dass eine intensive Beteiligung der Schülerinnen und Schüler am Klassengespräch auf einen wechselseitigen Diskurs zwischen Lehrperson und Klasse hinweist. Der Grad der Schülerbeteiligung wird über die Erfassung der Anteile für das Klassengespräch nicht unmittelbar deutlich, da das Gespräch in unterschiedlichem Ausmaß von der Lehrperson dominiert werden kann. Zur Erfassung der aktiven Schüler- und Lehrerbeteiligung wurden verschiedene Äußerungsformen (z.B. Fragen stellen, Fragen beantworten, Beiträge erläutern) unterschieden und damit die Gesprächsanteile der Schüler bzw. der Lehrpersonen ermittelt.

Die deskriptiven Befunde zu den Äußerungsformen auf Lehrer- und Schülerseite verdeutlichen die enge Verteilung der Gesprächsrollen: Die Lehrkräfte stellen Fragen und erklären Lerninhalte. Die Schülerinnen und Schüler antworten kurz bzw. erläutern ihre Beiträge. Fragen werden beispielsweise von den Lernenden in allen Klassen selten gestellt ($M = 0,4$ Min; $s = 0,6$ Min). Zieht man Befunde zur Elaboration und Organisation von Lerninhalten beim Klassengespräch hinzu, dann zeigt sich, dass im Unterricht insbesondere die Wiedergabe von Fakten und Begriffen betont wird.

Das Ausmaß der aktiven Beteiligung der Schülerinnen und Schüler stellt einen relativ „grobkörnigen“ Indikator für Lernbedingungen während des Klassengesprächs dar. Allerdings deuten Befunde zur Qualität der Schüleräußerungen und zum Umgang der Lehrpersonen mit Schüleräußerungen darauf hin, dass diese in einem erheblichen Ausmaß von der aktiven Beteiligung der Lernenden abhängen (Seidel 2002). Ähnlich wie die genutzte Unterrichtszeit Lerngelegenheiten im Unterricht eröffnet (Seidel u.a. 2002), stellt die aktive Beteiligung der Lernenden die Basis bereit für den lernunterstützenden

Umgang der Lehrpersonen mit den Beiträgen der Schülerinnen und Schüler. Der Vergleich zwischen den Klassen mit einer Schülerbeteiligung über dem Gesamtmittelwert und den Klassen unter dem Gesamtmittelwert zeigt dementsprechend nennenswerte Differenzen bezüglich des Lernzuwachses in diesen Klassen (Seidel u.a. in Vorb.). Demnach erreichen die Klassen mit einer stärker ausgeprägten Schülerbeteiligung höhere Lernzuwächse im Vergleich zu Klassen mit einer niedrigeren Schülerbeteiligung. Dieser Befund unterstützt die Annahme, dass die aktive Beteiligung der Schülerinnen und Schüler u.a. einen weiteren Indikator für Lerngelegenheiten während des Klassengesprächs darstellt.

4. Diskussion

Die dargestellten Verfahren, Auswertungen und Befunde zeigen den Ertrag des gewählten Ansatzes zur Untersuchung von Lehr-Lernprozessen im Unterricht. Die Kombination und Integration verschiedener Perspektiven und Verfahren scheint geeignet zu sein, um das komplexe Wechselspiel unterrichtlicher Prozesse zu durchdringen. Es wurden insbesondere in folgenden Bereichen Fortschritte erzielt:

Mit der Entwicklung und Anpassung von Beobachtungsverfahren zur Erfassung unterrichtlicher Lehr-Lernprozesse ist es gelungen, erste Problemzonen im Physikunterricht zu identifizieren und Hinweise auf ihre Bedeutung für Lernprozesse und Bildungsergebnisse zu erhalten. Die Methodenentwicklung schloss auch die Entwicklung und Anpassung von Schülerfragebogen und -tests sowie Lehrerinterviews ein. Mit dem Abschluss der ersten Projektphase liegt ein differenziertes Instrumentarium zur Untersuchung von Lehr-Lernprozessen im Physikunterricht vor (Darstellung aller entwickelten und verwendeten Verfahren: Prenzel u.a. 2001).

Die analysierten Daten belegen hohe Übereinstimmungen zwischen den über die Videoanalysen gewählten Indikatoren für Unterrichtsqualität und den Einschätzungen der Schülerinnen und Schüler. Die Schülereinschätzungen validieren damit die Ergebnisse der Videoanalysen. Man kann annehmen, dass es sich bei den gewählten Unterrichtsindikatoren unter Umständen um Merkmale unterrichtlicher Lehr-Lernbedingungen handelt, die innerhalb der Lerngemeinschaften „geteilt“ werden.

Die Frage „kulturell geteilter“ Unterrichtsmerkmale wird im Projekt mit dem Konzept der „Unterrichtsskripts“ (Stigler u.a. 1999) bzw. „Choreographien unterrichtlichen Handelns“ (Oser/Patry 1990) aufgegriffen. Es wird angenommen, dass mit diesem Erklärungsansatz sich wechselseitig bedingende Lehr- und Lernprozesse erfasst und untersucht werden können. In der einschlägigen Forschungsliteratur findet man zwar Hinweise auf einzelne unterrichtliche Bedingungsfaktoren. Allerdings fehlen für den Physikunterricht bisher Studien mit entsprechend komplexen Erklärungsansätzen. Zum Projektbeginn gab es beispielsweise noch keine systematische und zuverlässige Datenbasis über vorherrschende Merkmale und Besonderheiten des Physikunterrichts in Deutschland. Deshalb bestand im ersten Projektzeitraum das Ziel, anhand einer ausgewählten Stichprobe von 13 Schulklassen explorativ Muster der Unterrichtsorganisation

und unterrichtlicher Gelegenheitsstrukturen für Lern- und Bildungsprozesse zu untersuchen.

Die empirischen Befunde zu Mustern der Unterrichtsorganisation und unterrichtlicher Lehr-Lernbedingungen eröffnen einen Einblick in Lehr-Lernskripts im Anfangsunterricht in Physik. Der kreidelastige „Demonstrationsunterricht“, von dem Baumert und Köller (2000) in der gymnasialen Oberstufe sprechen, kennzeichnet ebenfalls einen Teil der von uns untersuchten Klassen. Die Form der Unterrichtsführung stellt dabei einen betont lehrergeleiteten Unterricht dar. Baumert und Köller (2000) beschreiben dieses Muster aus Sicht der Schülerinnen und Schüler in der Sekundarstufe II, also am Ende des Physikschulunterrichts. Die Befunde des hier vorgestellten Projekts basieren dagegen auf den Videoanalysen im Anfangsunterricht in Physik. Es zeigt sich, dass der beschriebene „Demonstrationsunterricht“ bereits im ersten Physikschuljahr eine gängige unterrichtliche „Choreographie“ darstellt. Ein zweites Muster der Unterrichtsorganisation, das in seiner äußeren Form stärker schülerzentriert zu sein scheint, findet sich für einen weiteren Teil der Klassen. Zwar nimmt auch hier das Klassengespräch eine dominante Rolle ein, allerdings wird es durch Schülerarbeitsphasen durchbrochen, in denen häufig Schülergruppenexperimente durchgeführt werden.

Betrachtet man die Entwicklungen der Lernenden in den einzelnen Klassen, lassen sich nennenswerte Differenzen feststellen, die jedoch nicht in einem systematischen Zusammenhang zu den beschriebenen Mustern der Unterrichtsorganisation stehen. Vielmehr werden in *beiden* Unterrichtsmustern bei einigen Klassen lernunterstützende Gelegenheitsstrukturen realisiert. Hinweise darauf liefern die dargestellten Ergebnisse zur Beziehung zwischen Indikatoren für Zielorientierung im Unterricht und dem Lernzuwachs: Wird der Unterricht über Videoratings bezüglich „Nachvollziehbarkeit“ und „Schlüssigkeit“ positiv eingeschätzt, erreichen die Schülerinnen und Schüler im Verlauf des Schuljahres eine positive Leistungsentwicklung. Bezogen auf den aktuellen Unterricht berichten sie über unterstützende Lehr-Lernbedingungen, über Formen selbstbestimmter Lernmotivation sowie über ein hohes Maß an kognitiven Lernaktivitäten. Einen weiteren Hinweis auf unterrichtliche Gelegenheitsstrukturen liefern die Auswertungen zur Schülerbeteiligung beim Klassengespräch. Auch hier zeigt sich, dass die erhöhte Beteiligung der Lernenden in Beziehung zu einer positiven Einschätzung der Lehr-Lernbedingungen und zur Leistungsentwicklung im ersten Physikschuljahr steht.

Die Befunde der Videoanalysen lenken den Blick auf besondere Merkmale der Unterrichtsorganisation. Bei einigen Indikatoren, die als entscheidend für unterrichtliche Gelegenheitsstrukturen gelten (Zielorientierung, Begleitung des Lernens, Rolle der Experimente, Umgang mit Schülervorstellungen und Fehlern), werden zwischen den einbezogenen Klassen relativ geringe Streuungen festgestellt. Diese geringen Differenzen zwischen den Klassen weisen u.U. auf kulturell „geteilte“ Merkmale unterrichtlicher Lehr-Lernskripts hin.

In diesem Zusammenhang ist die eingeschränkte Aussagefähigkeit intra-kultureller Beschreibungen von Lehr-Lernskripts zu diskutieren (Seidel 2002). Das dargestellte Projekt war bisher auf eine deutsche Stichprobe beschränkt. Die Befunde der TIMS-Videostudie zu Unterrichtsskripts basieren dagegen auf interkulturellen Vergleichen

(Baumert u.a. 1997; Stigler u.a. 1999). Es besteht dabei die Annahme, dass routinisierte Muster unterrichtlicher Handlungsweisen und Vorstellungen über „gelingenden“ Unterricht innerhalb einer Kultur geteilt und von den einzelnen Mitgliedern nicht mehr bewusst wahrgenommen werden (Reusser 2001; Stigler 2001). Bei der Einbeziehung einer nationalen Stichprobe ist zu berücksichtigen, dass gerade geringe Streuungen zwischen den Klassen besonders auffällige Charakteristika eines Landes widerspiegeln können. Ziel des zweiten Projektzeitraums wird es deshalb sein, die bisherigen Befunde zu unterrichtlichen Charakteristika an einer größeren nationalen Stichprobe zu überprüfen und sie mittels eines Schweiz-Deutschland-Vergleichs interkulturell einzubetten. Die deutschsprachige Schweiz stellt ein interessantes Vergleichsland dar, um den Einfluss wichtiger Kontextfaktoren (z.B. Stellenwert naturwissenschaftlicher Bildung) auf Unterrichtsmuster und den Stand der erreichten naturwissenschaftlichen Grundbildung zu untersuchen (Prenzel u.a. 2001).

Videostudien sind mit einem großen Aufwand verbunden. Dies betrifft die Datenerhebungen ebenso wie die Entwicklung und Anwendung von Beobachtungsverfahren. Trotzdem besteht ein weiter Konsens darüber, dass nur durch die Integration verschiedener Perspektiven erklärungskräftige Erkenntnisse über das Wechselspiel unterrichtlicher Lehr- und Lernprozesse zu erwarten sind. Bisher basieren viele Studien auf unterrichtsbezogenen Einschätzungen der Lehrenden und Lernenden. Diese Einschätzungen liefern wichtige Erkenntnisse über internal ablaufende Prozesse und über wahrgenommene Merkmale von Unterricht. Einschränkungen entstehen allerdings dann, wenn die konkrete Realisation des Unterrichts fokussiert werden soll. Die Identifikation konkreter Unterrichtssequenzen, von Beispielen gelungener und weniger gelungener Lehr-Lernprozesse ist aus Sicht verschiedener BIQUA-Projekte (vgl. auch Beiträge von Fischer, Fischler und Klieme) für die Reflexion und Weiterentwicklung von Unterricht von entscheidender Bedeutung.

Literatur

- Aebli, H. (1998): Zwölf Grundformen des Lehrens. Stuttgart: Klett-Cotta.
- Baumert, J./Lehmann, R./Lehrke, M./Schmitz, B./Clausen, M./Hosenfeld, I./Köller, O./Neubrand, J. (1997): TIMSS-Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen: Leske + Budrich.
- 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.) TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Bd.2. Opladen: Leske + Budrich, S. 271–315.
- Bleichroth, W./Dahncke, H./Jung, W./Kuhn, W./Merzyn, G./Weltner, K. (1999): Fachdidaktik Physik. Köln: Aulis Verlag.
- Clausen, M. (2001): Unterrichtsqualität: eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität. Münster: Waxmann.
- Duit, R./Martin, O./Wachsmuth, J. (2001): Anleitung zur Erstellung von Sachstrukturdiagrammen. In: Prenzel, M./Duit, R./Euler, M./Lehrke, M./Seidel, T. (Hrsg.) Erhebungs- und Auswertungsverfahren des DFG-Projekts „Lehr-Lern-Prozesse im Physikunterricht – eine Videostudie“. Kiel: IPN, S. 135–142.

- Euler, M. (2001): Lernen durch Experimentieren. In: Ringelband, U. /Prenzel, M./Euler, M. (Hrsg.) Lernort Labor. Initiativen zur naturwissenschaftlichen Bildung zwischen Schule, Forschung und Wirtschaft. Kiel:IPN, S. 13–42.
- Fischler, H. (2002): Verfahren zur Erfassung von Lehrer-Vorstellungen zum Lehren und Lernen in den Naturwissenschaften. In: Zeitschrift für Didaktik der Naturwissenschaften, 7, S. 105–120.
- Gruehn, S. (1995): Vereinbarkeit kognitiver und nichtkognitiver Ziele im Unterricht. In: Zeitschrift für Pädagogik, 4, S. 531–553.
- Harlen, W. (1999): Effective teaching of science. Edinburgh, UK: The Scottish Council for Research in Education (SCRE).
- Helmke, A./Weinert, F.E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Weinert, F.E. (Hrsg.) Enzyklopädie der Psychologie. Bd.3: Psychologie des Unterrichts und der Schule. Göttingen: Hogrefe, S. 71–176.
- Hoffmann, L. (2002): Promoting girls` interest and achievement in physics classes for beginners. In: Learning and Instruction, 12, S. 447–465.
- Hoffmann, L./Häußler, P./Lehrke, M. (1998): Die IPN-Interessenstudie Physik. Kiel: IPN.
- Jacobs, J.K./Kawanaka, T./Stigler, J.W. (1999): Integrating qualitative and quantitative approaches to the analysis of video data on classroom teaching. In: International Journal of Educational Research, 31, S. 717–724.
- Klieme, E. (1999): Unterrichtsqualität und mathematisches Verständnis in verschiedenen Unterrichtskulturen. Projektantrag an die DFG. Berlin: MPI für Bildungsforschung.
- Lehtinen, E. (1994): Institutionelle und motivationale Rahmenbedingungen und Prozesse des Verstehens im Unterricht. In: Reusser, K. /Reusser-Weyeneth, M. (Hrsg.) Verstehen. Psychologischer Prozess und didaktische Aufgabe. Bern: Huber, S. 143–162
- Lunetta, V. (1998): The school science laboratory: historical perspectives and contexts for contemporary teaching. In: Fraser, B./Tobin, K. (Hrsg.) International handbook of science education. Dordrecht: Kluwer, S. 249–262.
- Oser, F./Patry, J.-L. (1990): Choreographien unterrichtlichen Lernens. Basismodelle des Unterrichts. Universität Fribourg.
- Pippig, R./Schneider, W. B. (1995): Beiträge der Physikdidaktik zum Experimentalunterricht und deren Akzeptanz. In: MNU, 48(1), S. 45–49.
- Prenzel, M. (2000). Steigerung der Effizienz des mathematisch- naturwissenschaftlichen Unterrichts: Ein Modellversuchsprogramm von Bund und Ländern. In: Unterrichtswissenschaft, 28(2), S. 103–126.
- Prenzel, M. u.a. (1999): Die Bildungsqualität von Schule: Fachliches und fächerübergreifendes Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten. Antrag an die DFG.
- Prenzel, M./Drechsel, B./Kramer, K. (1998): Lernmotivation im kaufmännischen Unterricht: Die Sicht von Auszubildenden und Lehrkräften. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 14, S. 169–187.
- Prenzel, M./Duit, R./Euler, M./Lehrke, M.(2001): „Lehr-Lern-Prozesse im Physikunterricht: Eine Videostudie“. Fortsetzungsantrag an die DFG. Kiel: IPN.
- Prenzel, M./Duit, R./Euler, M./Lehrke, M./Seidel, T. (Hrsg.) (2001): Erhebungs- und Auswertungsverfahren des DFG-Projekts „Lehr-Lern-Prozesse im Physikunterricht – eine Videostudie“. Kiel: IPN.
- Prenzel, M./Rost, J./Senkbeil, M./Häußler, P./Klopp, A. (2001): Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske & Budrich, S. 191–248.
- Reusser, K. (2001): The challenge of video-based research to cognitive instructional theory. Tagungsbeitrag: Konferenz der European Association for Research on Learning and Instruction (EARLI), Fribourg, Schweiz.

- Reusser, K. (1998): Denkstrukturen und Wissenserwerb in der Ontogenese. In: Klix, F./Spada, H. (Hrsg.) Enzyklopädie der Psychologie. C: Theorie und Forschung, II Kognition, Bd. 6 (Wissen). Göttingen: Hogrefe, S. 115–166.
- Seidel, T. (2002): Lehr-Lernskripts im Unterricht. Freiräume und Einschränkungen für kognitive und motivationale Prozesse beim Lernen – eine Videostudie im Physikunterricht. Universität Kiel: Unveröffentlichte Dissertation.
- Seidel, T./Prenzel, M./Duit, R./Euler, M./Geiser, H./Hoffmann, L./Lehrke, M./Müller, C./Rimmele, R. (2002): „Jetzt bitte alle nach vorne schauen!“ – Lehr-Lernskripts im Physikunterricht und damit verbundene Bedingungen für individuelle Lernprozesse. In: Unterrichtswissenschaft, 30(1), S. 52–77.
- Seidel, T./Prenzel, M., /Rimmele, R. (in Vorb.): Investigating teaching and learning processes by combining multiple perspectives – What to learn from videotape classroom studies?
- Shuell, T.J. (1996): Teaching and learning in a classroom context. In: Berliner, D.C./Calfee, R.C. (Hrsg.) Handbook of Educational Psychology. New York: Macmillan, S. 726–764.
- Staub, F./Stern, E. (in Druck): The nature of teachers` pedagogical content beliefs – matter for students` achievement gains: quasi-experimental evidence from elementary mathematics. In: Journal of Educational Psychology.
- Stigler, J.W. (2001): Using video surveys to compare classrooms and teaching across cultures: Foundations and results of the TIMSS and TIMSS-R video studies. Paper presented at the conference of the European Association for Research on Learning and Instruction (EARLI), Fribourg, Schweiz.
- Stigler, J.W./Gallimore, R./Hiebert, J. (2000): Using video surveys to compare classrooms and teaching across cultures: examples and lessons from the TIMSS Video Studies. In: Educational Psychologist, 35(2), S. 87–100.
- Stigler, J.W./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. Washington, D.C.: U.S. Department of Education.

Anschrift der Autoren:

- Prof. Dr. Manfred Prenzel, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Dr. Tina Seidel, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Dr. Manfred Lehrke, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Rolf Rimmele, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Prof. Dr. Reinders Duit, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Prof. Dr. Manfred Euler, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Helmut Geiser, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Dr. Lore Hoffmann, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Christoph Müller, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.
Ari Widodo, IPN, Universität Kiel, Olshausenstr. 62, 24098 Kiel.

Helmut Fischler/Hans-Joachim Schröder/Cornelia Tonhäuser/Peter Zedler

Unterrichtsskripts und Lehrerexpertise: Bedingungen ihrer Modifikation¹

1. Zielstellung des Projekts

Das Forschungsprojekt geht der Frage nach, wie sich ausgewählte, erfahrungsbewährte Anschauungen von Physiklehrern über die Gestaltung des (Physik-)Unterrichts zielgerichtet ändern lassen und wie sich Änderungen in zentralen didaktisch-methodischen Kognitionsfeldern auf das unterrichtliche Handeln der Lehrer sowie darüber auf die Schülerleistungen auswirken. Mit dieser Fragestellung ist die Grundannahme verbunden, dass das unterrichtliche Handeln von Lehrern in Subjektiven Theorien verankert ist, mit denen typische Situationen und Abläufe des Unterrichts interpretiert werden und die ihrerseits mit einem (erfahrungsbewährten) Inventar von Handlungsoptionen verknüpft sind. Soll die Qualität des Unterrichts verbessert werden, ist es unverzichtbar, die „Expertise“ der Lehrer ebenso wie davon abhängige Muster in der Gestaltung des Unterrichts (Unterrichtsskripte) zu ändern. Mit Blick auf das Ziel einer Erweiterung fachdidaktisch relevanter Urteilsstrukturen und situationstypischer Verhaltensweisen wird im Rahmen des Projekts experimentell untersucht, welche Formen und Verfahren einer unterstützenden Beratung (Coaching) eine Transformation von Kognitions- und Handlungsstrukturen begünstigen und wie sich erfolgte Änderungen auf den Lernerfolg von Schülern auswirken.

Die Zielstellung des Projekts schließt an Befunde sowohl der allgemeinen und fachdidaktischen Lehr-Lernforschung als auch der Lehrerforschung an. Nachdem sich in den vergangenen Jahren das Forschungsinteresse verstärkt auf schülerseitige Bedingungen des Leistungserfolgs (Alltagskonzepte, motivationale Faktoren, Selbstwirksamkeitserwartungen, Elterneinfluss etc.) gerichtet hat, lässt sich im Gefolge der TIMS-Studie ein zunehmendes Interesse an der „Lehrervariable“ feststellen. Nicht ohne Grund: Wie insbesondere die Ergebnisse der TIMS-Videostudie nahe legen, erklärt die vom Lehrhandeln abhängige Gestaltung des Unterrichts ein erhebliches Maß an der feststellbaren Leistungsvarianz von Schülern. Wird davon ausgegangen, dass die Gestaltung des Unterrichts wesentlich von den Vorstellungen der Lehrer über pädagogisches und didaktisch-methodisch angemessenes Handeln und dessen Wirkungen auf die Lernprozesse der Schüler abhängt, liegt es nahe, in der Lehrerexpertise und den darin verankerten Handlungsoptionen einen zentralen Ansatz für die Veränderung der Qualität des Un-

1 Die Studie wurde gefördert durch Mittel der DFG (FI 579/8-1 /FI 579/8-2, ZE 249/3-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

terrichts zu sehen. Die Expertise von Lehrern ebenso wie deren unterrichtliche Arbeitsroutinen als Determinanten des Leistungserfolgs von Schülern zu betrachten, ist nicht neu. Wiederholt wurde in Untersuchungen festgestellt und kritisch vermerkt, dass das Methodenrepertoire von Lehrern überwiegend nur eine geringe Bandbreite aufweist, Unterricht durch eine „didaktische Monostruktur“ geprägt ist (vgl. Bauer u.a. 1979; Hage u.a. 1985). Ebenfalls wurde bereits häufig festgestellt, dass Ergebnisse der Lehr-Lernforschung ebenso wie der Fachdidaktik nicht in dem Maße von Lehrern zur Kenntnis genommen und genutzt werden, wie dies wünschenswert erscheint. Zum anderen wird gleichwohl häufig ein nicht unerheblicher Beratungsbedarf der Lehrer im Hinblick auf den Umgang mit Aufgaben und Anforderungen des Unterrichts konstatiert (vgl. Spanhel/Hübner 1995; Combe/Buchen 1996; Böhm-Kasper u.a. 2001). Es stellt sich mit Blick auf die Befunde zur Rezeption und Implementation pädagogisch-psychologischer und fachdidaktischer Forschungsergebnisse die grundsätzliche Frage, auf welchem Wege eine stärkere Kopplung zwischen vorhandenem Wissen und unterrichtlicher Praxis erreicht werden kann. So legen eine Reihe von Forschungsbefunden zu Weiterbildungskursen und Trainings von Lehrern nahe, dass herkömmliche Formen der Lehrerweiterbildung in der Regel nicht dazu führen, die Unterrichtspraxis nachhaltig zu verändern.

Sowohl Berichte über erfolgreiche Bemühungen, Veränderungen von Lehrervorstellungen zu bewirken (Hand/Treagust 1994), als auch über Misserfolge (Tillema/Knol 1997; Yerrick u.a. 1997) resümieren übereinstimmend, dass nur solche Programme zur Modifizierung von Lehrervorstellungen und Lehrerhandeln eine Erfolgchance besitzen, die ihren Ausgang von den individuellen Erfahrungen und Vorstellungen der Lehrer nehmen (Tillema 1994). Überwiegend wird hieraus die Forderung abgeleitet, dass fachdidaktische Lehrveranstaltungen in der Lehrerausbildung Merkmale aufweisen sollten, die denen eines „konstruktivistisch“ orientierten naturwissenschaftlichen Unterrichts ähnlich sind (Fensham 1987; Marion u.a. 1994; Parsons-Chatman 1990), das heißt unter Bezug auf je gegebene Vorstellungen und Erfahrungen der Lehrer neue Vorstellungen zu konstruieren erlauben. Dieses Prinzip wird im Allgemeinen auch für die Arbeit mit erfahrenen Lehrern als richtungsweisend erachtet (z.B. Thonhauser 1987).

In vielen Forschungsberichten wird konstatiert, dass die gesetzten Ziele von Aus- und Fortbildungskursen nicht erreicht wurden (Gustafson/Rowell 1995; Tillema/Knol 1997; Yerrick u.a. 1997; Tillema 2000; Boulten-Lewis u.a. 2001). Weisen einige Interventionsstudien darauf hin, dass sich zentrale pädagogisch-didaktische Vorstellungen und Überzeugungen einer intendierten Veränderung entziehen, so zeigen andere Studien, dass selbst dort, wo eine Änderung pädagogisch-didaktischer Vorstellungen erfolgreich ist, ungewiss bleibt, ob die Änderungen Auswirkungen auf die Unterrichtspraxis haben.

Hewson/Hewson (1997) konstatieren nach einem Kurs im Rahmen der Lehrerausbildung, in dem die Bedeutung von Schülervorstellungen für den Unterricht betont wurde, eine Veränderung im Denken der Lehrerstudenten in Richtung auf ein „conceptual change model“ des Unterrichtens. Aber: „The most one can say is that the student teachers found these activities intelligible. The results do not guarantee that they found them plausible or fruitful, i.e. that they would employ them in the classroom“ (S. 436).

Lässt sich einerseits dem Forschungsstand entnehmen, dass sich nachhaltige Erfolge bei der Veränderung von leitenden Handlungsorientierungen im Rahmen herkömmlicher Kurse und Trainings nur selten erreichen lassen, so ist andererseits zugleich festzustellen, dass in bisherigen Untersuchungen nicht geklärt wurde, ob und in welchem Maße geäußerte Vorstellungen über didaktisch angemessenes Handeln dem tatsächlichen Handeln der Lehrer im Unterricht korrespondieren. Bekannt ist lediglich, dass Differenzen zwischen Absicht und Handeln häufig in Situationen zu finden sind, in denen „ein Handeln unter Druck“ (vgl. Wahl 1991) stattfindet, das bedingt durch Faktoren wie z.B. Zeitknappheit und Stofffülle zu anderen als den selbst für angemessen erachteten Reaktionen führt.

Vor dem Hintergrund der genannten Befunde wurde für das Forschungsprojekt ein forschungsmethodischer Rahmen entwickelt, der es erlaubt,

- den pädagogischen und didaktisch-methodischen Referenzrahmen der Lehrer für (fach-)typische Situationen des Physikunterrichts zu erfassen,
- von den Lehrern benannte Handlungsweisen für die Gestaltung fachdidaktisch typischer Unterrichtssituationen auf ihre faktisch vorhandene Typizität im Unterricht zu kontrollieren,
- fachdidaktisch-problematische oder dysfunktionale Kognitionen und Verhaltensweisen mittels spezifischer Beratungsverfahren zu korrigieren und
- Beratungsinterventionen auf ihre Wirksamkeit hinsichtlich einer Erweiterung der didaktischen Expertise, der Veränderung von Unterrichtsskripts sowie davon abhängiger Schülerleistungen zu kontrollieren.

2. Interventionsspezifik: Beratung als fachdidaktisch orientiertes Coaching

Im Bereich der Wirtschaft ist es Alltag, dass geänderte Anforderungen im beruflichen Aufgabenfeld bisherige Routinen und Denkweisen in Frage stellen und Arbeitsplatzinhaber noch nicht in zureichendem Maße über die Voraussetzungen verfügen, die für die Erfüllung jeweiliger Aufgaben wünschenswert erscheinen. Den erwähnten Befunden über Lehrertrainings analog gelagerte Erfahrungen zu den Transferproblemen von Weiterbildungsmaßnahmen führten dazu, dass im Bereich der Beratung von Unternehmen Formen und Verfahren einer individualisierten Stützung der Kompetenzerweiterung entwickelt wurden, die unter dem Begriff des „Coaching“ firmieren. Coaching ist „personenbezogene Einzelberatung von Menschen in der Arbeitswelt“ (Looss 1991, S. 13), eine „professionelle Form der Managementberatung“ (Schreyögg 1999, S. 7), die „Beratung von Führungskräften, Experten, Mitarbeitern bei der Erreichung von Zielen im beruflichen Bereich“ (König/Volmer 2002, S. 11). Coaching erfolgt als Prozess- und Expertenberatung, das heißt die Beratung erfolgt sowohl als methodisch-kontrollierter Prozess der Unterstützung des Klienten bei der Gewinnung von Problemlösungen als auch als Anregung des Beraters in Hinblick auf jeweilige Problemlösungsalternativen.

Theoretische Grundlage in der Methodik der Prozessberatung sind durchgängig Ansätze und Konzepte, die insbesondere im Bereich der kognitiven Verhaltenstherapie (vgl. Lückert/Lückert 1994; Ellis 1993; Dryden 1995; Bandler/Grinder 1992; Meichenbaum 1994) sowie im Feld systemischer Beratung (vgl. Watzlawick u.a. 1972, 1997; König/Volmer 1996) entwickelt wurden. Diese Ansätze werden – entsprechend den Unterschieden in Veranlassung und Zielstellung – im Coachingprozess nur in einer zeitlich und methodisch abgeflachten Version verwendet. So unterschiedlich die Ansätze und Konzepte hinsichtlich des Zugriffs auf eine Veränderung kognitionsabhängiger Verhaltensweisen sind, ihnen gemeinsam ist, dass sie den für ein Problem konstitutiven Referenzrahmen und ein damit je konnotiertes Verhalten fokussieren und methodisch kontrolliert zu transformieren erlauben.

Für die Verwendung der verschiedenen Ansätze und Konzepte des Coachings, das heißt für den Weg der Beratung mit dem Ziel einer Erweiterung der didaktischen Expertise sowie der Möglichkeiten einer pädagogisch-didaktischen Gestaltung des Unterrichts von Lehrern, sind drei Punkte zentral:

- a) Ausgangspunkt für die Beratung sind (selbst- oder fremdbeobachtete) Einschränkungen hinsichtlich der Aufgabenerfüllung; geäußert z.B. als Differenz von Absicht und Wirkung des Unterrichts seitens des Lehrers, als Unzufriedenheit mit der eigenen Rolle, mit dem (Miss-)Erfolg des Unterrichts oder mit dem pädagogischen und didaktischen Verhaltensrepertoire im Umgang mit wiederholt auftretenden Problemsituationen (z.B. unmotivierte oder störende Schüler).
- b) Bezugspunkt der Beratung sind fachdidaktisch typische Situationen und Phasen des Unterrichts. Expertenberatung, verstanden als Anregung von Problemlösungsalternativen (Denk- und Handlungsmöglichkeiten), orientiert sich an den pädagogischen und fachdidaktischen Problemen des Lehrers und dem Wissen über die fallbezogen erkennbaren Wirkungszusammenhänge von pädagogischen und didaktisch-methodischen Denk- und Handlungsweisen.
- c) Der Beratungsprozess hat sein Ziel erreicht, wenn das Ausgangsproblem in einer vom Lehrer als erfolgreich erachteten Weise bewältigt wird.

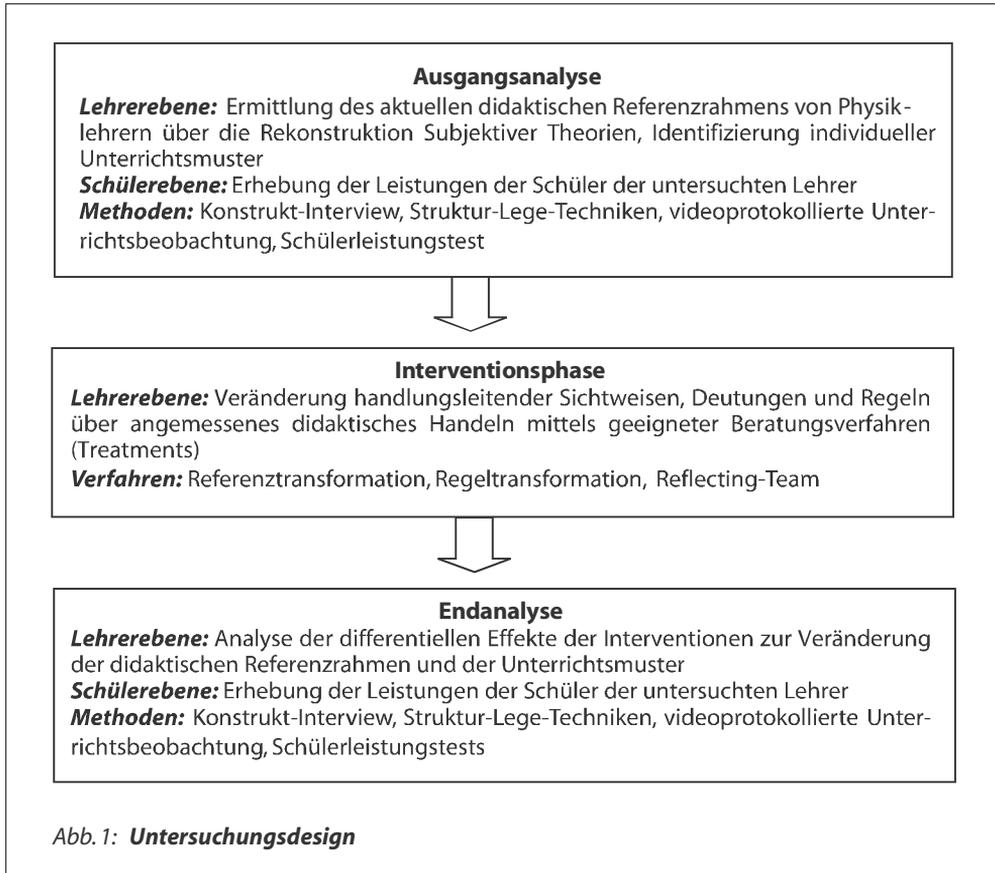
Mit Blick auf die genannten Maßgaben wurden in Vorbereitung der Beratungsgespräche in den bisherigen Fallstudien

- Schwachstellen des Unterrichts aus der Sicht der ihn durchführenden Lehrer ermittelt und die Lehrer daraufhin befragt, welche didaktischen Kompetenzwünsche vorhanden sind, verknüpft mit Nachfragen zum Belastungserleben durch den Unterricht; entsprechende Angaben dienen als Gesprächsanlässe in der Beratung sowie als Grundlage der Auswahl und Priorisierung von Interventionszielen;
- erste Schnittstellen der verschiedenen Coachingansätze und Verfahren bestimmt, die eine symptom- und zielabhängige Verwendung verschiedener Verfahren zu begründen erlauben und damit eine zielgerichtete Differenzierung der geplanten Interventionsmaßnahmen zur Erweiterung der didaktischen Kompetenz von Lehrern unter Bezug auf definierte Schwachstellen der Unterrichtsgestaltung zulassen.

3. Methoden

3.1 Untersuchungsdesign und Erhebungsmethodik

Die zunächst als explorative Untersuchung angelegte Studie folgt einem dreiphasigen Design (vgl. Abb. 1).



3.2 Instrumente und Verfahren

Interview

Für die Durchführung des Projektes ist es von entscheidender Bedeutung, solche Subjektiven Theorien zu erfassen, die handlungswirksam sind. Mit dieser Zielsetzung ist die Auswahl der Erhebungsmethode in entscheidender Weise festgelegt. Da in einem Interview durch gezieltes Nachfragen die Situationsspezifität der Lehrerkognitionen gut erfasst und die Begriffsinterpretation des Probanden genau geklärt werden kann, wird es in dieser Untersuchung als Erhebungsmethode verwendet.

In dem eingesetzten Interviewverfahren dienen konstruierte videografierte Unterrichtsszenen als Gesprächsanlässe. Auf diese Weise soll eine möglichst große Situations- und Handlungsnähe der Lehreraussagen erreicht werden, ohne die befragten Lehrer in einen Rechtfertigungsdruck zu bringen, wie er eventuell bei der Kommentierung eigenen Handelns im Unterricht (z.B. beim Stimulated Recall Interview) entstehen kann.

Die Konstruktion der Unterrichtsszenen, die typologische Auswahl der in den Situationen thematisierten methodisch-didaktischen Probleme und die Gestaltung der Situationen orientieren sich an zwei Bezugsfeldern:

- Die Situationen sprechen die ausgewählten Kognitionsbereiche (s.u.) der Lehrer an und
- sie beziehen sich auf die gängige Praxis des Physikunterrichts mit ihrer organisatorischen, inhaltlichen und methodischen Vielfalt.

Anhand dieser beiden Bezugsfelder wurde eine Matrix mit 14 Zeilen (Kognitionsbereichen) und 10 Spalten (typische Unterrichtssituationen) erzeugt, deren einzelne Felder durch neun konstruierte und von Schülern und Lehrern gespielte drei- bis fünfminütige Videoszenen abgedeckt werden.

In einer dieser Szenen wird zum Beispiel gezeigt, wie ein an der Tafel stehender Schüler am Beginn einer Unterrichtsstunde Lernprobleme aufgrund stabiler Alltagsvorstellungen hat, die trotz mehrerer Unterrichtsstunden zu diesem Thema nicht ausgeräumt werden konnten. Typische Fragen im Anschluss an die Präsentation dieser Szene sind u.a.:

- Wie kann man sich solche Verständnisprobleme der Schüler erklären?
- Wie würden Sie die Unterrichtsstunde fortsetzen?

Die Szene stellt eine typische Wiederholungsphase am Stundenbeginn dar und berücksichtigt im Zusammenhang mit den Fragen die Kognitionsbereiche: a) Vorstellungen über das Lernen von Physik und b) Vorstellungen über das Lehren von Physik (s. u.).

Es wird erwartet, dass mit der bildhaften Darstellung der standardisierten Unterrichtsszenen wesentlich mehr Informationen über die Gesamtheit von Kontextvariablen angeboten werden können als mit jeder anderen Darstellung (z.B. verbale Beschreibung von Szenen), die ohne Reduktion der Komplexität des Unterrichtsgeschehens nicht möglich ist.

Kognitionsbereiche

Für die Ausdifferenzierung des komplexen Konstrukts „Subjektive Theorie“ wurden Kognitionsbereiche ausgewählt, deren Eingrenzungen aufgrund von Analysen der forschungsbasierten lehrerbezogenen fachdidaktischen Literatur vorgenommen wurden. Berücksichtigt wurden u.a. (1) Empfehlungen für das Lehrerhandeln, die die fachdidaktische Forschung als Folgerungen aus ihren Untersuchungen über Schülervorstellungen zu naturwissenschaftlichen Begriffen und Zusammenhängen nahe legt (Überblick in

Duit/Treagust 1998), (2) die in der Folge der TIMS-Studie erarbeiteten Vorschläge zur Verbesserung des Unterrichts (BLK 1997), (3) die insbesondere in angelsächsischen Forschungsgruppen vorgenommenen Analysen zur Lehrerkompetenz und Untersuchungen der Bedingungen für ihre Veränderung (Übersichten bei Clark/Peterson 1986; Pajares 1992) und (4) die für die Hand des Lehrers gedachte allgemein- und fachdidaktische Literatur zur Planung, Durchführung und Analyse des Unterrichts (Beispiele: Bleichroth u.a. 1999; Häußler u.a. 1998; Kircher u.a. 2000).

Die Analyse dieser Bezugsfelder begründet in ausreichendem Maße die hypothetische Bestimmung von fünf Bereichen von Lehrerkognitionen, deren genauere Untersuchung für die Ziele des Projektes als wichtig erscheint (vgl. Fischler 2001):

- 1) Vorstellungen über das *Lernen* von Physik: Explizite und implizite Theorien über den Ablauf und die Beeinflussbarkeit von Lernprozessen im Fach Physik unter Berücksichtigung von Alltagsvorstellungen der Schüler, geschlechts- und altersspezifischen Unterschieden und Leistungsattributionen;
- 2) Vorstellungen über das *Lehren* von Physik: Vorstellungen über Ziele des Physikunterrichts, didaktisch-methodische Strategien und deren Auswirkungen sowie über den Umgang mit „Lernproblemen“ bzw. „Lernhemmnissen“, insbesondere über den Umgang mit Alltagsvorstellungen;
- 3) *Wissenschaftstheoretische Vorstellungen*: Wissenschaftstheoretischer Hintergrund, insbesondere Bedeutung des Experiments für die Erkenntnisgewinnung, Konsequenzen für den Unterricht;
- 4) *Selbstsicht* als Lehrer: Selbstverständnis hinsichtlich der primären Aufgaben des Lehrers (Wissensvermittler, Lehrplanverwalter, Unterstützer von selbsttätigen Lernprozessen, Entwicklungsberater für Lernprozesse u.a.);
- 5) *Kompetenzwünsche* der Lehrer: Rollenselbstverständnis, Restriktionen für eine Verbesserung der Schülerleistungen im Fach Physik (restriktive Randbedingungen in Schule und Unterricht, z.B. Selektivität der Schülerklientel, Stofffülle des Lehrplans, Ausstattungsdefizite).

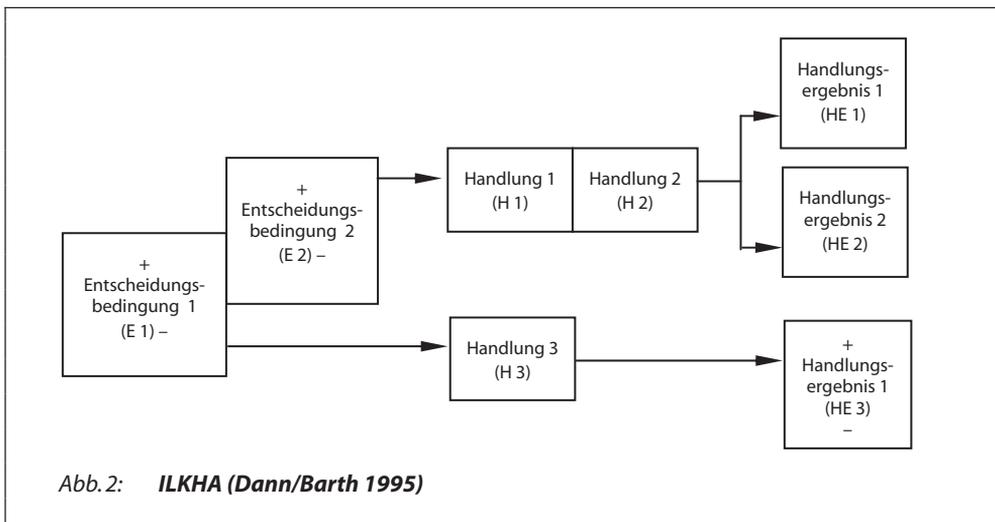
3.3 Rekonstruktion der Subjektiven Theorien und deren Strukturdarstellungen

Die an den videografierten Unterrichtssituationen orientierten Interviewfragen haben zur Folge, dass ein großer Teil der Äußerungen der interviewten Lehrer auf diese Situationen bezogen und daher sehr handlungsorientiert sind. Die zum Zwecke ihrer besseren Kommunizierbarkeit und der Fokussierung auf grundlegende Aussagen vorgenommene Strukturierung dieser Äußerungen erfolgt aus diesen Gründen auf der Basis einer operativen Struktur, in der die von dem Interviewten als Reaktion auf die vorge-spielten Szenen und auf die anschließenden Fragen genannten Handlungsentwürfe abgebildet werden. Die aus den operativen Strukturen der Unterrichtsgestaltung erschlossenen situationsübergreifenden Handlungsmuster, die sowohl Deutungs- und Erklärungsmuster sowie Begründungen enthalten, werden als Elemente Subjektiver Theorien

größerer Reichweite betrachtet. Von diesen Subjektiven Theorien wird angenommen, dass sie die im Interview geäußerten Handlungsalternativen generieren.

Für die Darstellung der operativen Struktur der Unterrichtsgestaltung wird als Verfahren die *Interview- und Legetechnik zur Rekonstruktion kognitiver Handlungsstrukturen* (ILKHA, Dann/Barth 1995) verwendet, die einerseits den konkreten Handlungsablauf berücksichtigt und andererseits darüber hinaus bereits auf situationsübergreifende Strukturen verweist. Mit dieser ILKHA-Struktur (Abb. 2) werden ausgehend von verschiedenen Entscheidungsbedingungen, die teilweise von den Unterrichtsszenen vorgegeben sind und teilweise von den Befragten hinzugefügt werden, Handlungsabfolgen und erhoffte Handlungsergebnisse dargestellt, die von den Lehrern im Interview genannt wurden.

Beispiel: Wenn ein Schüler am Ende einer Unterrichtssequenz Fehlvorstellungen äußert (E1+) und sich die Mehrzahl der Schüler dieser Vorstellung anschließt (E2+), dann ermahne ich die Schüler, im Unterricht besser aufzupassen (H1) und wiederhole die entsprechenden Fakten in einem Lehrervortrag (H2), um die Fehlvorstellungen abzubauen und den weiteren Unterrichtsverlauf zu sichern (HE1 und HE2). Werden von keinem der Schüler Fehlvorstellungen geäußert (E1-), dann erfrage ich sie auch nicht (H3), damit mir der rote Faden im Unterricht nicht verloren geht (HE3).



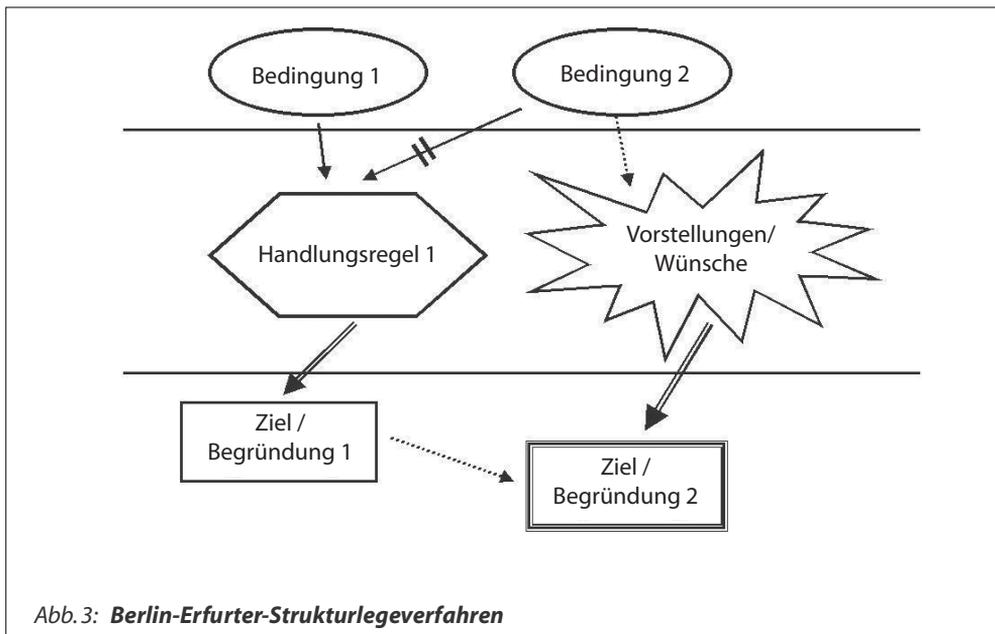
Die Prüfung der in der Literatur beschriebenen Verfahren zur Rekonstruktion Subjektiver Theorien ergab, dass keiner dieser Vorschläge verwendbar war, da sie entweder so kompliziert in der Handhabung sind, dass den an der Untersuchung teilnehmenden Lehrern eine Einarbeitung nicht zugemutet werden konnte, oder die Fragestellung der Untersuchung nicht hinreichend abdeckten. Ein für das Projekt geeignetes Verfahren sollte:

- wegen der Vielzahl der für jeden Probanden zu erfassenden Subjektiven Theorien zeitökonomisch einsetzbar sein und
- eine Darstellungsform bieten, die es ermöglicht, in der Analyse des (videografierten) Unterrichts der Probanden die handlungswirksamen Subjektiven Theorien zu erschließen.

Unter diesen Prämissen wurde eine Struktur für die Darstellung Subjektiver Theorien entwickelt, die drei Beschreibungsebenen enthält:

- Ebene der Handlungsvoraussetzungen und -bedingungen,
- Ebene der Handlungsregeln und Kompetenzwünsche,
- Ebene der Handlungsziele und Handlungsbegründungen.

Da auf der oberen Ebene die Voraussetzungen und Bedingungen so verallgemeinert werden, dass nicht nur einzelne Situationen (z.B.: Ein bestimmter Schüler aus der Klasse 8d stört den Unterricht.), sondern umfassende Situationsklassen (z.B.: Unterrichtsstörungen in der SEK I) erfasst werden und auf der mittleren Ebene situationsübergreifende Handlungsregeln dargestellt sind, beschreibt dieses Verfahren (Berlin-Erfurter-Strukturlegetechnik, BEST; Abb. 3) Subjektive Theorien größerer Reichweite.



Beispiel: Trifft Bedingung 1 zu (Unterrichtsstörungen in der SEK1), nicht aber Bedingung 2 (Schüler sind überfordert), dann muss man als Lehrer nach der Handlungsregel 1 verfahren (Man muss die Anforderungen an die Schüler erhöhen), um das Ziel 1

(Aufmerksamkeit der Schüler) zu erreichen und somit auch Ziel 2 (effektiver Unterricht) zu sichern. Wenn nur die Bedingung 2 zutrifft, dann wünscht sich der Lehrer Teilungsunterricht, um das Ziel 2 zu erreichen.

3.4 *Kommunikative Validierung*

Die Vielzahl der in dem Interview angesprochenen Kognitionsbereiche hat zur Folge, dass eine große Zahl an rekonstruierten Subjektiven Theorien induziert wird. Da diese Rekonstruktion durch die Forscher erfolgt, muss im weiteren Verlauf der Untersuchung sichergestellt werden, dass die erstellten Strukturen auch tatsächlich den Ansichten des Lehrers entsprechen. Außerdem bietet das Validierungsverfahren die Möglichkeit, in dem Interview genannte Rechtfertigungen bzw. sozial erwünschtes Antwortverhalten des Probanden zu identifizieren. Für diesen Prozess der Validierung musste ein Verfahren gewählt werden, das zum einen den Korpus aller Subjektiven Theorien (zu den einzelnen Kognitionsbereichen) lückenlos erfasst und zum anderen das Zeitbudget der Probanden nicht übermäßig belastet. Für die Validierungsphase im Dialog-Konsens-Verfahren wurde daher ein zweistufiges Verfahren gewählt: (1) Präsentation der mit ILKHA dargestellten operativen Struktur der Handlungsentwürfe der Probanden und ihre Diskussion und (2) gemeinsame Strukturlegung der Subjektiven Theorien auf der Grundlage von Vorschlägen der Forscher.

3.5 *Überprüfung der Subjektiven Theorien auf ihre Handlungswirksamkeit*

Grundsätzliche Überlegungen und Entscheidungen

Die von den Lehrern im Interview geäußerten Vorstellungen und Überzeugungen müssen im Unterricht nicht zwingend handlungswirksam werden. Um festzustellen, in welchem Maße die aus dem Interview rekonstruierten und validierten operativen Handlungsstrukturen und Subjektiven Theorien eines Lehrers im Unterricht handlungsleitend sind, werden in mehreren Schritten Bezüge zwischen den Subjektiven Theorien und konkreten Unterrichtssituationen hergestellt, die zu einer Handlungsvalidierung führen sollen. Von jedem an der Untersuchung teilnehmenden Lehrer werden zwei Unterrichtsstunden vor der Interventionsphase und zwei Unterrichtsstunden nach Abschluss der Intervention mit dem Ziel aufgenommen, mögliche Veränderungen in den Unterrichtsmustern der Probanden als Effekte der Treatments nachzuweisen. Die Unterrichtsaufnahmen erfolgen in Anlehnung an die Aufnahmebedingungen der TIMS-Videostudie innerhalb eines Monats nach den Interviews.

Weitere für das Projekt wichtige Daten (z.B. berufsbiografische Daten, Angaben zur videografierten Unterrichtsstunde usw.) werden mit Fragebögen erhoben.

Die aufgezeichneten Unterrichtsstunden werden dazu genutzt zu überprüfen, welche der im Rahmen der Ausgangsanalyse eruierten Handlungsregeln beim Lehrer handlungswirksam und im Unterricht identifizierbar sind. Da eine Handlungsregel verschied-

dene Verhaltensweisen zulässt, ist es Ziel der Unterrichtsanalyse zu untersuchen, ob die Präferenz für Handlungsregeln, die der Lehrer im Interview genannt hat, in der faktischen Gestaltung des Unterrichts durch den Lehrer nachweisbar ist.

Die Handlungswirksamkeit der Subjektiven Theorien wird anhand der im Unterricht beobachteten Verhaltensweisen, welche über die operative Struktur mit den validierten Handlungsregeln gekoppelt sind, ermittelt. Dazu wird zunächst das im Unterrichtsvideo beobachtbare Verhalten auf Lehrer- und Schülerseite mit Verhaltenstermini beschrieben. Ausgehend von den ermittelten Handlungsregeln und den sich darauf beziehenden Subjektiven Theorien werden die Beziehungen und Interdependenzen zu den im Video sichtbaren Verhaltensweisen des Lehrers und den dazu komplementären Verhaltensweisen der Schüler analysiert.

4. Ausgewählte Ergebnisse anhand einer Fallstudie

Im folgenden Abschnitt werden die verwendeten Erhebungsmethoden an dem Beispiel eines Lehrers X eines Physikgrundkurses der Jahrgangsstufe 12 anwendungsbezogen dargestellt.

4.1 Diagnose Subjektiver Theorien

Entsprechend den fünf ausgewählten Kognitionsbereichen ließen sich in dieser Fallstudie 21 validierte Elemente Subjektiver Theorien (z.B. ein Pfad in Abb. 3) bei dem Lehrer nachweisen. Davon sind 12 Elemente handlungsleitend (z.B. Theorieelemente über die Gestaltung von Wiederholungsphasen oder über den Einsatz von Schülerexperimenten), eines ist nicht in allen Unterrichtssituationen, in denen es anwendbar ist, handlungsleitend. Vier validierte Elemente Subjektiver Theorien widersprechen dem beobachteten Handeln und weitere vier konnten noch nicht geprüft werden, da der beobachtete Unterricht nicht die entsprechenden Bedingungen aufwies. Es spricht für die Erhebungsmethoden, dass von den 17 bisher überprüften Elementen Subjektiver Theorien sich 12 (70%) als handlungswirksam herausgestellt haben. Davon sind drei aus der Sicht von Ergebnissen fachdidaktischer Forschung als bedenklich anzusehen; sie basieren auf der Grundannahme des Probanden, dass effektiver Unterricht immer mit einem Zuwachs an „abrechenbarem Wissen“ (Faktenwissen) verbunden ist, und wurden als Anlass für das Treatment genutzt.

4.2 Authentizität der videografierten Unterrichtsstunden

Knapp zwei Wochen nach dem Interview erfolgte die Aufzeichnung von zwei Unterrichtsstunden in einer 9. Klasse und einem Grundkurs, welche von dem Probanden unterrichtet werden. Die Auswertung der eingesetzten Fragebögen zeigte, dass die Schüler

(n = 45) die aufgezeichneten Physikstunden als einen für den Lehrer typischen Physikunterricht in der entsprechenden Klasse ansehen. So gaben 88% der Schüler des in diesem Fallbeispiel untersuchten Grundkurses 12 an, dass der betreffende Unterricht „wie immer“ bzw. „fast wie immer“ war.

4.3 *Treatments*

Die Struktur von Beratungsgesprächen folgt der Zielsetzung, den Klienten darin zu unterstützen, für ein Problem eine Lösung zu gewinnen, die auf seinen individuellen Voraussetzungen aufbaut. Einer ersten Phase (Orientierungsphase), die der Kontaktaufnahme, dem Aufbau einer positiven Atmosphäre sowie der Vereinbarung über Themen, Rahmenbedingungen und Verfahren des Beratungsprozesses dient, folgt eine Phase, die der Klärung des Problems dient, das der Klient bearbeiten möchte (Klärungsphase). Ihr schließt sich eine Phase an, die auf die Erarbeitung einer Problemlösungsmöglichkeit (unter Einschluss neuer Sicht- und Verhaltensweisen) ausgerichtet ist, deren Ergebnisse in einer Abschlussphase festgehalten und durch Vereinbarungen über die nächsten Schritte des Klienten abgesichert werden (vgl. König/Volmer 2002, S. 15ff.).

Im vorab genannten Fall des Lehrers X wurde zu Beginn des Beratungsgesprächs zunächst Ablauf und Ziel von Coachings erläutert, das Einverständnis über Rahmenbedingungen des Gesprächs (Videoprotokollierung, Teilnahme von zwei Beobachtern u.a.m.) abgesichert, schließlich unter Bezug auf das Interview mit den Probanden ein Problem definiert, das ihn nach eigener Aussage mit dem Unterricht unzufrieden sein lässt. Möglicherweise – so der Lehrer – habe er einen „zu hohen wissenschaftlichen Anspruch im Unterricht“, der dazu führe, dass sowohl die Schüler (nach seiner Beobachtung) öfter mal unzufrieden mit dem Unterricht sind als auch er selbst. Eigentlich würde er in stärkerem Maße gerne die Rolle eines Lernberaters im Unterricht übernehmen, was jedoch – möglicherweise aufgrund seines wissenschaftlichen Anspruchs an den Unterricht – nicht in dem Maße gelinge, wie er sich das wünsche. In der Klärungsphase wurde das Problem mit Blick auf hierfür typische Unterrichtssituationen fokussiert sowie der Stellenwert des Problems für zentrale Gelenkstellen in der Steuerung des Unterrichts beleuchtet. Mit der Verständnisabsicherung des Problems und seiner Konsequenzen für den Verlauf des Unterrichts ging eine schrittweise Klärung von Einstellungen und Überzeugungen einher, die mit der Verlaufstypik des Unterrichts verknüpft sind und im konkreten Fall Bedingungen für den Konflikt sind, der den Lehrer unzufrieden macht. Dabei wurde deutlich, dass mit diesem Konflikt eine Reihe weiterer Probleme verknüpft sind, die aus der Sicht des Lehrers einen optimalen Unterrichtserfolg einschränken (z.B. unmotivierte Schüler). Im Ergebnis führte die ausführliche Klärungsphase dazu,

- dass mit dem „wissenschaftlichen Anspruch des Unterrichts“ seitens des Lehrers ein Konzept effektiven Unterrichts konnotiert ist, demzufolge am Ende jeder Unterrichtsstunde ein möglichst abrufbarer Kanon an physikalischen Wissens-elementen (Definitionen, Gesetze, Berechnungsverfahren) stehen sollte,

- dass dieses Verständnis effektiven Unterrichts die didaktische Strategie des Lehrers bis hinein in einzelne Verhaltensweisen der Interaktion zwischen Lehrer und Schüler bestimmt und in der Didaktik als engmaschige lehrerzentrierte Unterrichtsführung bekannt ist und
- dass dieses Verständnis dem Wunsch des Lehrers entgegensteht, in stärkerem Maße als „Lernberater“ der Schüler zu fungieren.

Für die in der Veränderungsphase anstehende Gewinnung neuer Problemlösungen war damit die Frage leitend, was der Lehrer tun könne, um seiner Lernberater-Rolle besser gerecht zu werden. Spätestens in dieser Phase hat der Coach für sich zu klären, auf welche „Instrumente“ der verschiedenen Ansätze er zurückgreift. In vorliegendem Fall erschien es für das erste Gespräch zweckmäßig, nicht die „beliefs“ hinsichtlich effektiven Unterrichts sowie damit verknüpfte Annahmen über den Lernprozess von Schülern zu bearbeiten, sondern zunächst zu klären, was im Sinne eines ersten Schrittes für den Probanden möglich ist, um dem mit der Lernberaterrolle assoziierten Ziel näher zu kommen.

Es wurden deshalb zunächst bisherige Lösungsversuche, das heißt Handlungsweisen thematisiert, die aus der Sicht des Lehrers mit einer stärkeren Wahrnehmung der Rolle als Lernberater verknüpft sind. Hierbei zeigte sich, dass dazu insbesondere Situationen und Verhaltensweisen gehörten, in denen der Lehrer den Schülern mehr Zeit für ein selbstständiges Nachdenken über Aufgabenstellungen und Problemlösungsvorschläge von Mitschülern gab. Unter Bezug auf die in der Klärungsphase ermittelten zirkulären Strukturen zwischen einer engmaschigen Unterrichtsführung und den Reaktionen der Schüler sowie der Funktion entsprechend gelagerter Interaktionsprozesse für den Erwerb von Kenntnissen und Kompetenzen wurde daraufhin überlegt und erörtert, welche Verhaltensweisen für typische Situationen und Phasen des Unterrichts geeignet sein könnten, um den Schülern mehr Raum für selbstständiges Nachdenken zu geben.

In der Abschlussphase wurde die gefundene Verhaltensweise hinsichtlich ihrer Umsetzbarkeit in den nächsten Unterrichtsstunden in den Blick genommen, eine konkrete Hilfestellung dazu vereinbart und ein Kontrakt über die infragestehende Verhaltensweise mit einem Verstärker belegt. Ebenfalls vereinbart wurde eine Feedback-Schleife hinsichtlich der Erfahrungen des Probanden mit der neuen Verhaltensweise sowie Termin und Ziel des nächsten Beratungsgesprächs.

4.4 Erste treatmentbedingte Veränderungen

In der Ausgangsanalyse wurde als eine zentrale handlungsleitende Grundannahme des Lehrers die Überzeugung ermittelt, dass Unterricht immer in dem Sinne effektiv sein muss, dass die Schüler am Ende einer Unterrichtseinheit über einen vom Lehrer kontrollierbaren Zuwachs an physikalischem (Fakten-)Wissen verfügen müssen. Diese Vorstellung erwies sich insofern als dysfunktional, als sie im vorliegenden Fall dazu führte, dass andere vom Lehrer ebenfalls als zentral erachtete Unterrichtsziele, wie die Förde-

rung von naturwissenschaftlichem Verständnis sowie von Selbstständigkeit und Problemlösefähigkeit damit in unzureichendem Maße berücksichtigt werden konnten.

Bereits während des ersten Coachings gelangte Lehrer X zu der Einsicht, dass Unterricht unter anderen Prämissen nicht weniger effektiv ist, wenn z.B. als Lernziel selbstständiges Problemlösen intendiert ist. Schon nach den ersten zwei Coaching-Gesprächen zeigte sich ein deutliches Bemühen des Lehrers um Veränderungen einzelner Handlungselemente, die auf der Grundannahme zu effektivem Unterricht beruhen. So hat Lehrer X in einer 12. Klassenstufe eine Unterrichtssequenz durchgeführt, in der die Schüler über mehrere Stunden vollkommen eigenständig einen neuen Inhalt anhand vorbereiteter Arbeitsblätter mit Instruktionen und Übungsaufgaben erarbeiteten. Mit diesem erfolgreichen Einsatz einer „neuen Unterrichtsmethode“ akzeptierte Lehrer X bei den Schülern eine selbstständige und problemzentrierte Erarbeitung von physikalischem Faktenwissen. Diese nach dem Treatment gezeigte Konzeption steht im Widerspruch zu der ursprünglich vorhandenen Grundannahme über effektiven Unterricht und deutet so auf die Veränderung subjektiver Theorien hin (z.B.: Wenn man sich für Experimente im Unterricht entscheidet, dann sollte man unter keinen Umständen offene Schülerexperimente im Sinne von „entdeckendem Lernen“ durchführen, um keine Zeit zu vergeuden und einen effektiven Unterricht mit deutlichem Ergebnis zu erreichen). Die am Ende dieser Unterrichtseinheit durchgeführte Leistungsüberprüfung in Form eines schriftlichen Tests zeigte, dass die Leistungen der Schüler sich im Zuge der geänderten didaktischen Arrangements keinesfalls verschlechterten. Dieses Ergebnis überzeugte Lehrer X davon, dass Unterrichtsformen, die den Schüler ein höheres Maß an Selbstständigkeit und Selbsttätigkeit abverlangen, mit dem eigenen Maßstab für „guten“ Unterricht verträglich sind.

5. Ausblick

Im Rahmen der weiteren Forschungsarbeit soll auf der Basis einer größeren Anzahl von Lehrern untersucht werden, in welchem Ausmaß die subjektiven Theorien von Physiklehrern differieren, wie stark sich subjektive Theorien und den Prozessverlauf des Unterrichts steuernde Handlungsmuster voneinander unterscheiden sowie schließlich, in welchem Ausmaß die operativen Strukturen des Unterrichts zwischen den Probanden differieren. Erwartet wird, dass sich anhand entsprechender Vergleiche typologische Strukturen in den Vorstellungs- und Handlungsschemata ermitteln lassen, die in einem weiteren Schritt auf davon abhängige Schülerleistungen hin untersucht werden sollen.

Ziel in Bezug auf die Coaching-Verfahren ist es, im weiteren Untersuchungsverlauf a) die Zielsetzungen für Interventionen zu systematisieren, b) die Formen und Verfahren der Treatments entsprechend den Zielsetzungen für Interventionen als ein Methodeninventar zu beschreiben und zu präzisieren sowie c) die Wirksamkeit verschiedener Treatments in Hinblick auf einzelne Ziele der Referenztransformation (Veränderung des Deutungs- und Interpretationsmusters) und der Veränderung didaktischer Handlungsschemata zu überprüfen.

Grundsätzlich erscheint es möglich, ein Modell der Entwicklung und Erweiterung didaktischer Kompetenz von (Physik-)Lehrern zu erarbeiten und zu testen.

Literatur

- Bandler, R./Grinder, J. (1992): Reframing. Paderborn: Junfermann.
- Bauer, R./Kaiser, H./Müsgens, R. (1979): Zur Methodik des Physikunterrichts in der gymnasialen Oberstufe. In: *physica didactica* 6, S. 115–124.
- Bleichroth, W./Dahncke, H./Jung, W./Kuhn, W./Merzyn, G./Weltner, K. (1999): *Fachdidaktik Physik*. Köln: Aulis.
- BLK (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung) (1997): Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“. Materialien zur Bildungsplanung und zur Forschungsförderung, Heft 60. Bonn.
- Böhm-Kasper, O./Bos, W./Körner, S.C./Weishaupt, H. (2001): Sind 12 Schuljahre stressiger? Belastung und Beanspruchung von Lehrern und Schülern am Gymnasium. Weinheim: Juventa.
- Boulton-Lewis, G.M./Smith, D.J.H./McCordle, A.R./Burnett, P.C./Campbell, K.J. (2001): Secondary teachers' conceptions of teaching and learning. *Learning and Instruction* 11, S. 35–51.
- Clark, C.M./Peterson, P.L. (1986): Teachers' Thought Processes. In: Wittrock, M.C. (Hrsg.): *Handbook of Research on Teaching*. New York: McMillan, S. 255–296.
- Combe, A./Buchen, S. (1996): Belastung von Lehrerinnen und Lehrern: Fallstudien zur Bedeutung alltäglicher Handlungsabläufe an unterschiedlichen Schulformen. Weinheim: Juventa.
- Dryden, W. (1995): *Brief Rational Emotive Behaviour Therapy*. Chichester: Wiley.
- Duit, R./Treagust, D. (1998): Learning in Science. From behaviourism towards social constructivism and beyond. In: Fraser, B./Tobin, K. (Hrsg.): *International Handbook of Science Education*. Dordrecht: Kluwer, S. 3–25.
- Ellis, A. (1993): *Die Rational-Emotive Therapie*. München: Pfeiffer.
- Fensham, P.J. (1987): Theory in Practice: How to Assist Science Teachers to Teach Constructively. In: Adey, P. (Hrsg.): *Adolescent Development and School Science*. New York u.a.: Falmer, S. 61–77.
- Fischler, H. (2001): Verfahren zur Erfassung von Lehrer-Vorstellungen zum Lehren und Lernen in den Naturwissenschaften. In: *Zeitschrift für Didaktik der Naturwissenschaften* 7, S. 163–178.
- Gustafson, B.J./Rowell, P.M. (1995): Elementary preservice teachers: constructing conceptions about learning science, teaching science and the nature of science. In: *International Journal of Science Education* 17, S. 589–605.
- Hage, K./Bischoff, H./Dichanz, H./Eubel, K.-D./Oehlschläger, H.-J./Schwittmann, D. (1985): *Das Methoden-Repertoire von Lehrern. Eine Untersuchung zum Unterrichtalltag in der Sekundarstufe I*. Opladen: Leske + Budrich.
- Hand, B./Treagust, D. (1994): Teachers' Thoughts about Changing to Constructivist Teaching/Learning Approaches within Junior Secondary Science Classrooms. In: *Journal of Education for Teaching* 20(1), S. 97–112.
- Häußler, P./Bünder, W./Duit, R./Gräber, W./Mayer, J. (1998): *Naturwissenschaftsdidaktische Forschung – Perspektiven für die Unterrichtspraxis*. Kiel: IPN.
- Hewson, P.W./Hewson, M.G.A'B. (1987): Science teachers' conceptions of teaching: Implications for teacher education. In: *International Journal of Science Education* 9, S. 425–440.
- Kircher, E./Girwidz, R./Häußler, P. (2000): *Physikdidaktik. Eine Einführung in Theorie und Praxis*. Braunschweig/Wiesbaden: Vieweg.
- König, E./Volmer, G. (1996): *Systemische Organisationsberatung*. Weinheim: Dt. Studien Verlag.
- König, E./Volmer, G. (2002): *Systemisches Coaching. Handbuch für Führungskräfte, Berater und Trainer*. Weinheim: Beltz.

- Looss, W. (1991): *Coaching für Manager: Problembewältigung unter vier Augen*. Landsberg/Lech: Verlag Moderne Industrie.
- Lückert, H.R./Lückert, I. (1994): *Einführung in die kognitive Verhaltenstherapie*. München/Basel: Reinhardt.
- Marion, R./Hewson, P./Tabachnick, B./Blomker, K. (1994): *Teaching for Conceptual Change in Elementary and Secondary Science Methods Courses*. Paper Presented at the Annual Meeting of the National Association for Research in Science Teaching, New Orleans.
- Meichenbaum, D. (1979): *Kognitive Verhaltensmodifikation*. München: Urban & Schwarzenberg.
- Pajares, M.F. (1992): Teachers' beliefs and educational research: cleaning up a messy construct. In: *Review of Educational Research* 62, S. 307–332.
- Parsons-Chatman, S. (1990): *Making Sense of Constructivism in Preservice: A Case Study*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta.
- Schreyögg, A. (1999): *Coaching. Eine Einführung für Praxis und Ausbildung*. Frankfurt/New York: Campus.
- Spanhel, D./Hübner, H.-G. (1995): *Lehrersein heute: berufliche Belastungen und Wege zu deren Bewältigung*. Bad Heilbrunn: Klinkhardt.
- Tillema, H.H. (1994): *Training and Professional Expertise: Bridging the Gap between New Information and Pre-Existing Beliefs of Teachers*. *Teaching and Teacher Education* 10, S. 601–615.
- Tillema, H.H. (2000): *Belief change towards self-directed learning in student teachers: Immersion in practise or reflection on action*. In: *Teaching and Teacher Education* 16, S. 575–591.
- Tillema, H.H./Knol, W.E. (1997): *Promoting Student Teacher Learning Through Conceptual Change Or Direct Instruction*. In: *Teaching and Teacher Education* 13, S. 579–595.
- Thonhauser, J. (1987): *Lehrerbildung als Weiterentwicklung Subjektiver Theorien*. In: Schlee, J./Wahl, D. (Hrsg.): *Veränderung Subjektiver Theorien von Lehrern*. Oldenburg: Universität Oldenburg, S. 236–252.
- Wahl, D. (1991): *Handeln unter Druck. Der weite Weg vom Wissen zum Handeln bei Lehrern, Hochschullehrern und Erwachsenenbildern*. Weinheim: Dt. Studien Verlag.
- Watzlawick, P./Beavin, J.H./Jackson, D.D. (1972): *Menschliche Kommunikation: Formen, Störungen, Paradoxien*. Bern: Huber.
- Watzlawick, P. (Hrsg.) (1997): *Die erfundene Wirklichkeit*. München: Pieper.
- Yerrick, R./Parke, H./Nugent, J. (1997): *Struggling to Promote Deeply Rooted Change: The „Filtering Effect“ of Teachers' Beliefs on Understanding Transformational Views of Teaching Science*. In: *Science Education* 81, S. 137–159.

Anschrift der Autoren:

- Prof. Dr. Helmut Fischler, Freie Universität Berlin, Fachbereich Physik, Arnimallee 14, 14195 Berlin.
- Prof. Dr. Peter Zedler, Universität Erfurt, Institut für Allgemeine Erziehungswissenschaft und Empirische Bildungsforschung, Nordhäuser Str. 63, 99089 Erfurt.
- Hans-Joachim Schröder, Freie Universität Berlin, Fachbereich Physik, Arnimallee 14, 14195 Berlin.
- Cornelia Tönhäuser, Universität Erfurt, Institut für Allgemeine Erziehungswissenschaft und Empirische Bildungsforschung, Nordhäuser Str. 63, 99089 Erfurt.

Teil III: Entwicklung und Evaluation von Unterrichtsmodulen und Trainingsprogrammen

Schulische Lehr-Lernumgebungen und außerschulische Trainings zur Förderung fächerübergreifender Kompetenzen von Schülerinnen und Schülern

Bernhard Schmitz

Einleitung

In der Verarbeitung der Ergebnisse von TIMSS und PISA lassen sich zwei potenzielle Strategien als Extrempositionen ausmachen. Die eine Variante (vor der Schavan 2001 nachdrücklich warnt) tendiert zu schnellen und leichtfertigen Schlüssen und identifiziert simplifizierend einzelne Ursachen für das Abschneiden der deutschen Schülerinnen und Schüler. Die andere Position könnte darin bestehen, aufgrund der Komplexität und Vielgestalt möglicher Schlussfolgerungen – wie der Buridansche Esel vor seiner Futteralternative – in Erstarrung und völlige Entscheidungsunfähigkeit zu verfallen.

Die in diesem Teil zusammengefassten sechs Beiträge sind darauf angelegt, jeweils einen Ansatzpunkt zur Optimierung der Erziehungsanstrengungen zu benennen und die Wirksamkeit der vorgeschlagenen Maßnahmen empirisch zu belegen.

Alle hier vorzustellenden Beiträge führen Trainings- oder (Quasi-)Experimente durch bzw. stellen Untersuchungen dar, die solche Interventionen vorbereiten sollen. Es werden somit gezielte und systematische Versuche unternommen, nicht nur Zusammenhänge zu beschreiben und zu analysieren, sondern auch Unterrichtsgeschehen oder dessen Kontext zu beeinflussen. Dies geschieht unter kontrollierten Bedingungen und erlaubt wissenschaftlich weitreichende Schlussfolgerungen, insofern bei den hier präferierten experimentellen Ansätzen fundiertere (im Vergleich zu deskriptiven oder korrelativen Vorgehensweisen) kausale Schlussfolgerungen in Bezug auf die manipulierte unabhängige Variable gezogen werden können. Pekrun (2002) weist nachdrücklich daraufhin, dass weder TIMSS noch PISA aufgrund des primär deskriptiven und korrelativen Ansatzes Kausalerklärungen zulassen. Die hier berichteten Interventionsstudien versuchen somit eine empirische Basis für die Umsetzung von Konsequenzen aus TIMSS/PISA zu liefern.

Neben der methodischen Zugangsweise der Intervention bzw. der experimentellen Studie, sind die hier vorgestellten Beiträge durch eine weitere Besonderheit gekennzeichnet, nämlich die Art der zu analysierenden Kompetenz. Es werden über die häufig vermittelten Fertigkeiten weitergehende fachspezifische und/oder fachübergreifende Kompetenzen in den Blick genommen.

Die betrachteten Kompetenzen sind: ein Wissenschaftsverständnis (Sodian/Thoerner/Kircher/Grygier/Günther), ein tieferes naturwissenschaftliches Verständnis bereits in der Grundschule (Möller/Jonen/Hardy/Stern), das Lernen aus Beispielen und Selbsterklärungen (Renkl/Schworm), Selbstregulation (Leopold/Leutner sowie Gürtler/Perels/Schmitz/Bruder), selbstorganisiertes Lernen und selbstbestimmte Motivation (Sumfleth/Wild/Rumann/Exeler).

Im Folgenden werden kurz die Ziele und Besonderheiten der einzelnen Studien genannt: Der Artikel von Möller/Jonen/Hardy/Stern beschäftigt sich mit der Förderung des naturwissenschaftlichen Verständnisses durch den Sachunterricht in der Grund-

schule, um so schon frühzeitig ein Lernen im Vorfeld der Physik zu ermöglichen. Dabei haben die Autoren sich zum Ziel gesetzt, ein genaueres Verständnis über das naturwissenschaftsbezogene Denken und Lernen von Grundschulern zu erwerben und Möglichkeiten einer angemessenen und wirkungsvollen Förderung aufzuzeigen.

Die Studie von Sodian/Thoermer/Kircher/Grygier/Günther untersucht Vorstellungen über die Grundlagen des Experimentierens und das Wissenschaftsverständnis bei Grundschulern. Dabei wird in einem ersten Schritt das intuitive Wissenschaftsverständnis von Schülern der dritten und vierten Jahrgangsstufe beschrieben und durch kurze Interventionen überprüft, in welchem Ausmaß die epistemologischen Präkonzepte durch die Anregung wissenschafts- bzw. erkenntnistheoretischer Reflexion anregbar sind.

Sumfleth/Wild/Rumann/Exeler zeigen in ihrer Arbeit Wege zur Förderung der naturwissenschaftlichen Grundbildung im Chemieunterricht auf. Dabei werden die Wirkungen und Wechselwirkungen von Familienmerkmalen (Form der Unterstützung bei der Hausaufgabenbearbeitung) als außerschulischem Kontext und Instruktionsmerkmalen (kooperatives Lernen) als schulischem Kontext analysiert. Als abhängige Variablen werden dabei fachspezifische und fächerübergreifende Aspekte untersucht.

Zur Behebung von Defiziten in der Bearbeitung mathematischer Aufgaben soll in der Studie von Gürtler/Perels/Schmitz/Bruder überprüft werden, ob sich fachspezifische (Problemlösen in Mathematik) und fachübergreifende Kompetenzen (Selbstregulation) bei Gymnasiasten der achten Jahrgangsstufe durch Trainings verbessern lassen. Dabei soll weiterhin das Lernverhalten bei Hausaufgaben als wichtigem außerschulischem Kontext mit standardisierten Tagebüchern durch verstärkte Selbstreflexion gefördert werden.

Leopold/Leutner untersuchen die Verwendung selbstregulatorischer Lernstrategien in unterschiedlichen Jahrgangsstufen. Dabei legen sie ihren Schwerpunkt zunächst auf die Analyse der Veränderung des Lernstrategieinsatzes im Zeitverlauf und die Beziehungen zwischen Lernerfolg und Strategieinsatz.

In der Studie von Renkl/Schworm wird eine computerbasierte Lernumgebung für Lehrende entwickelt, mit der erlernt werden kann, wie im Unterricht effektives Lehren mit Lösungsbeispielen realisiert werden kann. Sie untersuchen dabei, ob die Bereitstellung unterschiedlicher Lernhilfen die angehenden Lehrer dabei unterstützt, sich entsprechendes Wissen anzueignen.

Literatur

- Schavan, A. (2001) Vorwort der Präsidentin der Kultusministerkonferenz. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000, Leske und Budrich: Opladen, S. 11–12.
- Pekrun, R. (2002) Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für zukünftige Bildungsforschung. Zeitschrift für Pädagogik, 48, S. 111–128.

Anschrift des Autors:

Bernhard Schmitz, Technische Universität Darmstadt, Fachbereich 3, Institut für Psychologie, Hochschulstraße 1, 64289 Darmstadt.

Kornelia Möller/Angela Jonen/Ilonca Hardy/Elsbeth Stern

Die Förderung von naturwissenschaftlichem Verständnis bei Grundschulkindern durch Strukturierung der Lernumgebung¹

1. Einleitung

1.1 Naturwissenschaftliches Lernen im Grundschulalter?

Im Zuge von TIMSS und PISA gerät auch die Grundschule in den Blick. Fördert die Grundschule angemessen die Begabungen unserer Kinder? Könnte eine Förderung schon in der Grundschule späteres naturwissenschaftliches Lernen erleichtern und einem Interessenverfall im Bereich der Naturwissenschaften entgegenwirken?

Manche Pädagogen stehen diesen Fragen kritisch gegenüber, gilt es doch, Bildungsprozesse in der Grundschule vor einer Vereinnahmung durch die weiterführenden Schulen zu schützen. Grundschule sei auf die gegenwärtigen Bedürfnisse und Interessen von Grundschulkindern auszurichten; Sachunterricht habe primär die Aufgabe, zur Klärung der von Kindern erlebten Welt und zu einer in dieser Welt erforderlichen Handlungsfähigkeit beizutragen.

Bereits in den 70er-Jahren forderten Bildungspolitiker und Pädagogen eine stärkere Wissenschaftsorientierung der Grundschule. Auf der Basis neuer lernpsychologischer Erkenntnisse erfolgte die Aufnahme naturwissenschaftlicher Themen in den Kanon des von da an als „Sachunterricht“ bezeichneten Faches. Schon wenige Jahre nach Beginn der Reform setzte allerdings heftige Kritik ein: Die neuen, durch Entwicklungen im Ausland beeinflussten Curricula vernachlässigten kindliche Sichtweisen und Interessen, überforderten die Kinder häufig im Verstehen der Inhalte und trafen zudem auf enorme Implementationsschwierigkeiten (Beck/Claussen 1979; Möller 2001a). Die Folge: Offene, verfahrensorientierte und lebensweltlich akzentuierte Konzeptionen verdrängten naturwissenschaftsbezogene Themen seit Mitte der 70er-Jahre und führten zu einer Vernachlässigung des naturwissenschaftlichen Bereiches, was sich in der geringen Bedeutung entsprechender Themen in Lehrplänen und in der Aus- und Fortbildung von Sachunterrichtslehrkräften widerspiegelt (Strunck/Lück/Demuth 1998).

Seit einiger Zeit, sicher auch als Folge der Ergebnisse der aktuellen Schulleistungsvergleiche und unterstützt durch neuere entwicklungspsychologische Befunde (siehe 1.2), findet die Forderung nach einem frühen Lernen im Vorfeld der Naturwissenschaften wieder verstärkte Aufmerksamkeit. Dabei stellt sich die Frage, wie Unterricht orga-

1 Die berichteten Studien wurden im Rahmen des DFG-Schwerpunktprogramms „Bildungsqualität von Schule (BIQUA)“, Az. MO 942/1-1 (Münster) Az. STE 539/9-1 (Berlin) gefördert.

nisiert sein sollte, um motivierendes naturwissenschaftsbezogenes Lernen zu ermöglichen, ohne Grundschul Kinder zu überfordern. In der fachdidaktischen Diskussion kristallisiert sich in der Tradition Martin Wagenscheins ein sog. konstruktiv-genetischer Unterricht als vielversprechend heraus (Thiel 1990; Köhnlein 1999), der auch unter dem Begriff eines konstruktivistisch orientierten, auf verstehendes, kooperatives und problemorientiertes Lernen ausgerichteten Unterrichts diskutiert wird (z.B. Duit/Treagust 1998; Möller 2001b; Gerstenmaier/Mandl 1995; Reinmann-Rothmeier/Mandl 1997). Kerngedanke dieser Ansätze ist die Annahme, dass Wissen nicht ohne kognitive Aktivität der Lernenden erworben werden kann, woraus sich eine notwendige Selbsttätigkeit der Lernenden bei der Konstruktion von Wissen ableiten lässt. Das hohe Maß an Selbststeuerung und Komplexität, das sich hieraus bei anwendungsorientierten Fragestellungen ergibt, birgt für leistungsschwächere Schüler bei anspruchsvollen Inhalten allerdings die Gefahr der Überforderung. Dennoch zeigt sich im Bereich der Grundschulmathematik, dass auch schwächere Kinder von anspruchsvollem Unterricht profitieren können (Staub/Stern 2002; Renkl/Stern 1994).

Für die Lehrkraft stellt sich die Aufgabe, Lernaktivitäten angemessen zu unterstützen. Die Frage nach dem Ausmaß der Strukturierung in einem auf kognitive Selbststeuerung ausgerichteten Unterricht erweist sich dabei als zentral (Möller *eing.*).

1.2 *Naturwissenschaftliches Lernen aus entwicklungs- und lernpsychologischer Sicht*

Eine weitverbreitete, auf die Rezeption des Werkes von Jean Piaget zurückgehende Überzeugung ist, dass sich das Denken von Grundschulkindern auf konkrete Bereiche beschränkt. Nach Piaget vollzieht sich die geistige Entwicklung des Kindes vom Konkreten zum Abstrakten, indem sich inhaltsübergreifende Denkschemata verändern, wobei die konkret operative Phase des Grundschulkindes durch die Unfähigkeit gekennzeichnet ist, nicht explizit verfügbare Informationen zu erschließen. Die Annahme einer bereichsübergreifenden Veränderung des Denkens und der Informationsverarbeitung gilt jedoch als überholt. Die Befunde zur Inhaltsspezifität der geistigen Entwicklung sind überwältigend (Sodian 1995; Stern 2002). Manche Kinder können in einem Inhaltsgebiet – z.B. in Mathematik – bereits sehr fortgeschrittene geistige Operationen vollziehen, während sie in anderen Inhaltsgebieten bei vergleichbar schwierigen Aufgaben überfordert sind. Generell zeigte sich, dass das Erklärungspotenzial von inhaltsunspezifischen Fähigkeiten wie der allgemeinen Intelligenz geringer ist als das Erklärungspotenzial von inhaltspezifischem Wissen (Stern 2001).

Versteht man kognitive Entwicklung als den Erwerb und die Umstrukturierung von Wissen, ist davon auszugehen, dass eine frühe Konfrontation mit anspruchsvollen Inhaltsgebieten einen Entwicklungsvorsprung verschafft. Dies belegen u.a. Untersuchungen zur Grundschulmathematik (Stern *im Druck*). Unter Wissenserwerb wird hier nicht allein die Akkumulation von Fakten, sondern vielmehr der Erwerb theoriegeleiteter Begriffe und Konzepte verstanden (Carey 1985).

Lernen wird dabei insbesondere in der naturwissenschaftsdidaktischen Forschung als Konzeptwechsel beschrieben (Duit 1999; Duschl/Hamilton 1998). Konzeptwechsel sind das Ergebnis von aktiven Umstrukturierungsprozessen (Vosniadou/Ioannides/Dimitrakopoulou 2001). Sie sind abhängig von den Vorstellungen und Erklärungen, die Schüler in eine Lernsituation einbringen, wie auch von den Eigenschaften der Lernumgebung. Naturwissenschaftliche Phänomene regen Kinder oft spontan zur Bildung von Erklärungen an; viele dieser häufig robusten Vorstellungen sind aber inadäquat oder unvollständig. Ziel einer am konstruktivistischen Lernbegriff orientierten Lehr-Lernumgebung ist es, Kinder auf inadäquate, nicht belastbare Vorstellungen aufmerksam zu machen und ihnen Möglichkeiten zur Konstruktion adäquaterer, qualitativer physikalischer Erklärungen zu bieten (Möller 1999).

In unserem Forschungsprojekt geht es um die Untersuchung einer angemessenen und wirkungsvollen Förderung naturwissenschaftlichen Lernens von Grundschulkindern. Wir konzentrieren uns in unserer Studie auf die Untersuchung konzeptuellen Wissens; das ebenfalls in BIQUA angesiedelte Projekt von Sodian und Kircher untersucht die Förderung von Wissenschaftsverständnis durch unterrichtliche Interventionen in der Grundschule (vgl. Sodian u.a. in diesem Heft).

2. Strukturierung als Mittel der Optimierung in konstruktivistisch orientierten Lehr-Lernumgebungen

2.1 Konzeptwechsel beim Themenbereich „Schwimmen und Sinken“

Ein in Bezug auf Grundschulkindern kontrovers diskutierter Bereich ist das Themengebiet „Schwimmen und Sinken“ mit den zugrunde liegenden Begriffen „Dichte“ und „Auftriebskraft“ (Smith/Carey/Wiser 1985; Thiel 1990; Janke 1995; Möller 1999; Haru 2000). Nicht adäquate Vorstellungen, wie z.B. die Abhängigkeit der Auftriebskraft von der nach oben „ziehenden“ Luft („die Luft will nach oben“) oder die Abhängigkeit des Schwimmens von der Größe („alles, was groß ist schwimmt“) oder der Masse („alles, was leicht ist, schwimmt“) eines Körpers müssen als nicht belastbar erkannt werden, um ein qualitatives Verständnis von Dichte und Auftriebskraft aufbauen zu können. Besondere Probleme bereitet das für den Dichtebegriff erforderliche proportionale Verständnis. Da Grundschulkindern dazu tendieren, nur eine der beiden relevanten Größen Masse und Volumen in ihren Erklärungen zu berücksichtigen, ist diese Art des konzeptuellen Verständnisses besonders schwierig (Smith u.a. 1997; Möller 1999).

2.2 Optimierung durch Sequenzierung und kognitiv strukturierende Gesprächsführung (Schulstudie)

Aus Forschungen zu unterrichtlichen Bedingungen für Konzeptwechsel (z.B. Scott/Driver 1998; Hewson/Beeth/Thorley 1998; Duschl/Hamilton 1998) ergeben sich Unterrichtsmerkmale wie Erfahrungsorientierung, Selbststeuerung, Anwendungsorientierung

und die Berücksichtigung multipler Perspektiven. Eine zu große Komplexität von Lernsituationen wie auch ein zu großes Ausmaß an Selbststeuerung allerdings kann zu Problemen führen (vgl. Weinert 1982; Weinert 1996; Stark/Gruber/Mandl 1998; Gräsel/Mandl 1993; Renkl 1995). Im Grundschulbereich wird ein häufig geringes Niveau der kognitiven Verarbeitung in stark selbstgesteuerten und wenig strukturierten Lernumgebungen kritisiert (Einsiedler 1978, 1996). Auch die selbstgesteuerte Fokussierung relevanter Merkmale bereitet insbesondere jüngeren Lernenden mit geringem, wenig strukturiertem Vorwissen in einem Inhaltsbereich, für den nur wenige Lernstrategien bereit stehen, Schwierigkeiten. Zu vermuten ist daher, dass ein auf Konzeptwechsel ausgerichteter Unterricht einer Verknüpfung von Instruktion und Konstruktion auf der methodischen Ebene der Unterrichtsgestaltung bedarf, um die erforderliche Balance von Selbststeuerung und Fremdsteuerung sicherzustellen (Friedrich/Mandl 1997; Reinmann-Rothmeier/Mandl 1999; Dubs 1997; Möller 2001b).

Der Schulstudie liegt eine problemorientierte Lernumgebung mit einem hohen Maß an angestrebter kognitiver Eigenaktivität zugrunde, in der die Schüler Konzepte entwickeln, kooperativ überprüfen und kommunikativ aushandeln. Gezielte Strukturierungselemente werden in diesen Prozess eingebunden, um die kognitiven Aktivitäten der Schüler durch Komplexitätsreduktion zu unterstützen. Sie beziehen sich auf eine inhaltliche Sequenzierung einer anwendungsorientierten Fragestellung und auf eine kognitiv strukturierende Gesprächsführung, unter der wir das Aufmerksammachen auf Widersprüche im Denken, das Einfordern von Begründungen und Überprüfungen, das Anregen von Anwendungen, das Fokussieren einzelner Aspekte wie auch das Anregen der Integration von Konzepten verstehen.

2.3 Optimierung durch externe Repräsentation (Laborstudie)

Eine weitere Art der Strukturierung stellen externe Repräsentationsformen dar. Diese bieten die Möglichkeit zum Aufbau konzeptuellen Verständnisses, indem einerseits an bereits bestehendes (oft intuitives) Wissen von Lernenden angeknüpft und andererseits die Beachtung relevanter Merkmale unterstützt wird (siehe auch das BIQUA-Projekt Martignon u.a. in diesem Heft). Im Bereich des proportionalen Denkens erlauben Repräsentationsformen, wie z.B. Koordinatensysteme, die Beziehung zwischen zwei Größen als Steigung eines Graphen zu visualisieren, wobei jede der beiden Größen repräsentiert, also auch vom Lerner beachtet werden muss. Die Visualisierung von Strukturen unterstützt den Lernenden bei der bewussten Wahrnehmung relevanter Situationseigenschaften, die zuvor entweder noch nicht beachtet oder nur implizit in den Lösungsprozess integriert wurden (Cox 1999). Das Ausmaß, in dem Repräsentationsformen bestimmte Eigenschaften einer Situation hervorheben, beeinflusst wiederum die Integration dieser Merkmale in ein bereits bestehendes Situationsverständnis und damit die Entwicklung eines tieferen konzeptuellen Verständnisses (Greeno/Smith/Moore 1993; Stern/Hardy/Koerber 2002). Bei der Entwicklung eines konzeptuellen Verständnisses von Schwimmen und Sinken sollte es durch die Nutzung geeigneter Repräsentationsformen möglich

sein, die für das Konzept der Dichte relevanten Situationsmerkmale Masse und Volumen hervorzuheben. Entscheidend ist hierbei, dass die Nutzung von Repräsentationsformen ein konstruktivistisch orientiertes Unterrichten im oben dargelegten Sinne erlaubt. Das bedeutet, dass die Nutzungsweise einer Repräsentationsform möglichst intuitiv verständlich sein sollte, so dass Lernende durch Anknüpfung an ihr Vorwissen ein integriertes Verständnis von Repräsentationsform und repräsentiertem Inhalt aufbauen können. In einer Studie von Koerber (2000) zeigte sich, dass insbesondere die Balkenwaage aufgrund ihrer leichten Interpretierbarkeit von Grundschulkindern erfolgreich für proportionale Vergleiche genutzt werden konnte. In einem Unterricht, der auf die Erklärung von Schwimmen und Sinken ausgerichtet ist, bieten Repräsentationsformen einerseits die Möglichkeit, die Unangemessenheit von Fehlvorstellungen zu erfahren, beispielsweise indem erkenntlich ist, dass die Größe von Gegenständen zusätzlich zu ihrem Gewicht beachtet werden muss. Andererseits wird der Aufbau qualitativer Erklärungen unterstützt, indem das Verhalten von Gegenständen im Wasser durch den Dichtevergleich korrekt vorhergesagt werden kann.

3. Empirische Untersuchungen

In zwei Untersuchungen (Schulstudie und Laborstudie) wurde jeweils an unabhängigen Gruppen die Strukturierung der Lernumgebung variiert. Die Effekte dieser Variationen auf das physikalische Verständnis wurden in Abhängigkeit von den kognitiven Eingangsvoraussetzungen untersucht. Wir erwarten, dass insbesondere Schüler mit ungünstigen Eingangsvoraussetzungen von einem höheren Maß an Strukturierung profitieren.

3.1 Schulstudie

Methode

In einem Messwiederholungsdesign mit zwei Experimentalgruppen und einer nicht trainierten Basisgruppe wurde untersucht, welchen Einfluss Strukturierung in konstruktivistisch orientierten Lehr-Lernumgebungen auf den Erwerb eines Basisverständnisses der physikalischen Konzepte „Dichte“ und „Auftrieb“ in Abhängigkeit von kognitiven Eingangsunterschieden hat.

Unterricht MIT und OHNE Strukturierung: Es wurden zwei konstruktivistisch orientierte Unterrichtseinheiten zum Thema „Wie kommt es, dass ein Schiff schwimmt?“ entwickelt. In der weniger stark strukturierten Einheit (kurz: OHNE) wurde den Kindern eine Lernumgebung in Form eines offenen Materialangebots (Knete, Wasserbecken, Töpfe, Metallboote, Waagen, Holzknöpfe, Metallplatten, Stecknadeln, Kerzen etc.) zur Verfügung gestellt, in der sie ihre Ideen überprüfen und ihre Fragen beantworten konnten. Außerdem gab es ein komplexes Stationenangebot mit Anregungen für Versuche, um

neue Erfahrungen zu ermöglichen und auf Phänomene aufmerksam zu machen. Während der Stationenarbeit gab die Lehrkraft individuelle prozessbezogene Hilfestellungen. Entdeckungen und Probleme wurden in Kleingruppen und in Klassengesprächen präsentiert und diskutiert. Die Kinder bestimmten hierbei weitgehend den Ablauf der Gespräche. In seinem Aufbau ähnelte der Unterricht einem Werkstattangebot. Beide Unterrichtseinheiten erstreckten sich über acht Doppelstunden; eine Doppelstunde fand im Schwimmbad statt.

In der stärker strukturierten Unterrichtseinheit (kurz: MIT) wurde durch eine Sequenzierung nach Teilfragen die Komplexität des Themas verringert. Die Teilfragen (Was schwimmt, was sinkt? Wie kommt es, dass Wachs schwimmt und Metall untergeht? Was passiert mit dem Wasser, wenn man einen Gegenstand hineintaucht? Was macht das Wasser mit den Gegenständen?) wurden, eingebettet in die komplexe Fragestellung: „Wie kommt es, dass ein Schiff schwimmt?“, nacheinander erforscht; die in Stationen- oder Gruppenarbeit angebotenen Materialien, Aufgaben und Versuche beschränkten sich hierbei auf die jeweiligen Teilfragen. Zudem setzte die Lehrkraft in den häufigeren und längeren Klassengesprächen verstärkt kognitiv strukturierende Gesprächshilfen ein, um den aktiven Konzeptaufbau, insbesondere auch Konzeptwechsel, zu fördern. Sie hob z.B. widersprüchliche oder wichtige Aussagen hervor, forderte Begründungen, Überprüfungen und Anwendungen ein, verdeutlichte Erklärungen, gab prozessbezogene Hilfen und wiederholte wichtige Zusammenfassungen.

Teilnehmende Klassen: Sechs dritte Klassen aus drei vergleichbaren Schulen mit insgesamt 149 Kindern (65 Mädchen, 84 Jungen) dienten als Experimentalgruppen, zwei dritte Klassen aus zwei weiteren Schulen mit 41 Kindern (27 Mädchen, 14 Jungen) dienten als Vergleichsgruppe ohne Unterricht (Basisgruppe). Die Durchführung des Unterrichts erfolgte in drei Wellen von je vier Wochen. Pro Welle wurde in je einer Schule Unterricht MIT und Unterricht OHNE in zwei Parallelklassen je zwei Wochen nacheinander unterrichtet. Die Zuordnung MIT/OHNE erfolgte zufällig. In den Basisklassen wurde ausschließlich die Testbatterie durchgeführt. Der gesamte Unterricht in allen sechs Untersuchungsklassen wurde videografiert.

Screening des Unterrichts: Zur Überprüfung der Einhaltung der Unterschiede zwischen den beiden Unterrichtsdesigns wurden zwei Screeningverfahren angewendet. Die Videoaufnahmen der Unterrichtsstunden wurden von 12 unabhängigen Beobachtern in einem Blind-Verfahren anhand der Kriterien für den Unterricht MIT bzw. OHNE bei 45 von 48 Doppelstunden den Bedingungen richtig zugeordnet.

Um die Variationskriterien Sequenzierung und kognitiv strukturierende Gesprächsführung nachzuweisen, wurde ein Kategoriensystem entwickelt, mit dem 30% der Klassengespräche in einem 10-Sekunden-Takt analysiert wurden; bei der zufälligen Auswahl wurden alle Stunden der Einheiten wie auch alle beteiligten Klassen gleichermaßen berücksichtigt. Die Auswertung der Äußerungen weist auf einen signifikant höheren Anteil von strukturierenden Lehreräußerungen im Unterricht MIT hin. Die Ergebnisse der Screenings lassen auf eine Einhaltung der Gruppenunterschiede schließen.

Erhebungsinstrumente

Test zum Schwimmen und Sinken: Der als Prä-Post-Test eingesetzte, auf der Basis von Pilotstudien entwickelte Test umfasst insgesamt 17 Aufgaben mit 14 Multiple-Choice-Items und drei Items mit offenem Antwortformat. Die Fragen beziehen sich auf typische Präkonzepte von Drittklässlern (nicht belastbare Konzepte wie Gewichtskonzept, Größenkonzept, Formkonzept, Luftkonzept), auf belastbare Konzepte (wie Materialkonzept, Hohlkörperkonzept) und auf qualitative physikalische Konzepte (Verdrängung, Dichtevergleich, Auftrieb). Signalwörter wie „leichter/schwerer als“ und „drücken“, die häufig im Unterricht verwendet werden, wurden jeweils innerhalb eines Items in einer richtigen und in einer falschen Wendung angeboten. Es ergab sich eine interne Konsistenz von $\alpha = .73$ bei Berechnung von Aufgabensummenwerten, die sowohl die korrekte Ablehnung eines nicht belastbaren Konzeptes (z.B. Luftkonzept, Gewichtskonzept) bepunkteten wie auch die korrekte Annahme von Antworten auf qualitativem Verständnisniveau (Materialkonzept, Hohlkörperkonzept) und explizitem Verständnisniveau (qualitative physikalische Erklärungen mit Dichte und Auftrieb).

Transfertest zum Schwimmen und Sinken: Die Fragen dieses Tests beziehen sich auf die Anwendung des Dichtekonzepts und des Auftriebskonzepts in neuen, nicht im Unterricht behandelten Kontexten anhand von drei Items mit offenem Antwortformat und 12 Multiple-Choice Items ($\alpha = .60$ bei der Auswertung mit Aufgabensummenwerten).

Kriterienkatalog zur Schülereinschätzung: Die Klassenlehrerinnen schätzten die Kinder vor der Studie in Bezug auf insgesamt 13 Merkmale ein. Auf einer 16-stufigen Notenskala wurde für jedes Kind die Fähigkeit zum „kreativen, selbstständigen und problemlösenden Denken“ erfasst.

Ergebnisse

Test zum Schwimmen und Sinken: Zur Erfassung des Lerngewinns in den Unterrichtsformen MIT und OHNE wurden 2 (Zeit) X 2 (Gruppe) Messwiederholungsanalysen mit den Summenwerten des Tests zum Schwimmen und Sinken durchgeführt. Der Summenwert beinhaltet eine Bepunktung sowohl der korrekt abgelehnten nicht belastbaren Konzepte wie auch der korrekt angenommenen physikalischen Erklärungen bzgl. Dichte und Auftrieb. Tabelle 1 enthält die Mittelwerte und Standardabweichungen der Gruppen MIT, OHNE und der Basisgruppe für Prätest und Posttest. Durch die vorgenommene Bepunktung beim Summenwert ergibt sich eine mittlere Ratewahrscheinlichkeit von 48%, die bei der Interpretation der Mittelwertsunterschiede berücksichtigt werden sollte. Es ergab sich ein signifikanter Zeiteffekt, $F_{(1, 179)} = 339,17$, $p < .001$, $\eta^2 \times 100 = 67$, und eine signifikante Interaktion Zeit x Gruppe, $F_{(2, 179)} = 36,98$, $p < .001$, $\eta^2 \times 100 = 29$, wobei die Gruppe MIT den höchsten Lernzuwachs aufwies. Geplante Kontraste zeigten signifikante Unterschiede zwischen den beiden Experimentalgruppen und der Basisgruppe ($p < .001$) und zwischen der Unterrichtsgruppe MIT und OHNE ($p < .01$). Auch die Basisgruppe verbesserte sich leicht, aber signifikant von Prä- zu Posttest, was v.a. auf die Zunahme richtiger Antworten beim Materialkonzept zurückzuführen

ist und durch Alltagserfahrungen erklärt werden kann, deren Reflexion durch den Test angeregt wurden.

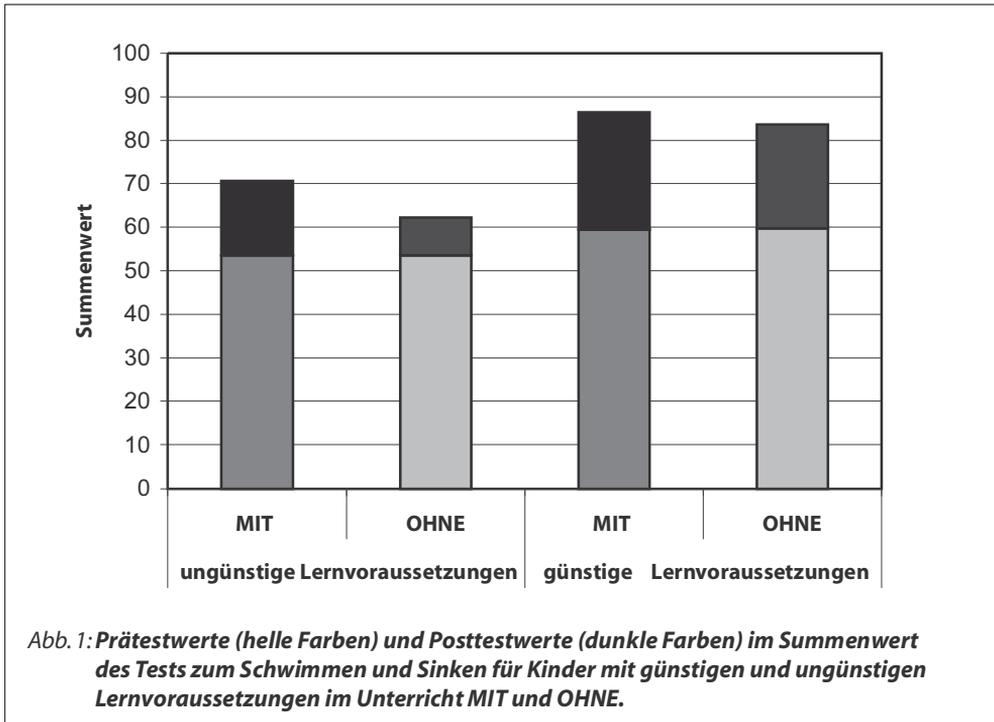
Tab. 1: Mittelwerte (Standardabweichungen) der Unterrichtsgruppen im Summenwert des Prä- und Posttests zum Schwimmen und Sinken.

	Gruppe MIT (N = 71)	Gruppe OHNE (N = 70)	Basisgruppe (N = 41)
Prätest	56,72 (8,35)	56,94 (6,94)	56,88 (8,05)
Posttest	77,04 (11,27)	73,23 (11,14)	61,41 (7,28)

Um insbesondere den Abbau von Fehlkonzepten differenziert zu erfassen, wurde ein Summenwert für alle Antworten auf dem Level nicht belastbarer Konzepte gebildet. Wie erwartet, zeigte sich eine signifikante Abnahme der Fehlkonzepte für beide Experimentalgruppen, nicht aber für die Basisgruppe, wobei die Gruppe MIT signifikant mehr Fehlkonzepte abbaute als die Gruppe OHNE (Hardy u.a. 2001).

Transfertest: In einer univariaten Varianzanalyse mit anschließenden Post-hoc Tests zeigten sich zwischen allen drei Gruppen Unterschiede, $F_{(1, 187)} = 45,61$, $p < .001$, $\eta^2 \times 100 = 33$, mit M (MIT) = 30,80 ($s = 5,45$), M (OHNE) = 28,11 ($s = 5,10$) und M (Basisgruppe) = 22,02 ($s = 3,68$).

Subgruppenanalyse: Eine Analyse von Extremgruppen, die aufgrund von Einschätzungen durch die Lehrkräfte bezüglich „kreativen, selbstständigen, problemlösenden Denkens“ auf einer Skala von 1 bis 16 gebildet wurden, sollte Aufschluss darüber geben, ob Kinder mit besonders günstigen bzw. ungünstigen Leistungsvoraussetzungen unterschiedlich stark von den beiden Unterrichtsformen profitieren. Die Korrelation dieser Einschätzung mit dem Posttestsummenwert beträgt $r_{(144)} = .55$, $p < .01$. Ausgehend von der Verteilung des Lehrerurteils bezüglich problemlösenden Denkens ($M = 9,4$, $s = 2,8$) wurden zwei Extremgruppen an den beiden Enden der Verteilung gebildet. In den Gruppen mit leistungsschwachen und leistungsstarken Kindern sind Mädchen und Jungen zu ähnlichen Anteilen vertreten und alle untersuchten Schulen berücksichtigt. Die Mittelwerte in der Testleistung können Abbildung 1 (S. 184) entnommen werden. Für jede Subgruppe wurden 2 (Zeit) \times 2 (Gruppe) Varianzanalysen mit dem Summenwert des Tests zum Schwimmen und Sinken durchgeführt. Für die Gruppe mit ungünstigen Voraussetzungen ergab sich ein signifikanter Zeiteffekt, $F_{(1, 32)} = 93,39$, $p < .001$, und eine signifikante Interaktion Zeit \times Gruppe, $F_{(1, 32)} = 12,16$, $p < .001$. Wie in der gesamten Stichprobe lernten auch hier die Kinder der Gruppe MIT mehr als die Kinder der Gruppe OHNE. Im Gegensatz zu dieser Gruppe wurde bei der Gruppe der Kinder mit günstigen Voraussetzungen nur der Zeiteffekt, $F_{(1, 19)} = 199,68$, $p < .001$, nicht jedoch die Interaktion statistisch signifikant. Gemäß unseren Erwartungen profitierten Kinder mit guten Lernvoraussetzungen gleichermaßen von beiden Unterrichtsformen, während der Unterricht MIT insbesondere für leistungsschwächere Kinder von Vorteil war.



Erreichtes qualitatives Verständnis: Vergleicht man die im Prä- und Posttest gegebenen schriftlichen, offenen Antworten zu der Frage „Wie kommt es, dass ein schweres Schiff aus Eisen nicht untergeht?“, so zeigt sich, dass nicht belastbare Konzepte wie „Motor“, „Kapitän“, aber auch das Alltagskonzept „Luft“ die Erklärungen in der Präbefragung dominierten. Drei typische Antwortbeispiele der Schüler sind: Kind A: „Auf dem Schiff ist ein Kapiten. Das Schiff trägt schwere sache. Zum beischbil Fische, Öl und Kole“, Kind B: „Weil vielleicht im Schiff Luft drin ist oder weil es bestimmte Motoren hat.“, Kind C: „Vär leich wegen den Luft.“

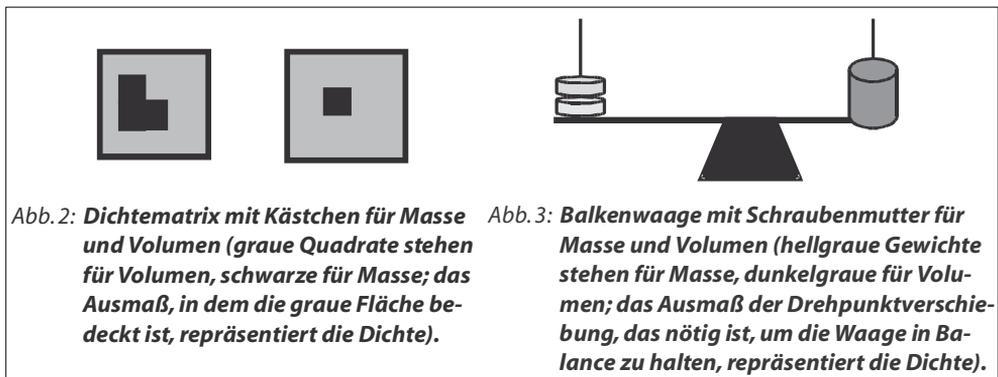
In der Nachbefragung wird dagegen die entscheidende Rolle des Wassers für die Auftriebskraft betont. Dabei werden Zusammenhänge zwischen der Verdrängung, der Auftriebskraft und dem Schwimmen und Sinken von Gegenständen hergestellt: Kind A: „Das Wasser will auf sein alten Platz zurück, und das Wasser drückt es (das Schiff) nach oben.“, Kind B: „Das Schiff drängt ja Wasser weg und dieses Wasser trägt das Schiff, weil das Wasser schwerer und stärker ist, hat es mehr Kraft das Schiff zu tragen. Wenn das Wasser weniger wiegt als das Schiff dann würde das Schiff untergehen.“ Dabei ereignen sich auch reflektierte Konzeptwechsel: Kind C: „Das ligt Nicht an der luft, das ligt auch Nicht an das glachgewicht, es ligt an den Wasser.“

Zusammenfassung der Ergebnisse: In unserer Variation eines konstruktivistisch orientierten Unterrichts zum Thema „Wie kommt es, dass ein Schiff schwimmt?“ mit unterschiedlichem Ausmaß an Strukturierung (Gruppe MIT und Gruppe OHNE) zeigten

beide Unterrichtsgruppen einen deutlichen Lerngewinn, wobei ein konstruktivistisch orientierter Unterricht mit inhaltlicher Sequenzierung und kognitiv strukturierender Gesprächsführung einem stärker selbstgesteuerten, komplexeren Werkstattunterricht sowohl im Mittelwertvergleich des Prä-Posttests zum Schwimmen und Sinken als auch im Transfertest signifikant überlegen war. Die Überlegenheit zeigte sich insbesondere im Abbau von Fehlkonzepten und bei Kindern mit schwächeren Leistungsvoraussetzungen.

3.2 Laborstudie

In der Laborstudie wurde in einem Messwiederholungsdesign mit drei Trainingsgruppen untersucht, welchen unterschiedlichen Einfluss die Nutzung externer Repräsentationsformen (Dichtematrix, Abb. 2; Balkenwaage, Abb. 3) im Vergleich zu einer Kontrollgruppe auf das konzeptuelle Verständnis von „Schwimmen und Sinken“ von Drittklässlern hat.



Wir erwarteten, dass die beiden Gruppen, die mit einer externen Repräsentation arbeiteten, ein tieferes Verständnis der dem Schwimmen und Sinken von Objekten zugrunde liegenden physikalischen Prinzipien erwerben würden als die Kontrollgruppe, die mit einer numerischen Repräsentation arbeitete. Insbesondere war von Interesse, ob die Arbeit mit einer externen Repräsentation auch für Kinder mit schwächeren Eingangsvoraussetzungen von Vorteil sein würde.

Methode

Die Trainings fanden an zwei aufeinander folgenden Nachmittagen statt und dauerten jeweils drei Stunden, inklusive Testzeit. Alle Trainings erforderten ein hohes Maß an Eigenaktivität der Kinder, wobei die Aufgabenabfolge vorab festgelegt war, um in allen Gruppen vergleichbare Trainingsbedingungen zu realisieren. Das Training am ersten Nachmittag zielte auf die Entwicklung des Materialkonzepts ab. Hierzu wurde eine über die Bedingungen konstante Aufgabensequenz durchgeführt, die sich nur in der Anwen-

derung der jeweiligen Repräsentationsform zur Darstellung der Größen Masse und Volumen unterschied.

Im Training am zweiten Nachmittag lernten die Kinder, das von Gegenständen unterschiedlicher Größe verdrängte Wasser zu bestimmen und vorherzusagen, ob ein Gegenstand im Wasser schwimmt oder sinkt. Die hierzu erforderlichen Vergleiche zwischen der Dichte des Wassers und des Gegenstandes erfordern proportionales Denken, das durch die Nutzung der Repräsentationsformen unterstützt wird, welche die zu vergleichenden Verhältnisse abbilden. Wiederum wurde in einer über die Bedingungen konstanten Aufgabenabfolge unterrichtet. Um zu erfassen, ob die Funktionsweise der Repräsentationsformen verstanden wurde, wurden für jede Gruppe Implementationschecks vorgenommen.

Als Prä- und Posttests wurden drei Tests vorgegeben: 1) Ausgewählte Aufgaben aus dem Test zum Schwimmen und Sinken (s. Schulstudie), 2) Aufgaben zum qualitativen Dichteverständnis, 3) Aufgaben zum proportionalen Denken. Versuchspersonen waren 49 Drittklässler (24 Jungen, 25 Mädchen). Diese wurden zufällig einer der drei Trainingsbedingungen mit der Balkenwaage ($N = 17$), der Matrix ($N = 16$) oder der Kontrollbedingung mit Zahlen ($N = 16$) zugewiesen.

Ergebnisse

Zur Analyse des Trainingseffekts der Nutzung unterschiedlicher Repräsentationsformen für das konzeptuelle Verständnis von Schwimmen und Sinken wurden 3 (Gruppe) \times 2 (Zeit) Varianzanalysen mit dem Summenwert des Tests zum Schwimmen und Sinken durchgeführt. Als Kovariaten wurden die Prätestwerte des numerischen und qualitativen Proportionalitätstests einbezogen. Dies ergab sich aus der Überlegung, das mathematische Eingangsverständnis der Kinder bezüglich proportionaler Größen werde das Training bzw. den Lernerfolg mit proportionalen Repräsentationsformen je nach Ausgangsleistung unterschiedlich beeinflussen. Tabelle 2 zeigt die Mittelwerte und Standardabweichungen für die drei Gruppen im Prätest und Posttest.

Tab. 2: Mittelwerte (Standardabweichungen) der Trainingsgruppen im Summenwert des Prä- und Posttests zum Schwimmen und Sinken.

	Balkenwaage ($N = 17$)	Dichtematrix ($N = 16$)	Kontrollgruppe ($N = 16$)
Prätest	35,00 (4,43)	36,37 (7,15)	33,25 (4,92)
Posttest	45,76 (6,07)	44,06 (6,20)	40,94 (7,67)

Es ergab sich ein signifikanter Lerngewinn der Kinder aller drei Gruppen, $F_{(1, 46)} = 89,62$, $p < .001$, $\eta^2 \times 100 = 23$, mit einem signifikanten Beitrag der Kovariate Proportionalitätstest ($p < .10$). Die Interaktion zwischen Zeit und Gruppe ist jedoch trotz einer Überlegenheit der Balkenwaagengruppe nicht signifikant.

Bei einer getrennten Analyse bezüglich der Abnahme der im Test angekreuzten Fehlkonzepte zeigte sich, dass die Balkenwaagengruppe im Posttest signifikant weniger Fehlkonzepte

annahme als die Kontrollgruppe, während sich die Matrixgruppe von keiner der beiden Gruppen unterschied.

Subgruppenanalyse: Anhand der Prätestwerte der Tests zum proportionalen Denken und zum qualitativen Dichteverständnis wurden zwei Gruppen von Kindern gebildet, die sich aus Kindern mit günstigen und ungünstigen Lernvoraussetzungen für ein proportionales Dichtetraining zusammensetzten. Das Kriterium zur Gruppenbildung lag bei 50% korrekten Lösungen der Prätestaufgaben. Eine 2 (Zeit) x 3 (Gruppe) x 2 (Subgruppe) Messwiederholungsanalyse ergab einen signifikanten Zeiteffekt, $F_{(1, 43)} = 81,45$, $p < .001$, und eine signifikante Zeit x Gruppe x Subgruppe Interaktion, $F_{(1, 43)} = 4,57$, $p < .05$. Tabelle 3 zeigt Mittelwerte und Standardabweichungen für den Lernzuwachs von Prätest zu Posttest der Gruppen und Subgruppen.

Tab. 3: Mittelwerte (Standardabweichungen) der Trainingssubgruppen im Lernzuwachs beim Summenwert von Prä- und Posttest zum Schwimmen und Sinken.

	Balkenwaage		Dichtematrix		Kontrollgruppe	
Proportionalitätswerte	Hoch (N = 7)	Niedrig (N = 10)	Hoch (N = 9)	Niedrig (N = 7)	Hoch (N = 5)	Niedrig (N = 11)
Lernzuwachs Prätest/ Posttest	9,14 (6,28)	11,90 (7,88)	4,11 (5,39)	12,28 (8,42)	12,60 (10,36)	5,45 (4,25)

Obwohl die Zellenbesetzung gering ist und etwas variiert, können aus dieser Analyse interessante Trends abgelesen werden. Während bei der Balkenwaagengruppe Kinder mit hohen und niedrigen Proportionalitätswerten etwa in gleichem Maße vom Training profitierten, gibt es bei Matrixgruppe und Kontrollgruppe signifikante Unterschiede hinsichtlich des Lernzuwachses in den beiden Subgruppen. In der Kontrollgruppe profitierten erwartungsgemäß besonders diejenigen Kinder, die schon vor dem Training ein gutes Proportionalitätsverständnis aufwiesen; die Kinder dieser Subgruppe zeigten jedoch im Matrixtraining einen besonders geringen Lernzuwachs. Im Gegensatz hierzu profitierten in der Matrixbedingung insbesondere die Kinder mit niedrigen Proportionalitätswerten.

Die unterschiedliche Lernentwicklung der beiden Subgruppen in den drei Bedingungen zeigte sich auch in getrennten Varianzanalysen: Kinder mit niedrigen Proportionalitätswerten profitierten signifikant mehr von der Teilnahme an der Balkenwaagenbedingung und Dichtematrixbedingung im Vergleich zur Kontrollbedingung, was beim Vergleich des Lernzuwachses von Prätest zu Posttest ersichtlich wird, $F_{(2, 25)} = 3,13$, $p < .05$. Kinder mit hohen Proportionalitätswerten profitierten gleichermaßen von der Teilnahme am Balkenwaagentraining und dem numerischen Training der Kontrollbedingung, während sie im Dichtematrixtraining weniger profitierten, $F_{(2, 18)} = 2,51$ n.s. In Gruppenvergleichen ergibt sich hier ein signifikanter Unterschied zwischen Dichte-

matrix- und Kontrollbedingung ($p < .05$), während sich die Balkenwaagenbedingung nicht von den anderen beiden Gruppen unterscheidet.

Aus diesen Analysen ist zu erkennen, dass vom Balkenwaagentraining sowohl Kinder mit günstigen und ungünstigen Lernvoraussetzungen im proportionalen Denken profitierten, während das Dichtematrixtraining nur für Kinder mit schlechteren und das Kontrollgruppentraining nur für Kinder mit besseren Werten im Proportionalitätsverständnis erfolgreich war.

Zusammenfassung der Ergebnisse: Es zeigte sich, dass ein relativ kurzes Training in allen drei Gruppen zu einem signifikanten Lernzuwachs bezüglich des Verständnisses von Schwimmen und Sinken führte. Die tendenzielle Überlegenheit der Balkenwaagengruppe im Summenwert des Tests zum Schwimmen und Sinken ist vor allem auf die signifikant häufigere Aufgabe von Fehlkonzepten zurückzuführen. Die Nutzung der Balkenwaage war sowohl für Kinder mit schwächeren als auch mit guten Eingangsvoraussetzungen bezüglich des proportionalen Verständnisses von Vorteil. Dies ist vermutlich darauf zurückzuführen, dass gleichzeitig visuelle und numerische Interpretationen an der Balkenwaage vorgenommen werden konnten, wobei insbesondere die Anwendung des Gleichgewichtsprinzips intuitiv interpretierbar ist.

4. Diskussion der Ergebnisse und Konsequenzen für den Grundschulunterricht

Die Schul- wie auch die Laborstudie zeigen, dass ein konzeptuelles Verständnis von physikalischen Inhalten wie Dichte und Auftrieb schon im Grundschulalter aufgebaut werden kann. In beiden Studien wurde deutlich, dass insbesondere Strukturierungselemente in Lehr-Lernumgebungen den Konzeptwechsel erleichtern, weil sie die Kinder dabei unterstützen, nicht belastbare Präkonzepte aufzugeben.

Strukturierungsangebote erlauben sowohl die Fokussierung relevanter Dimensionen bei der Wahrnehmung und Interpretation von Phänomenen wie auch die direkte Konfrontation (z.B. durch Visualisierung) mit nicht erwarteten Ergebnissen, die zur Aufgabe von nicht adäquaten Erklärungsmustern anregen. Strukturierungselemente schränken somit die zu interpretierende Umwelt so ein, dass neue Erklärungsansätze überzeugend aufgebaut und in ein bestehendes Situationsverständnis integriert werden können. Sie wirken sich insbesondere positiv auf die Förderung leistungsschwächerer Kinder aus.

Die nichtleistungsbezogenen Ergebnisse in der Schulstudie zeigen zudem, dass strukturierte Lernangebote, die zugleich ausgiebig Möglichkeiten zur aktiven Entwicklung und Überprüfung von Konzepten geben, eine multikriteriale Zielerreichung fördern (Hardy u.a. 2001). In einer weiterführenden Studie soll untersucht werden, ob eine Integration externer Repräsentationsformen in einen Unterricht mit inhaltlicher Sequenzierung und kognitiv strukturierender Gesprächsführung zu einer weiteren Optimierung von Lernumgebungen beiträgt.

Wir fühlen uns aufgrund der Untersuchungen in der Annahme bestätigt, dass der Sachunterricht der Grundschule stärker als bisher für ein Lernen im Vorfeld der Physik genutzt werden sollte. Der Unterricht muss Kindern dabei ausreichend Gelegenheit geben, ihre Ideen und Deutungen in Gesprächen zu entwickeln, diese in Experimenten zu überprüfen und Erklärungen zu konstruieren. Auf eigensprachliche Formulierungen ist Wert zu legen, auch wenn diese dem physikalischen Sprachgebrauch noch nicht entsprechen. Als Gegenstand des Unterrichts sollten Phänomene des Alltags gewählt werden, die Grundschul Kinder in Erstaunen versetzen. Über das Anwenden des Gelernten und das Verstehen weiterer Phänomene können Kinder Kompetenz erleben – eine wichtige Voraussetzung für motiviertes Lernen und für eine positive Einschätzung der eigenen, bereichsspezifischen Fähigkeiten. Besonders wichtig ist die Auseinandersetzung mit den vorhandenen Präkonzepten und deren Förderung durch strukturierende Elemente. Ein konzeptwechselorientierter Unterricht braucht Zeit; diese ist nur durch einen exemplarisch angelegten Unterricht zu gewinnen.

Die von den Schülern in einem solchen Unterricht aufgebauten Vorstellungen und Denkweisen wie auch die entwickelte Lernzuversicht und Lernmotivation bieten gute Voraussetzungen für den physikalischen Unterricht in den weiterführenden Schulen. In diesem Sinne kann die Grundschule einen Beitrag zur Förderung des naturwissenschaftlichen Unterrichts leisten, der angesichts der Ergebnisse der PISA-Studie sicherlich wünschenswert ist.

Literatur

- Beck, G./Claussen, C. (1979): Einführung in die Probleme des Sachunterrichts. Kronberg/Ts.: Scriptor.
- Carey, S. (1985): *Conceptual change in childhood*. Cambridge: MIT Press.
- Cox, R. (1999): Representation construction, externalised cognition and individual differences. In: *Learning and Instruction* 9, S. 343–363.
- Dubs, R. (1997): Der Konstruktivismus im Unterricht. In: *Schweizer Schule* 6, S. 26–36.
- Duit, R./Treagust, D.F. (1998): Learning in science – From behaviourism towards social constructivism and beyond. In: Fraser, B.J./Tobin, K.G. (Eds.): *International Handbook of Science Education*. Dordrecht; Boston; London: Kluwer Academic Publishers, S. 3–26.
- Duit, R. (1999): Conceptual change approaches in science education. In: Schnotz, W./Vosniadou, S./Carretero, M. (Eds.): *New Perspectives on conceptual change*. Amsterdam; New York; Oxford: Pergamon, S. 263–282.
- Duschl, R.A./Hamilton, R.J. (1998): Conceptual change in science and in the learning of science. In: Fraser, B.J./Tobin, K.G. (Eds.): *International Handbook of Science Education*. Dordrecht; Boston; London: Kluwer Academic Publishers, S. 1047–1065.
- Einsiedler, W. (1978): Selbststeuerung und Lernhilfen im Unterricht. In: Neber, H./Wagner, A.C./Einsiedler, W. (Hrsg.): *Selbstgesteuertes Lernen. Psychologische und pädagogische Aspekte eines handlungsorientierten Lernens*. Weinheim; Basel: Beltz, S. 192–213.
- Einsiedler, W. (1996): Wissensstrukturierung im Unterricht. Neuere Forschung zur Wissensrepräsentation und ihre Anwendung in der Didaktik. In: *Zeitschrift für Pädagogik* 42, S. 167–192.
- Friedrich, H.F./Mandl, H. (1997): Analyse und Förderung selbstgesteuerten Lernens. In: Weinert, F.E./Mandl, H. (Hrsg.): *Psychologie der Erwachsenenbildung. Enzyklopädie der Psychologie*, Bd. 4. Göttingen: Hogrefe, S. 237–293.

- Gerstenmaier J./Mandl, H. (1995): Wissenserwerb unter konstruktivistischer Perspektive. In: *Zeitschrift für Pädagogik* 41, S. 867–887.
- Gräsel, C./Mandl, H. (1993): Förderung des Erwerbs diagnostischer Strategien in fallbasierten Lernumgebungen. In: *Unterrichtswissenschaft* 21, S. 355–369.
- Greeno, J./Smith, D./Moore, J. (1993): Transfer of situated learning. In: Detterman, D./Sternberg, R. (Eds.): *Transfer on trial: Intelligence, cognition and instruction*. Norwood: Ablex Publishing Company, S. 99–167.
- Hardy, I./Jonen, A./Möller, K./Stern, E./Blumberg, E. (2001): Abschlussbericht BIQUA Projekt Münster/Berlin.
- Haru, S. (2000): Changes in Children's Conceptions through Social Interaction in Pre-school Science Education. Joensuu: University of Joensuu (= Publications in Education No: 60)
- Hewson, P.W./Beeth, M.E./Thorley, N.R. (1998): Teaching for conceptual change. In: Fraser, B.J./Tobin, K.G. (Eds.): *International Handbook of Science Education*. Dordrecht; Boston; London: Kluwer Academic Publishers, S. 199–218.
- Janke, B. (1995): Entwicklung naiven Wissens über den physikalischen Auftrieb. Warum schwimmen Schiffe? In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 27, H. 2, Göttingen: Hogrefe, S. 122–138.
- Köhnlein, W. (1999): Vielperspektivität und Ansatzpunkte naturwissenschaftlichen Denkens. Analysen von Unterrichtsbeispielen unter dem Gesichtspunkt des Verstehens. In: Köhnlein, W./Marquardt-Mau, B./Schreier, H. (Hrsg.): *Vielperspektivisches Denken im Sachunterricht*. Bad Heilbrunn: Klinkhardt, S. 88–124.
- Koerber, S. (2000): Der Einfluss externer Repräsentationsformen auf proportionales Denken im Grundschulalter. Dissertation an der Technischen Universität Berlin.
- Möller, K. (1999): Konstruktivistisch orientierte Lehr-Lernprozessforschung im naturwissenschaftlich-technischen Bereich des Sachunterrichts. In: Köhnlein, W. (Hrsg.): *Vielperspektivisches Denken im Sachunterricht*. Bad Heilbrunn: Klinkhardt, S. 125–191 (= *Forschungen zur Didaktik des Sachunterrichts*, Bd. 3).
- Möller, K. (2001a): Lernen im Vorfeld der Naturwissenschaften. Zielsetzungen und Forschungsergebnisse. In: Köhnlein, W./Schreier, H. (Hrsg.): *Innovation Sachunterricht. Befragung der Anfänge nach zukunftsfähigen Beständen*. Bad Heilbrunn: Klinkhardt, S. 275–298 (= *Forschungen zur Didaktik des Sachunterrichts*, Bd. 4).
- Möller, K. (2001b): Konstruktivistische Sichtweisen für das Lernen in der Grundschule? In: Czerwenka, K.; Nölle, K.; Roßbach, H.-G. (Hrsg.): *Jahrbuch Grundschulforschung*, Bd. 4. Opladen: Leske und Buderich, S. 16–31.
- Möller, K. (einger.): Anspruchsvolles Lernen in der Grundschule – am Beispiel naturwissenschaftlich-technischer Inhalte. In: *Pädagogische Rundschau*.
- Reinmann-Rothmeier, G./Mandl, H. (1997): Wissensvermittlung. Ansätze zur Förderung des Wissenserwerbs. In: Birbaumer, N. u.a. (Hrsg.): *Enzyklopädie der Psychologie*, Bd. 6. Göttingen; Bern; Toronto; Seattle: Hogrefe, S. 457–500.
- Reinmann-Rothmeier, G./Mandl, H. (1999): Instruktion. In: Perleth, C./Ziegler, A. (Hrsg.): *Pädagogische Psychologie. Grundlagen und Anwendungsfelder*. Bern; Göttingen; Toronto; Seattle: Huber, S. 207–215.
- Renkl, A./Stern, E. (1994): Die Bedeutung von kognitiven Eingangsvoraussetzungen und Lernaufgaben für das Lösen von einfachen und komplexen Textaufgaben. In: *Zeitschrift für Pädagogische Psychologie* 8, S. 27–39.
- Renkl, A. (1995): Learning for later teaching: An exploration of mediational links between teaching expectancy and learning results. In: *Learning and Instruction* 5, S. 21–36.
- Scott, P.H./Driver, R.H. (1998): Learning about science teaching. Perspectives from an action research project. In: Fraser, B.J./Tobin, K.G. (Eds.): *International Handbook of Science Education*. Dordrecht; Boston; London: Kluwer Academic Publishers, S. 67–80.

- Smith, C./Carey, S./Wiser, M. (1985): On differentiation: A case study of the development of the concepts of size, weight, and density. In: *Cognition* 21, S. 177–237.
- Smith, C./Maclin, D./Grosslight, L./Davis, H. (1997): Teaching for understanding. A study of students' preinstruction theories of matter and a comparison of the effectiveness of two approaches to teaching about matter and density. In: *Cognition and Instruction* 15, S. 317–393.
- Sodian, B. (1995): Entwicklung bereichsspezifischen Wissens. In: Oerter, R./Montada, L. (Hrsg.): *Entwicklungspsychologie*. Weinheim: Psychologie Verlags Union, S. 622–653.
- Stark, R./Gruber, H./Mandl, H. (1998): Motivationale und kognitive Passungsprobleme beim komplexen situierten Lernen. In: *Psychologie in Erziehung und Unterricht* 45, S. 202–215.
- Staub, F./Stern, E. (2002): The nature of teachers' pedagogical content beliefs matters for students' achievement gains. Quasi-experimental evidence from elementary mathematics. In: *The Journal of Educational Psychology* 93, S. 144–155.
- Stern, E. (2001): Intelligenz, Wissen, Transfer und der Umgang mit Symbolsystemen. In: Stern, E./Guthke, J. (Hrsg.): *Perspektiven der Intelligenzforschung*. Lengerich: Pabst Publisher, S. 163–204.
- Stern, E./Hardy, I./Koerber, S. (2002): Die Nutzung grafisch-visueller Repräsentationsformen im Sachunterricht. In: Spreckelsen, K./Harteringer, A./Möller, K. (Hrsg.): *Ansätze und Methoden empirischer Forschung zum Sachunterricht*. Bad Heilbrunn: Klinkhardt, S. 119–131.
- Stern, E. (2002): Wie abstrakt lernt das Grundschulkind? Neuere Ergebnisse der entwicklungspsychologischen Forschung. In: Petillon, H. (Hrsg.): *Handbuch der Grundschulforschung*, Bd. 5. Leverkusen: Leske und Budrich, S. 22–28.
- Stern, E. (im Druck): Früh übt sich. Neuere Ergebnisse aus der LOGIK-Studie zum Lösen mathematischer Textaufgaben in der Grundschule. In: Fritz, A./Ricken, G./Schmidt, S. (Hrsg.): *Handbuch Rechenschwäche, Lernwege, Schwierigkeiten und Hilfen*. Weinheim: Beltz.
- Strunck, U./Lück, G./Demuth, R. (1998): Der naturwissenschaftliche Sachunterricht in Lehrplänen, Unterrichtsmaterialien und Schulpraxis. Eine quantitative Analyse der Entwicklung in den letzten 25 Jahren. In: *Zeitschrift für Didaktik der Naturwissenschaften*. Biologie, Chemie, Physik 4, H. 1, S. 69–81.
- Thiel, S. (1990): Grundschulkindern zwischen Umgangserfahrung und Naturwissenschaft. In: Wagen-schein, M.: *Kinder auf dem Weg zur Physik*. Weinheim; Basel: Beltz, S. 90–180.
- Vosniadou, S./Ioannides, C./Dimitrakopoulou, A./Papademetriou, E. (2001): Designing learning environments to promote conceptual change in science. In: *Learning and Instruction* 15, S. 317–419.
- Weinert, F.E. (1982): Selbstgesteuertes Lernen als Voraussetzung, Methode und Ziel des Unterrichts. In: *Unterrichtswissenschaft* 10, S. 99–110.
- Weinert, F.E. (1996): Für und Wider die „neuen Lerntheorien“ als Grundlagen pädagogisch-psychologischer Forschung. In: *Zeitschrift für Pädagogische Psychologie* 10, S. 1–12.

Anschrift der Autorinnen:

Prof. Dr. Kornelia Möller, Seminar für Didaktik des Sachunterrichts, Westfälische-Wilhelms-Universität Münster, Leonardo Campus 11, 48149 Münster.

Angela Jonen, Seminar für Didaktik des Sachunterrichts, Westfälische-Wilhelms-Universität Münster, Leonardo Campus 11, 48149 Münster.

Dr. Ilonca Hardy, Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, 14195 Berlin.

Prof. Dr. Elsbeth Stern, Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, 14195 Berlin.

Beate Sodian/Claudia Thoermer/Ernst Kircher/Patricia Grygier/Johannes Günther

Vermittlung von Wissenschaftsverständnis in der Grundschule¹

1. Einleitung

1.1 Was versteht man unter „Wissenschaftsverständnis“?

„Wissenschaftsverständnis“ („Understanding the Nature of Science“) beinhaltet Einsicht in erkenntnistheoretische, wissenschaftstheoretische und wissenschaftsethische Grundlagen der Naturwissenschaften. Im Zentrum unseres Interesses stehen wissenschaftstheoretische Fragen, z.B.: Wie entsteht naturwissenschaftliches Wissen? Wie kommen Wissenschaftler² zu neuen Erkenntnissen? Was sind Experimente? Warum werden Theorien wieder geändert oder verworfen? Können die Daten eines Experiments durch verschiedene Theorien erklärt werden? Nach einer Analyse von offiziellen Dokumenten von Schulbehörden und naturwissenschaftlichen Gesellschaften aus anglophonen Staaten (McComas/McClough/Almaroza 1998) gehört zum Kanon des Wissens über Wissenschaft u.a. Einsicht in den tentativen Charakter und die rationale Begründung wissenschaftlichen Wissens, die Bedeutung von Beobachtungen und experimentellen Ergebnissen, die naturwissenschaftliche Methodik (wobei betont wird, dass es keine allgemein gültige Methode gibt, Wissenschaft zu betreiben), die Rolle von Theorien (und die Unterscheidung zwischen Theorien und Gesetzen) im wissenschaftlichen Erkenntnisprozess, den öffentlichen Status wissenschaftlichen Wissens und den evolutionären wie auch revolutionären Charakter der Wissenschaften.

Pädagogische und didaktische Begründungen für einen Unterricht, der die Natur der Naturwissenschaften thematisiert, existieren seit langem im deutschen und im angloamerikanischen Sprachraum (Litt 1959; Dewey 1916/1964). In den vergangenen zwanzig Jahren hat diese Thematik vor allem in den USA und in England eine beträchtliche Rolle sowohl in der Naturwissenschaftsdidaktik als auch – unter dem Einfluss der Forschung zu Prozessen des begrifflichen Wandels („Conceptual Change“) in der Pädagogischen Psychologie gespielt.

Driver u.a. (1996, S. 16ff.) begründen die Thematisierung wissenschaftstheoretischer Aspekte im naturwissenschaftlichen Unterricht durch ethische und gesellschaftspolitische, sowie kognitionspsychologische Argumente. Wir beschränken uns im Folgenden auf Letztere.

- 1 Das Projekt wird durch eine Sachbeihilfe der DFG (So 213/20-1, Ki 320/3-1) im Rahmen des Schwerpunktprogramms „Bildungsqualität von Schule“ gefördert.
- 2 Im gesamten Manuskript verwenden wir aus Gründen der Lesbarkeit lediglich die männliche Form. Hiermit sind selbstverständlich männliche und weibliche Personen gemeint.

1.2 Ist Wissenschaftsverständnis wichtig für naturwissenschaftliches Lernen?

Moderne kognitionspsychologische Theorien analysieren Wissen in naturwissenschaftlichen Domänen oder Bereichen (physikalisches, chemisches, biologisches Wissen) als kohärente Systeme interdependenter, theoriegeladener Begriffe (Carey 1986; Vosniadou 1994; Chi/Slotta/deLeeuw 1994). Anders als ältere lernpsychologische Theorien, die von allgemeinen, bereichsübergreifenden Lerngesetzen ausgingen, impliziert diese Sichtweise die Domänenspezifität der kognitiven Prozesse beim Lernenden: Physikalische Theorien postulieren andere Erklärungsmodelle als biologische oder chemische. Daher müssen Schüler bereichsspezifische begriffliche Systeme und Erklärungsmodelle erwerben. Dies ist – so die heute in der Kognitionspsychologie ebenso wie in den Naturwissenschaftsdidaktiken vertretene Position – deshalb so schwierig, weil die Präkonzepte der Lernenden häufig mit den wissenschaftlichen Zielkonzepten inkompatibel sind. Daher wird das Lernen der Naturwissenschaften, insbesondere der Physik, gelegentlich mit dem Einleben in eine fremde Kultur, mit dem Erwerb einer neuen Sprache verglichen (Duit 2002). Das bedeutet, dass die aus der Alltagssprache übernommenen Ausdrücke wie „Arbeit“, „Kraft“, „Geschwindigkeit“, „Masse“ usw. mit komplexen, abstrakten Theorien zu tun haben und erst dann im physikalischen Sinne verstanden werden können, wenn das gesamte Begriffssystem einer Theorie verstanden ist, – zu obigen Begriffen die Newtonsche Mechanik (Kircher 1993). Die TIMS-Studie (Baumert u.a. 2000) hat gezeigt, dass begrifflicher Wandel (Conceptual Change) im Sinne einer vollständigen Restrukturierung intuitiver Alltagsvorstellungen mit dem Ergebnis eines tiefen Verständnisses naturwissenschaftlicher Domänen häufig auch im naturwissenschaftlichen Unterricht der Sekundarstufe II nicht stattfindet: Aufgaben, die eine Änderung von Fehlvorstellungen überprüfen, bereiteten selbst Abiturienten die größten Schwierigkeiten.

Für dieses Versagen der naturwissenschaftlichen Bildung können neben Mängeln in der Vermittlung bereichsspezifischen Wissens auch Defizite in der Überwindung bereichsübergreifender Verständnishindernisse verantwortlich gemacht werden: Um physikalische oder biologische Rahmentheorien („Weltbilder“) zu verändern, genügt es nicht, domänen-spezifische Begriffe zu vermitteln. Vielmehr gehören zum naturwissenschaftlichen „Weltbild“ auch adäquate Vorstellungen über die Natur der Naturwissenschaften, sowie über das Zustandekommen und die Begründbarkeit naturwissenschaftlichen Wissens. Um den Theorie-Evidenz-Bezug in verschiedenen Wissenschaftsdisziplinen zu verstehen, ist ein solches metabegriffliches Verständnis dieses Zusammenhangs notwendig – oder zumindest förderlich. Daher wird der Vermittlung eines adäquaten Wissenschaftsverständnisses ein hoher Stellenwert bei der erfolgreichen Vermittlung naturwissenschaftlichen Wissens zugeschrieben (Carey/Smith 1993).

1.3 Zum Stand der Forschung: Wissenschaftsverständnis bei Schülern und Lehrern

Empirische Studien zu den intuitiven Vorstellungen von Schülern und Lehrern über Wissenschaft zeigen, dass das Wissenschaftsverständnis von Schülern der Sekundar-

stufe I (aber auch II) meist einer unreflektierten „knowledge unproblematic“ epistemologischen Position entspricht (Carey/Smith 1993). Diese ist gekennzeichnet durch eine mangelnde Differenzierung zwischen Theorien/Hypothesen einerseits und empirischer Evidenz andererseits sowie durch ein unzureichendes Verständnis des zyklischen und kumulativen Charakters naturwissenschaftlichen Wissens (Ryan/Aikenhead 1992; Lederman 1992; Kircher 1995, Driver u.a. 1996; McComas 1998). Korrespondierende Defizite wurden auf Lehrerseite festgestellt (Gallagher 1991; Lederman 1992; Pomeroy 1993; Höttecke 2001), wobei Sekundarstufenlehrer dazu tendieren, den „objektiven“ Charakter der Naturwissenschaften zu betonen und den tentativen und konstruktiven Charakter wissenschaftlicher Theoriebildung zu vernachlässigen, während Primarstufenlehrer eher eine nicht-traditionelle, antipositivistische Sichtweise zeigen (Pomeroy 1993), die möglicherweise auf radikal-relativistischen Missverständnissen einer konstruktivistischen Wissenschaftsphilosophie basieren.

Eine eingehende Untersuchung des Wissenschaftsverständnisses amerikanischer Siebtklässler (Junior High School Schüler) wurde von Carey u.a. (1989) mittels eines Interviews zum Wissenschaftsverständnis (Nature-of-Science-Interview) durchgeführt. Das Interview enthält Fragen zu den Zielen von Wissenschaft, der Art von Fragen, die Wissenschaftler bearbeiten, der Rolle von Ideen und Theorien bei der Planung von Experimenten, der Interpretation von Ergebnissen, und der Revision und Weiterentwicklung von Theorien.

Die Mehrheit der Siebtklässler in der Studie von Carey u.a. (1989) zeigte kein Verständnis des Theorie-Evidenz-Bezugs. Auch Schüler der elften Klasse (amerikanische High-School) antworteten nicht konsistent auf einem Niveau, das Verständnis von Wissenschaft als die Suche nach Erklärungen und wissenschaftliches Wissen als das Ergebnis der Prüfung von Theorien und Hypothesen anzeigt (Honda 1994). Ein Verständnis des zyklisch-kumulativen Charakters wissenschaftlicher Erkenntnis wird selbst von Erwachsenen mit naturwissenschaftlicher Ausbildung nur selten erreicht (Thoermer/Sodian, im Druck).

1.4 Effekte von Unterricht

Eine Reihe neuerer Grund- und Sekundarschulcurricula zu unterschiedlichen naturwissenschaftlichen Inhaltsbereichen enthält wissenschaftstheoretische Komponenten, d.h., die explizite Reflexion über die Bildung, Prüfung und Revision von Theorien und Hypothesen und den Hypothese-Evidenz-Zusammenhang (Brown/Campione 1994; White/Frederiksen 1998). Förderliche Effekte solcher metakognitiver Komponenten auf das fachliche Lernen gerade schwächerer Schüler konnten nachgewiesen werden (White/Frederiksen 1998). Über die Behandlung der Natur naturwissenschaftlichen Wissens als Unterrichtsfach in unterschiedlichen Schulsystemen geben McComas/Olson (1998) einen Überblick.

Aufgrund einer Defizitanalyse herkömmlicher Nature-of-Science-Curricula entwickelten Carey u.a. (1989) eine siebenstündige Unterrichtseinheit zur Vermittlung wis-

wissenschaftsphilosophischer Grundbegriffe. Am Beispiel der Exploration der Frage, warum Brotteig aufgeht, wurden Siebtklässler in eine gesteuerte Reflexion der Bildung und Prüfung ihrer Theorien und Hypothesen involviert, wobei Wert auf die explizite Vermittlung von metatheoretischen Konzepten (Hypothese, Experiment, Theorie) gelegt wurde. Im Prä-Posttest-Vergleich wurden anhand des „Nature-of-Science-Interviews“ von Carey u.a. (1989) die Effekte des Curriculums auf den individuellen Fortschritt im Verständnis von Wissenschaft untersucht. Durchschnittlich konnte ein Fortschritt von einem Verständnis von Wissenschaft als reine Aktivität oder Faktensammlung hin zu einem rudimentären Verständnis von Wissenschaft als Erklärung erreicht werden.

Die Effekte *langfristigen* wissenschaftstheoretisch reflektierten Unterrichts dokumentiert eine Fallstudie von Smith u.a. (2000), die anhand eines Vergleichs zweier sechster Klassen die Effekte unterschiedlicher „Wissenschaftsphilosophien“ der Lehrerinnen auf das epistemologische Verständnis ihrer Schüler belegt. Die beiden hinsichtlich soziodemografischer Merkmale vergleichbaren Klassen wurden über einen Zeitraum von mehreren Jahren in Naturwissenschaften von Lehrerinnen unterrichtet, die unterschiedliche epistemologische Überzeugungen (explizit konstruktivistisch vs. eher traditionell induktivistisch) hatten und diese in ihrem unterrichtlichen Handeln auch umsetzten. Ein Vergleich der Schülerepistemologien (mit dem Nature-of-Science-Interview von Carey u.a. 1989) zeigte, dass die Schüler, die aus einer wissenschaftstheoretisch fundierten Perspektive unterrichtet wurden, mehrheitlich ein Verständnis von Wissenschaft als Suche nach Erklärungen entwickelt hatten, während die traditionell (eher induktivistisch) unterrichteten Schüler Wissenschaft primär als das Sammeln von Fakten verstanden. Dieser Befund weist auf die langfristige Bedeutung eines frühen wissenschaftstheoretisch fundierten Unterrichts über die Natur der Naturwissenschaften schon in der Grundschule hin.

1.5 Vermittlung von Wissenschaftsverständnis schon in der Grundschule?

Die bisher vorliegenden Interventionsstudien im Unterricht und in der Lehrerbildung beschränken sich auf die Sekundarstufen (Meyling 1997; Solomon 1991) und die dort unterrichtenden Lehrerinnen und Lehrer (Höttecke 2001). Für die Grundschule liegen keine entsprechenden Studien vor, was u.a. darauf zurückzuführen ist, dass in der traditionellen, an Piaget orientierten Entwicklungspsychologie Kindern im Grundschulalter die Fähigkeit zum naturwissenschaftlichen Denken im Sinne der systematischen Bildung, Prüfung und Revision von Theorien und Hypothesen abgesprochen wurde. Die Fähigkeit, hypothetisch (besser: theoretisch) zu denken, entwickle sich erst im Jugendalter, setze formale Operationen voraus und sei gekennzeichnet durch die inhaltsunabhängige Anwendung von formalen Standards wissenschaftlicher Rationalität. In der neueren Forschung wurde sowohl das Konzept der formalen Operationen kritisiert als auch die damit verbundene Vorstellung von bereichsunabhängig anwendbaren idealen Standards des wissenschaftlichen Denkens (Chinn/Brewer 1993; siehe auch Möller u.a., in diesem Band). Ferner wurde gezeigt, dass Kinder im Grundschulalter systemati-

sche Strategien der Hypothesenprüfung (Variablenisolation und -kontrolle) erlernen können (Case 1974; Chen/Klahr 1999) und dass ihr Verständnis der Merkmale eines „guten“ Experiments ihrer Fähigkeit, solche Experimente selbst zu produzieren, vorausläuft (Bullock/Ziegler 1999). In unterstützenden Aufgaben-Kontexten differenzieren ältere Grundschüler zwischen Hypothese und Evidenz und zeigen ein Grundverständnis der Logik der Hypothesenprüfung (Sodian/Zaitchik/Carey 1991; Schrempp/Sodian 1999). Ferner zeigen ältere Grundschüler rudimentäre Fähigkeiten zur Evaluation von Theorien und zur Einschätzung des Erklärungswerts konkurrierender Theorien (Samarapungavan 1992; Leach 1999). Diese Befunde deuten darauf hin, dass die begrifflichen Voraussetzungen für Grundzüge wissenschaftstheoretischer Reflexion bereits bei älteren Grundschulern vorhanden sind.

Daher versuchen wir im vorliegenden Projekt in einem ersten Schritt das intuitive Wissenschaftsverständnis von Schülern der dritten und vierten Jahrgangsstufe differenziert zu beschreiben und durch kurzfristige Interventionen mittels experimenteller Curricula zu prüfen, ob, und wenn ja, in welchem Ausmaß, die epistemologischen Präkonzepte durch die Anregung wissenschafts- bzw. erkenntnistheoretischer Reflexion beeinflussbar sind. Im zweiten Schritt prüfen wir, ob reflektiertes Wissenschaftsverständnis für den Erwerb inhaltlichen naturwissenschaftlichen Wissens förderlich ist (ob also eine Klasse, die ein wissenschaftstheoretisches Curriculum erhalten hat, einer Kontrollklasse im Erwerb inhaltlichen Wissens in einem physikalischen Themenbereich überlegen ist). Hier berührt unser Projekt die Fragestellungen des Projekts von Möller & Stern (Möller u.a., in diesem Band), die die Wirksamkeit eines moderat konstruktivistischen Unterrichts auf den konzeptuellen Wandel in der Physik bei Grundschulern der 3. und 4. Jahrgangsstufe belegt haben³. Da bedeutsame Beziehungen zwischen dem Wissenschaftsverständnis von Schülern und dem ihrer Lehrer gezeigt worden sind, untersuchen wir das Wissenschaftsverständnis von Grundschullehrerinnen und -lehrern als bedeutendste Kontextbedingung für das Wissenschaftsverständnis der Schüler. Aufbauend auf der Deskription der Lehrerepistemologien ist eine Interventionsstudie im Rahmen der Lehrerfortbildung geplant. Außerdem werden als Kontextvariablen Schul- und Jugendsachbücher analysiert.

Im vorliegenden Artikel berichten wir über die erste Studie des Projekts, die der Beschreibung des Wissenschaftsverständnisses von Grundschulern der vierten Klassenstufe und der Prüfung der Effekte eines wissenschaftstheoretischen Curriculums diene.

3 Eine weitere Kooperationsbeziehung (nicht an das Grundschulalter gebunden) besteht zum Projekt von Reiss/Hellmich/Thomas (in diesem Band), in dem wissenschaftstheoretische Aspekte mathematischen Verständnisses untersucht werden.

2. Methode

2.1 Probanden

Es nahmen insgesamt 35 Schüler aus zwei vierten Klassen ($M = 10,5$ Jahre) der selben Grundschule aus dem Würzburger Umland teil. Die Trainingsgruppe (TG) umfasste je 10 Jungen und 10 Mädchen, die Kontrollgruppe (KG) 11 Mädchen und 4 Jungen.

2.2 Design und Durchführung

Bei der berichteten Studie handelt es sich um eine Trainingsstudie im Prä-Posttest-design mit Kontrollgruppe. In beiden Gruppen wurden im Vor- und Nachtest ein Interview zum Wissenschaftsverständnis sowie eine Experimentieraufgabe durchgeführt (siehe auch 2.3 Instrumente). Zwischen den beiden Befragungen lagen $2\frac{1}{2}$ Wochen Unterrichtszeit und 2 Wochen Ferien. Die Trainingsgruppe erhielt im Sachunterricht eine spezielle Unterrichtseinheit über Wissenschaftsverständnis (siehe hierzu Abschnitt 2.5, Abbildung 1), während die Kontrollgruppe in dieser Zeit von ihrem Klassenlehrer entsprechend dem Lehrplan dieser Jahrgangstufe unterrichtet wurde, d.h. ohne eine gezielte Thematisierung von Wissenschaftsverständnis. Die Vor- und Nacherhebung für jedes Kind einzeln in einem ruhigen Raum von einem erwachsenen Interviewer durchgeführt. Die Befragung dauerte pro Schüler etwa 30 Minuten. Sie wurde auf Tonband aufgezeichnet und anschließend transkribiert.

2.3 Instrumente

Das Nature-of-Science-Interview (NoS)

Das Nature of Science Interview (in Anlehnung an Carey u.a. 1989) erfragt Überzeugungen in Bezug auf Ziele und Fragen der Wissenschaften, sowie die Bedeutung zentraler Begriffe wie „Experiment“, „Hypothese“ und „Theorie“ und deren Funktion im wissenschaftlichen Forschungsprozess. Im Fokus der hier referierten Studie stehen jedoch die Interviewteile zu Schülervorstellungen über Ziele und Fragen der Wissenschaft und der Wissenschaftler und die Relationen zwischen den oben aufgeführten relevanten Begriffen. (Das vollständige Interview kann von den Autoren angefordert werden.) Das Transkript des Interviews wurde anschließend von zwei unabhängigen Kodierern ausgewertet (siehe Abschnitt „Kodiersystem Nature-of-Science-Interview“, S. 198).

Die Experimentieraufgabe

Bei der verwendeten Experimentieraufgabe handelt es sich um die so genannte „Flugzeugaufgabe“ nach Bullock/Ziegler (1999). Die Flugzeugaufgabe erfasst das Verständnis der Logik des kontrollierten Experiments an einem konstruierten Beispiel. Die Aufgabe ist eingebettet in einen Geschichtenkontext: Ein Ingenieur soll möglichst sparsame

Flugzeuge konstruieren und kann hierzu drei Variablen (Form der Flugzeugnase, Flügel, Position des Höhenruders) mit je zwei Ausprägungen (Nase spitz vs. rund, einfache Flügel vs. Doppeldeckerflügel, hohe vs. tiefe Position des Höhenruders) variieren. Die Probanden wurden aufgefordert, eine Methode zur Prüfung der Effekte einer vorgegebenen Variable (Position des Höhenruders) auf den Treibstoffverbrauch vorzuschlagen. Zunächst wurden die spontanen Vorschläge erhoben, dann zwei Aufgaben mit unterstützenden Hinweisen vorgelegt: In der „Folien“-Aufgabe konnten die Kinder durch Kombination von Flugzeugteilen zur Prüfung der vorgegebenen Fragestellung geeignete Flugzeuge bauen. In der Wahlaufgabe (Choice) wurde eine Serie von acht Flugzeugen vorgelegt und gefragt, welche der Ingenieur bauen sollte, um den Einfluss der Höhenruderposition auf den Treibstoffverbrauch zu prüfen.

2.4 Auswertung

Kodiersystem Nature-of-Science-Interview

Die Interviewtranskripte wurden von zwei unabhängigen Kodierern qualitativ ausgewertet. Es wurde zunächst jede Frage isoliert betrachtet, anschließend Scores für Themenkomplexe durch Zusammenziehen der Einzelfragen gebildet. Das Kodiersystem wurde in Anlehnung an die von von Carey u.a. (1989) und Smith u.a. (2000) theoriegeleitet erstellten Systeme an die jüngere Stichprobe angepasst. Es unterscheidet zwischen folgenden Ebenen von Wissenschaftsverständnis: Auf Ebene (1a) lassen die Antworten der Personen keine klare Trennung zwischen Realität und Repräsentation (also Welt und Wissen) erkennen, entsprechend wird Wissenschaft rein aktional bzw. als Produktion positiver Effekte gesehen. Auf der nächsten Ebene (1b) wird eine Trennung zwischen Welt und Wissen erkennbar: Wir wissen noch nicht genug über die Welt. Im Wesentlichen besteht wissenschaftliches Arbeiten im Sammeln direkt zugänglicher Fakten; wissenschaftliches Arbeiten zeichnet sich in erster Linie durch die Verwendung hoch entwickelter, exakter Messgeräte aus. Ab der Ebene 1.5 beginnen die Personen zumindest implizit zu erkennen, dass Wissenschaften das Ziel haben, durch die Suche nach Zusammenhängen beobachtbare Phänomene zu erklären. Wissenschaft wird als Prozess der aktiven Konstruktion von Wissen und nicht mehr als reine Informationsaufnahme verstanden. Auf Ebene (2) beginnen Probanden explizit, Wissenschaft als Suche nach Erklärungen zu verstehen und erkennen die Bedeutung der Überprüfbarkeit wissenschaftlicher Annahmen. Auf Ebene (3) schließlich wird zusätzlich noch der zyklisch-kumulative Charakter wissenschaftlichen Erkenntnisfortschritts erkannt. Aufgrund der geringen Auftretenshäufigkeit wurde in der vorliegenden Stichprobe nicht weiter differenziert, inwieweit die sehr jungen Kinder die Aspekte entsprechend den Levels 2 und 3 explizit ausführten, sondern als „Level 1.5 und höher“ zusammengefasst. Antwortete ein Proband auf eine Frage, sodass die Antworten mehrere Kodierungen zuließen, wurde das höchste erreichte Niveau gewertet⁴.

4 Materialien und Kodiersysteme können von den Autoren angefordert werden.

Das Kodiersystem für die Experimentieraufgabe

Das verwendete Kodiersystem folgt der von Bullock/Ziegler (1997) entwickelten Version. Dabei wird zwischen spontaner Antwort, Folientest, und Choice-Test unterschieden. Es wird zwischen vier Stufen differenziert: keine oder nicht kodierbare Antwort; konfundiertes Experiment (alle Dimensionen variiert, ohne die fokale zu beachten); kontrastiver Test (Fokus auf die entscheidende Dimension, die weiteren werden nicht beachtet); kontrollierter Test (Fokus auf die entscheidende Dimension, weitere werden konstant gehalten).

2.5 Trainingsintervention: Die Unterrichtseinheit

Für die Intervention wurde die Unterrichtseinheit „Warum geht der Brotteig auf? – Wir arbeiten wie Wissenschaftler“ für die 4. Jahrgangsstufe entwickelt (s. Abb. 1), bestehend aus fünf (z.T. Doppel-)Stunden; die Unterrichtseinheit orientierte sich an dem von Carey u.a. 1989) entwickelten „Hefe-Curriculum“. Am Beispiel der Frage „Warum geht der Brotteig auf?“ wurden wissenschaftstheoretische Aspekte besonders thematisiert und vertieft: Es wurden Vermutungen über die Gründe für das Aufgehen des Brotteigs gesucht sowie Hypothesen über die Natur von Hefe (Chemikalie oder Lebewesen) aufgestellt, experimentell systematisch geprüft und aufgrund der gefundenen Ergebnisse wiederum weiterführende Hypothesen aufgestellt. Kurz: Die Kinder arbeiteten „wie Wissenschaftler“. Dabei wurde zum einen die Logik des Experimentierens, zum anderen auch der kumulativ-zyklische Charakter wissenschaftlichen Arbeitens verdeutlicht.

3. Ergebnisse

3.1 Nature of Science Interview (NoS)

Wir beschränken uns hier auf die Darstellung der Ergebnisse der Fragenkomplexe zu Gegenstand, Zielen und Fragen der Wissenschaft sowie Experimenten und Hypothesen. Es werden zunächst die Antworten der Kontroll- und Trainingsgruppe zu den Messzeitpunkten (Vor- und Nachtest) verglichen, im zweiten Schritt die Veränderungen innerhalb der Gruppen über die beiden Messzeitpunkte. Wie oben erwähnt, wurden die Interviews von zwei unabhängigen Kodierern ausgewertet; die Interrater-Reliabilität betrug übereinstimmend für beide Messzeitpunkte $Kappa = .79$ und war damit sehr zufriedenstellend. Nicht-Übereinstimmungen wurden durch Diskussionen gelöst.

Im Vortest war der Großteil der Antworten bezüglich „Ziele und Fragen der Wissenschaft“ aus beiden Gruppen in Kategorie 1b (Wissenschaft als Faktensammlung) einzuordnen (TG: 60%, KG: 52%). Es zeigten sich keine Rangdifferenzen zwischen Kontroll- und Trainingsgruppe. Auch auf die Fragen zu Experimenten und Hypothesen antwortete über die Hälfte beider Gruppen auf Niveau 1a und 1b, wobei der Anteil von Antworten auf Niveau 1a bei der KG nicht-signifikant höher als bei der Trainingsgruppe war

Tag	Lernziele	Inhalte und Methode	Medien
1.Tag 60min	LZ: Kennen lernen der Fachbegriffe (siehe Wortkarten*)	<ul style="list-style-type: none"> - Warum geht der Brotteig auf? - Finden der Grundzutaten - SDV (Schülerdemonstrationsversuch): Mischen der Zutaten und Beobachtung der Gasentwicklung - Reflexion 	roher und gebackener Brotteig Experimentiermaterialien (Zutaten: Mehl, Hefe, Salz, Zucker, Wasser) div. Wortkarten*
2.Tag 45min	LZ: Aufstellen einer sinnvollen Versuchsplanung (Variablenisolation)	<ul style="list-style-type: none"> - Welche der Grundzutaten sind tatsächlich notwendig, damit Gasblasen entstehen? - ggf. Durchführung eines konfundierten Experimentes- gemeinsames Erstellen des Versuchsplanes - Reflexion, - Transferbeispiel (AB) 	Experimentiermaterialien  AB (Arbeitsblatt)
3.Tag 90min	1.LZ: Korrekte Umsetzung des Versuchsplanes in arbeitsteiliger Gruppenarbeit (GA) 2.LZ: Erkenntnis der Bedeutung von Hypothesen, Evidenzen und daraus resultierenden Schlussfolgerungen	Schüler experimentieren in GA ➔ <i>Ohne Wasser und Hefe entstehen keine Gasbläschen im Brotteig.</i> Erweiterung der Experimentfolge durch neuen Versuchsplan und SDV ➔ <i>Hefe und Wasser bilden zusammen mit Mehl oder Zucker Gasblasen.</i> Reflexion	Versuchsplan 1 (AB) Experimentiermaterialien (s.o.) Versuchsplan 2 (AB) Experimentiermaterialien (s.o.)
4.Tag 90min	1.LZ: Suche nach möglichen Erklärungen für beobachtetes Phänomen/ Definition von Lebewesen. 2.LZ: Angehendes Verständnis des Theorie-Begriffes 3.LZ: Einsicht, dass wissenschaftlicher Streit nicht immer sofort entschieden werden kann. 4.LZ: Angehendes Verständnis, dass Evidenzen zu verschiedenen Theorien passen können.	<i>Warum entstehen Gasblasen?</i> Zwei konkurrierende Theorien (Rollenspiel der Lehrerin als Wissenschaftlerin): 1. Hefe ist ein Lebewesen . Sie produziert ein Gas als „Abfallprodukt“. ➔ DV: menschlicher Atem und „Hefegas“ trüben Kalkwasser (Nachweisreaktion für CO ₂) 2. Hefe ist kein Lebewesen . Die Gasblasen entstehen durch eine chemische Reaktion. ➔ DV: Zutatengemisch mit Backpulver statt Hefe (als Chemikalie) lässt ebenfalls CO ₂ entstehen.	Verschiedene Kleidung zur Darstellung verschiedener Wissenschaftlerinnen „Kalkwasser“ + bisherige Experimentiermaterialien Backpulver + bisherige Experimentiermaterialien
5.Tag 60min	1.LZ: Erkenntnis der Vorläufigkeit von Theorien. 2.LZ: Erkennen der Bedeutung von Theorie und Experiment für die Arbeit des Wissenschaftlers.	<ul style="list-style-type: none"> - historisches Beispiel: <i>Ist die Erde eine Scheibe?</i> - Vertiefung des „Hefeproblems“: - Welche entscheidenden Experimente könnten weiterhelfen? - Rückblick auf die Unterrichtseinheit: Nachdenken über Arbeitsweisen, Möglichkeiten und Grenzen der Wissenschaften und Wissenschaftler 	Bilder zur Visualisierung der damaligen Vorstellungen und heutige Aufnahmen von der Erde

*Im Unterricht verwendete Wortkarten:
 „Feststellung/ Beobachtung einer Tatsache“, „Frage“, „wissenschaftliche Vermutung (Hypothese)“, „wissenschaftlicher Versuch (Experiment)“, „Ergebnis“, „Schlussfolgerung“.

Abb. 1: Unterrichtseinheit: „Warum geht der Brotteig auf? Wir forschen wie Wissenschaftler“

(TG: 29%, KG: 42%). Im Nachtest antworten jedoch nur noch deutlich weniger Kinder der Trainingsgruppe (12,5%) auf diesem niedrigen Niveau als bei der Kontrollgruppe (43%; Mann-Whitney-U = 78.5; $z = -2.24$, $N = 33$, $p = .057$).

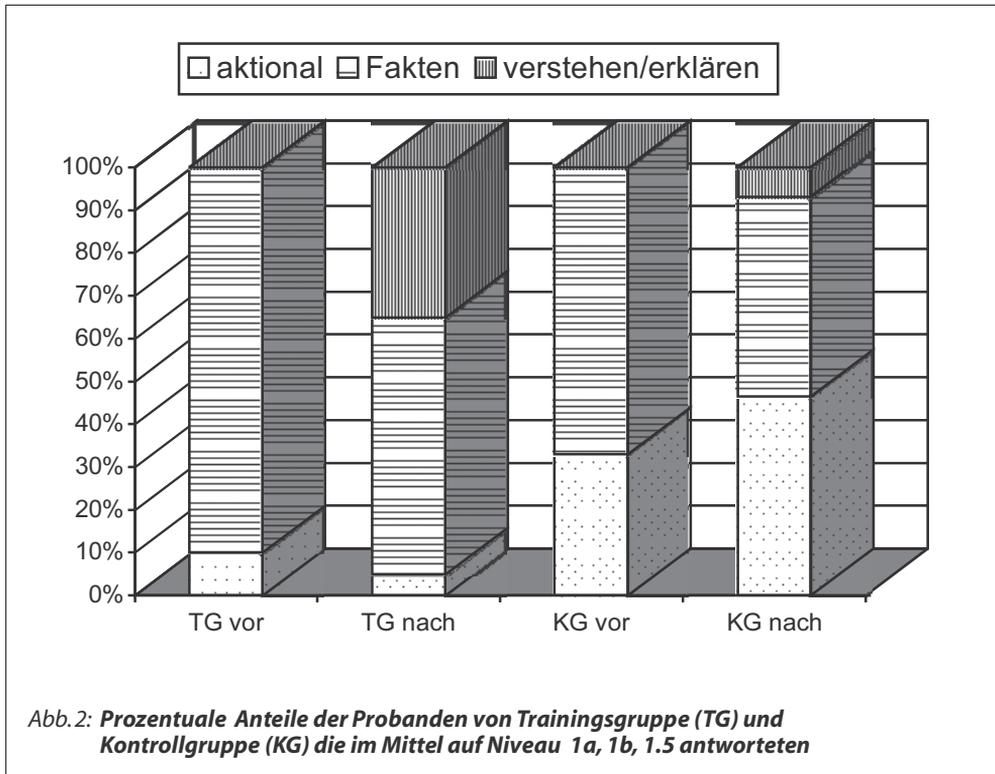


Abb. 2 zeigt den Anteil der Probanden (getrennt nach Gruppe und Messzeitpunkt), die im Mittel auf den verschiedenen Niveaus antworteten: Während die überwältigende Mehrheit beider Gruppen im Vortest auf Niveau 1b (Wissenschaft als Faktensammlung) antwortete, erreichte ein substantieller Anteil der Probanden der Trainingsgruppe, jedoch nicht der Kontrollgruppe, im Nachtest im Mittel Niveau 1.5 (TG: $\text{Chi}^2(20, 2) = 9,81$, $p = .007$; KG: $\text{Chi}^2(15, 2) = 8,6$, n.s.). Die Schüler der Trainingsgruppe zeigten also ein zumindest beginnendes Verständnis von Wissenschaft im Sinne einer Suche nach Erklärungen und Mechanismen. Analysiert man etwas feiner, nämlich getrennt nach „Zielen der Wissenschaft“, „Fragen der Wissenschaft“, „Experimenten in der Wissenschaft“ und „Hypothesen“, so zeigt sich im Vortest kein Unterschied, jedoch gibt es im Nachtest signifikante Unterschiede in Bezug auf „Ziele der Wissenschaft“ (Mann-Whitney $N = 34$; $p = .043$) und „Experimente“ (Mann-Whitney, $N = 34$, $p = .006$) dahingehend, dass die Trainingsgruppe im Nachtest auf höherem Niveau antwortet als die Kontrollgruppe.

Betrachtet man die Veränderungen innerhalb der Gruppen über die Messzeitpunkte, zeigt sich (siehe Abb. 2) dass in der Kontrollgruppe keine Verbesserung stattfand, wäh-

rend in der Trainingsgruppe im Nachtest weniger Schüler auf Niveau 1a und 1b, dafür deutlich mehr Schüler auf Niveau 1.5 antworteten. Auf der Ebene einzelner Fragen stellt sich heraus, dass vor allem die Fragen nach Experimenten und Hypothesen („Was ist ein Experiment? Machen Wissenschaftler Experimente? Wie kommt ein Wissenschaftler dazu, ein bestimmtes Experiment zu machen?“, „Was ist eine Hypothese?“, „Wie kommen Wissenschaftler zu ihren Vermutungen?“) von der Trainings-, jedoch nicht von der Kontrollgruppe im Nachtest bedeutend besser beantwortet wurden als im Vortest (TG: Wilcoxon-Test $z = -2.35$, $p = .019$; KG: Wilcoxon-Test $z = -.82$, n.s.), was auch vom Intragruppenvergleich über die Messzeitpunkte bestätigt wird (Wilcoxon-Test, $n = 20$, $z = -2.87$, $p = .004$; McNemar-Test $p = .004$). Wertet man hier jedoch wiederum getrennt nach dem Verständnis von Zielen, Fragen, Experimenten und Hypothesen aus, zeigt sich, dass dieser Effekt vorwiegend auf ein tieferes Verständnis von Experimenten im Sinne der gezielten Überprüfung von Hypothesen über Zusammenhänge zurückgeht (Wilcoxon-Test, $n = 20$, $z = -2.35$, $p = .019$): Während im Vortest nur 25% der TG auf Niveau 1.5 antworteten, taten dies im Nachtest 70%, in der Kontrollgruppe veränderte sich der Anteil nur von 13% auf 20%.

3.2 Experimentieraufgabe

Tab. 1: Prozentuale Anteile der Probanden, die bei den Unteraufgaben der Flugzeugaufgabe ein kontrastives oder kontrolliertes Experiment vorschlugen bzw. auswählten

Gruppe (Vor-Nachtest)	Spontan	Folie	Wahl
TG vor	50%	40%	35%
KG vor	43%	35%	43%
TG nach	70%	50%	65%
KG nach	50%	22%	43%

Tabelle 1 zeigt die Anteile der beiden Gruppen getrennt nach Messzeitpunkt die spontan, bei der Folienaufgabe bzw. bei der Wahlaufgabe ein mindestens kontrastives Experiment vorschlugen oder auswählten. Während sich im Vortest für keine Aufgabe Gruppenunterschiede zeigten, produzierten im Nachtest deutlich mehr Kinder der Trainingsgruppe als der Kontrollgruppe in der Folienaufgabe ein kontrastives oder kontrolliertes Experiment (Mann-Whitney-U = 91.0, $N = 34$, $z = -1.94$, $p = .053$). Jedoch war die Verbesserung innerhalb der Trainingsgruppe in dieser Aufgabe vom Vor- zum Nachtest nicht signifikant. Auch bei der Wahl-Aufgabe wählte die Trainingsgruppe im Nachtest beim Choice-Test deutlich mehr kontrastive oder kontrollierte Experimente als im Vortest (McNemar: $N = 20$, $df = 1$, $p = .031$). Das Abschneiden der Kontrollgruppe veränderte sich hingegen in keinem Aufgabenteil signifikant.

3.3 Zusammenhänge zwischen Interview und Experimentieraufgabe

Es werden Rangkorrelationen (Spearman-Rho) für den mittleren Level im Interview und der Experimentieraufgabe (nur spontaner Test, Folien- und Wahlaufgabe) berichtet: Während im Vortest kein Zusammenhang besteht, zeigt sich für die Trainingsgruppe ein Zusammenhang ($r_s = .49$, $p = .029$) im Nachtest.

Aufgrund der Ergebnisse für die beiden Instrumente (NoS: Verbesserung des Verständnisses von Experimenten in der Wissenschaft; Experimentieraufgabe: Verbesserung in der Wahlaufgabe), erscheint es angebracht, diese beiden Aufgabenteile gesondert zu betrachten. Tatsächlich zeigt sich für die Trainingsgruppe sowohl im Vortest ($r_s = .46$, $p = .043$) wie auch im Nachtest ($r_s = .41$, $p = .074$) ein Zusammenhang zwischen dem Verständnis von Experimenten und der Wahlaufgabe. Tendenziell ging ein Verständnis des Experimentierens im Sinne des Testens von Hypothesen mit der Auswahl kontrastiver oder kontrollierter Tests einher, während ein Verständnis des Experiments im Sinne von „ausprobieren“ mit der Auswahl konfundierter Tests korrelierte.

4. Diskussion

Hauptziel der vorliegenden Studie war die Exploration des intuitiven Wissenschaftsverständnisses von Grundschulern der vierten Jahrgangsstufe sowie die Untersuchung von Effekten einer kurzfristigen curricularen Intervention. Ausgehend von den Befunden der Vorläuferstudie von Carey u.a. (1989) an Siebtklässlern erwarteten wir ein überwiegend aktionales Wissenschaftsverständnis (Ebene 1a) bei Viertklässlern (Verständnis von Wissenschaft im Sinne konkreter Aktivitäten, insbesondere der Produktion von Effekten). Die Befunde der vorliegenden Studie zeigen ein solches aktionales Wissenschaftsverständnis nur bei einer Minderheit von ca. 20% der Kinder. Die überwiegende Mehrheit der Grundschüler antwortete im Vortest auf Ebene 1b, d.h., sie verstanden Wissenschaft im Sinne des Sammelns von konkreter faktischer Information. Beide Ebenen (1a und 1b) sind durch das *Fehlen* eines auch nur rudimentären Verständnisses des Theorie-Evidenz-Zusammenhangs gekennzeichnet: Experimente werden als konkrete Aktivitäten ohne Bezug zu einer Theorie, Idee oder Vermutung verstanden, das Sammeln von faktischer Information erscheint ebenfalls als konkretes Ziel wissenschaftlicher Arbeit, das nicht in Bezug zur Erklärung von Phänomenen oder zur Prüfung von Hypothesen gebracht wird. Diese Interpretation der Defizite im Wissenschaftsverständnis der Grundschüler basiert nicht auf Missverständnissen der Interviewfragen oder einem hohen Anteil nicht kodierbarer bzw. fehlender Antworten; vielmehr ließ sich die weit überwiegende Mehrheit der Antworten reliabel Ebene 1b zuordnen.

Ebenso wie Carey u.a. (1989) bei Siebtklässlern erzielten auch wir – schon in der vierten Grundschulklasse – durch eine Unterrichtseinheit mit wissenschaftstheoretischen Aspekten einen im Prä-Posttestvergleich nachweisbaren Trainingseffekt, der sich gegen einen Testwiederholungseffekt absichern ließ. Der Anteil der Probanden der Trainingsklasse, die zumindest implizit ein Verständnis des *Testens* von Theorien und Hypo-

thesen erkennen ließen, und Wissenschaft im Sinne der Suche nach Erklärungen für Phänomene verstanden, stieg deutlich an: der Anteil von Niveau 1.5, stieg von ca 10% im Vortest auf 35–55% im Nachtest. Dabei zeigte sich ein besonders ausgeprägter Effekt des Unterrichts auf das Verständnis der Logik des Testens (weniger auf das Verständnis des Zusammenhangs zwischen Hypothesen und Experimenten). In der Kontrollklasse zeigte sich keine Veränderung gegenüber dem Ausgangsniveau. Der Trainingseffekt war nicht beschränkt auf das abstrakt-definitorische Verständnis der Logik der Hypothesenprüfung, sondern die Trainingsklasse verbesserte sich auch in der Wahl eines kontrollierten Experiments – in einer Aufgabe zum Verständnis des Experimentierens aus einer artifiziellen Domäne, die nicht Gegenstand des Unterrichts war. Wenn sich auch keine Effekte auf die Fähigkeit zur ungestützten (spontanen) Produktion eines kontrollierten Experiments zeigten, so spricht die Verbesserung der Trainingsklasse in der Wahlaufgabe doch für eine Generalisierung des *Verständnisses* des Hypothesentestens, das im aktiven Handeln (Durchführung geeigneter Tests) noch nicht voll umgesetzt werden kann. Für einen funktionalen Zusammenhang zwischen Wissenschaftsverständnis (mindestens Ebene 1.5) und Wahl eines geeigneten Experiments spricht auch die signifikante Korrelation zwischen Interviewscores und Abschneiden in der „Flugzeugaufgabe“ in der Trainingsgruppe im Nachtest. Insgesamt deuten die vorliegenden Befunde darauf hin, dass das Wissenschaftsverständnis von älteren Grundschulern durch kurzfristige curriculare Intervention beeinflussbar ist. Insbesondere scheint durch den Unterricht im Prozess der Konstruktion wissenschaftlichen Wissens ein rudimentäres Verständnis der Rolle empirischer Tests vermittelt werden zu können, das inhaltsunabhängig auf wissenschaftliche Problemstellungen angewandt wird. Der Befund, dass Grundschüler Experimentierstrategien (Variablenisolation und -kontrolle) lernen können, ist nicht neu; vielmehr haben seit Case (1975) mehrere experimentelle Strategietrainings stabile Effekte auf die Experimentierfähigkeiten von Grundschulern erbracht (Chen/Klahr 1999). Neu ist vielmehr, dass eine wissenschaftstheoretische Unterrichtseinheit, die kein spezifisches Strategietraining enthält, bereits im Grundschulalter inhaltsunabhängige Effekte auf das Verständnis des Experimentierens zeigt.

Wie erwartet, waren die Effekte unserer an fünf Tagen durchgeführten Unterrichtseinheit begrenzt. Nur etwa die Hälfte der Trainingsklasse profitierte von diesem Unterricht; dabei lag die Verbesserung bei durchschnittlich einem halben Niveau. Wenn das skizzierte Trainingsparadigma, wie im vorliegenden Projekt geplant, als Forschungsinstrument eingesetzt werden soll, dann sollte ein Trainingseffekt – im Sinne eines veränderten Wissenschaftsverständnisses – bei der Mehrheit (nicht lediglich der Hälfte) der unterrichteten Kinder angestrebt werden. Ferner sollte der Trainingseffekt nicht begrenzt bleiben auf den engeren Bereich der Logik der Hypothesenprüfung. Dies soll in Folgestudien durch die Vertiefung erkenntnistheoretischer Aspekte geschehen.

Literatur

- Baumert, J. u.a. (2000): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Leske & Budrich: Opladen.
- Brown, A.L./Campione, J.C. (1994): Guided discovery in a community of learners. In: McGilly, K. (Ed.) Classroom lessons: Integrating cognitive theory and educational practice. Cambridge, MA: MIT Press, S. 229–270.
- Bullock, M./Ziegler, A. (1999): Scientific reasoning: Developmental and individual differences. In: Weinert F.E./Schneider, W. (Eds.) Individual development from 3 to 12. Findings from the Munich Longitudinal Study. Cambridge: Cambridge University Press, S. 38–44.
- Carey, S. (1986): Cognitive Science and Science Education. In: American Psychologist, 41, S. 1123–1130.
- Carey, S./Evans, R./Honda, M./Jay, E./Unger, C. (1989): An experiment is when you try it and see if it works. A study of junior high school students' understanding of the construction of scientific knowledge. In: International Journal of Science Education, 11, S. 514–529.
- Carey, S./Smith, C. (1993): On understanding the nature of scientific knowledge. In: Educational Psychologist, 28, S. 235–251.
- Chi, M. T. H./Slotta, J. D./de Leeuw, N. (1994): From things to processes: A theory of conceptual change for learning science concepts. In: Learning and Instruction, 4, S. 27–43.
- Dewey, J. (1916/1964): Demokratie und Erziehung. Braunschweig.
- Driver, R. u.a. (1996): Young peoples images of science. Bristol: Open University Press.
- Case, R. (1974): Structures and strictures: Some functional limitations on the course of cognitive growth. In: Cognitive Psychology, 6, S. 544–573.
- Chen, Z./Klahr, D. (1999): All other things being equal: children's acquisition of the control of variables strategy. In: Child Development, 70, S. 1098–1120.
- Chi, M. T. H./Slotta, J. D./de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. In: Learning and Instruction, 4, S. 27–43.
- Chinn, C.A./Brewer, W.F. (1993): The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. In: Review of Educational Research, 63, S. 1–49.
- Duit, R. (2002): Alltagsvorstellungen und Physiklernen. In: Kircher, E./Schneider, W. B. (Hrsg.). Physikdidaktik in der Praxis. Heidelberg Berlin: Springer.
- Gallagher, J.J. (1991): Prospective and practising secondary school science teachers' knowledge and beliefs about the philosophy of science. In: Science Education, 75, S. 121–133.
- Höttecke, D. (2001). Die Natur der Naturwissenschaften historisch verstehen. Berlin: Logos Verlag.
- Honda, M. (1994): Linguistic inquiry in the science classroom: It is science, but it's not like a science problem in a book. Cambridge, MA: MIT Working Papers in Linguistics.
- Kircher, E. (1993): Warum ist Physiklernen schwierig?. In: Schneider, W.B. (Hrsg.). Wege in der Physikdidaktik. Erlangen: Palm & Enke, S. 124–134.
- Kircher, E. (1995): Studien zur Physikdidaktik – Erkenntnis- und wissenschaftstheoretische Grundlagen. Kiel: IPN.
- Leach, J. (1999): Students' understanding of the co-ordination of theory and evidence in science. In: International Journal of Science Education, 21, S. 789–806.
- Lederman, N.G. (1992): Students' and teachers' conceptions of the nature of science: A Review of the Research. In: Journal of Research of Science Teaching, 29, H.4, S. 331–359.
- Litt, T. (1959): Naturwissenschaft und Menschenbildung. Heidelberg: Quelle & Meyer.
- Mc Comas, W.F. (Ed.) (1998): The Nature of Science in Science Education. Dordrecht: Kluwer.
- Mc Comas, W.F./Mc Clough, M./Almaroza, H. (1998): The role and character of the nature of science. In: Mc Comas, W.F. (Ed.) The Nature of Science in Science Education. Dordrecht: Kluwer Academic Publishers, S. 3–39.

- McComas, W.F./Olson, J.K. (1998): The nature of science in international science education standards documents. In: McComas, W.F. (Ed.). *The Nature of Science in Science Education*. Dordrecht: Kluwer, S. 41–52.
- Meyling, H. (1997): How to change students' conceptions of the epistemology of science. In: *Science & Education*, 6, S. 397–416.
- Pomeroy, D. (1993): Implications of teachers' beliefs about the nature of science: Comparisons of the beliefs of scientists, secondary science teachers and elementary teachers. In: *Science Education*, 77 (3), S. 261–278.
- Ryan, A.G./Aikenhead, G.S. (1992): Students' preconceptions about the epistemology of science. In: *Science Education*, 76, S. 559–580.
- Samarapungavan, A. (1992): Children's judgments in theory choice tasks: Scientific rationality in childhood. In: *Cognition*, 45, S. 1–32.
- Schecker, H. (1985): *Das Schülervorverständnis zur Mechanik*. Dissertation, Universität Bremen.
- Schrempp, I./Sodian, B. (1999): Wissenschaftliches Denken im Grundschulalter. Die Fähigkeit zur Hypothesenprüfung und Evidenzevaluation im Kontext der Attribution von Leistungsergebnissen. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31, S. 67–77.
- Smith, C./Maclin, D./Houghton, C./Hennessey, M.G. (2000): Sixth grade students' epistemologies of science: the impact of school science experiences on epistemological development. In: *Cognition and Instruction*, 18, S. 349–422.
- Sodian, B./Zaitchik, D./Carey, S. (1991): Young children's differentiation of hypothetical beliefs from evidence. In: *Child Development*, 62, S. 753–766.
- Solomon, J. (1991): Teaching about the nature of science in the British National Curriculum. In: *Science Education* 75, H.1, S. 95–103.
- Thoermer, C./Sodian, B. (im Druck): Science undergraduates' and graduates' epistemologies of science: The notion of interpretive frameworks. In: *New Ideas in Psychology*.
- Vosniadou, S. (1994): Capturing and modelling the process of conceptual change. In: *Learning and Instruction*, 4, S. 45–69.
- White, B.Y./Frederiksen, J.R. (1998): Inquiry, modelling, and metacognition: Making science accessible to all students. In: *Cognition and Instruction*, 16, S. 3–118.

Anschrift der Autoren:

Prof. Dr. Beate Sodian, Lehrstuhl für Entwicklungs- und Pädagogische Psychologie, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München.

Dipl.-Psych. Claudia Thoermer, Lehrstuhl für Entwicklungs- und Pädagogische Psychologie, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München.

PD Dr. Ernst Kircher, Physikalisches Institut/Didaktik der Physik, Universität Würzburg, Am Hubland, 97074 Würzburg.

Lin. Patricia Grygier, Physikalisches Institut/Didaktik der Physik, Universität Würzburg, Am Hubland, 97074 Würzburg.

Dipl.-Phys. Johannes Günther, Physikalisches Institut/Didaktik der Physik, Universität Würzburg, Am Hubland, 97074 Würzburg.

Elke Sumfleth/Elke Wild/Stefan Rumann/Josef Exeler

Wege zur Förderung der naturwissenschaftlichen Grundbildung im Chemieunterricht¹

Kooperatives Problemlösen im schulischen und familialen Kontext zum Themenbereich Säure-Base

1. Einleitung

Die Ergebnisse der TIMS- und PISA-Studien (Baumert u.a. 1997; Baumert/Bos/Watermann 1998; Baumert u.a. 2001) belegen für den mathematisch-naturwissenschaftlichen Unterricht, dass deutsche Schüler besondere Schwierigkeiten mit Aufgaben haben, die konzeptuelles Verständnis und flexibel anwendbares Wissen erfordern. Eine Ursache hierfür wird im vorherrschenden fragend-entwickelnden Unterrichtsgespräch gesehen, das von den Lehrenden über enge Fragen konvergent gesteuert wird. Eine solche direktive Instruktion scheint nicht nur den Aufbau einer transferfähigen Wissensstruktur, sondern auch die Herausbildung und Aufrechterhaltung eines fachspezifischen Interesses und selbstbestimmter Formen der Lernmotivation zu verhindern. Dies zeigen nicht zuletzt Interessenstudien in den Naturwissenschaften (Hoffmann/Häußler/Lehrke 1998; Gräber 1992; Duit/Häußler 1997).

Eine Alternative sind Unterrichtsansätze, die selbstorganisierte und kooperative Lernprozesse betonen. Diese erscheinen wirksam zum Erreichen affektiv-motivationaler Ziele, gelten jedoch als weniger effizient zur Wissensvermittlung (Klieme/Clausen 1999). Analysen von „Positivklassen“ (Helmke/Schrader 1990), in denen kognitive und motivationale Ziele in überdurchschnittlichem Maße erreicht werden, zeigen, dass diese multikriteriale Zielerreichung vor allem eine Folge der effektiven Nutzung von Unterrichtszeit ist und der Fähigkeit der Lehrenden, Handlungsspielräume für die individuelle Berücksichtigung von unterschiedlichen Leistungsvoraussetzungen zu schaffen. Solche ex post Analysen geben aufgrund der atheoretischen Bestimmung der Einflussfaktoren nur bedingt Aufschluss über notwendige und hinreichende instruktionale Maßnahmen.

In diesem Projekt stehen die kognitiven und motivationalen Entwicklungsverläufe im Vordergrund, die im Unterricht durch Lehrer-Schüler- und Schüler-Schüler-Interaktionen sowie im familialen Kontext durch das Interesse der Eltern und deren Hausaufgabenunterstützung angestoßen werden. Es wird davon ausgegangen, dass eine Steigerung des Schülerinteresses und ein tiefergehendes Verständnis chemischer Probleme durch ein problemorientiertes, kooperatives, selbstbestimmtes Arbeiten der Schüler er-

¹ Die Studie wurde gefördert durch Mittel der DFG (Aktenzeichen SU 187/5-1 & WI 1607/3-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

zielt werden kann, wobei das richtige Verhältnis von Offenheit und Strukturiertheit entscheidend ist. Dieses wird durch die in den Lernmaterialien enthaltenen instruktionalen Hilfen ausbalanciert.

Vor dem Hintergrund des im Einleitungskapitel zu diesem Heft dargestellten mehrbenenanalytischen Rahmenmodells werden in diesem Projekt die Wirkungen und Wechselwirkungen von Familienmerkmalen (Arbeitsalltag, Hausaufgaben, Elterninteresse, Elternunterstützung, Geschwister) als außerschulischem Kontext und Unterrichtsmerkmalen (Lerngelegenheiten, Quantität und Qualität der Instruktion, Lehrer-Schüler-Beziehung, Schüler-Schüler-Beziehung) als schulischem Kontext analysiert, wobei die Bedeutung dieser Merkmale am Erreichen fachlicher und fächerübergreifender Zielkriterien mathematisch-naturwissenschaftlicher Bildung gemessen wird. Neben der themenspezifischen Kompetenz, operationalisiert durch die erworbenen Fachkenntnisse und wissenschaftlich angemessenen Vorstellungen sowie durch die themenspezifische Problemlösekompetenz und Argumentationsfähigkeit ist hiermit auch die Fähigkeit und Bereitschaft zur selbstregulierten Auseinandersetzung mit naturwissenschaftlichen Inhalten angesprochen. Insgesamt werden insofern *fachspezifische und fächerübergreifende Aspekte* betrachtet.

2. Theoretische Grundlagen

2.1 Kognitive und motivationale Vorteile eines kooperativen problemorientierten Lernens im Unterricht

Unterrichtsanalysen im Rahmen der TIMS-Videostudie lassen einen Zusammenhang zwischen den Spitzenleistungen japanischer Schüler und dem aktiven situierten problemorientierten Lernen in schulischen wie außerschulischen Kontexten vermuten (Klieme/Clausen 1999; Schümer 1998). Dies korrespondiert mit der Forderung nach kooperativen Lehr-Lernarrangements und multiplen Perspektiven und Kontexten (Reinmann-Rothmeier/Mandl 1997), um die Ausbildung „trägen Wissens“ zu reduzieren. Die Kritik am herkömmlichen Unterricht setzt entsprechend an der Dominanz des Lehrers im Unterrichtsgespräch und der mangelnden Berücksichtigung vorhandener Schülervorstellungen an. Lernende sollten dann wissenschaftlich angemessene Vorstellungen entwickeln, wenn ihnen selbstständige Problemlöseleistungen abverlangt werden. Kooperation mit anderen Lernenden unterstützt durch die sprachliche Auseinandersetzung mit anderen die kritische Reflektion der eigenen Vorstellungen. Entscheidend hierfür sind geeignete Denkanstöße und genügend Zeit für kreative Denkprozesse.

Strittig ist das optimale Verhältnis von Selbst- zu Fremdsteuerung, da Lernende in komplexen und offenen Lernumgebungen wiederholt schlechtere Leistungen erzielten als in herkömmlichen Lehr-Lernarrangements (Gräsel 1997). Erklärt wird die geringere Effektivität mit einer Überforderung der Lernenden, die dann dazu neigen, zu lösende Probleme nur oberflächlich zu explorieren, nur wenige Hypothesen zu bilden und diese nicht systematisch zu prüfen (z.B. Stark u.a 1995; Fischer u.a. 1997; Leutner 1992). Der

Nutzen instruktionaler Hilfen scheint umso geringer, je mehr diese die kognitive Belastung erhöhen (Sweller 1994).

Vor allem Arbeiten zum „conceptual change“ zeigen, dass entscheidende Denkanstöße aus der Interaktion mit einem kompetenteren Partner („guided participation“) resultieren können (Rogoff 1990; Krappmann 1994), andererseits aber auch wahrscheinlich sind, wenn Heranwachsende mit Gleichaltrigen sozial induzierte kognitive Konflikte bearbeiten. Vermutlich verleitet die anerkannte Überlegenheit des Lehrers oder anderer erwachsener Bezugspersonen leicht zur unreflektierten Anpassung statt zum verstärkten Nachdenken. Andererseits laufen Lernprozesse in Gruppen aufgrund mangelnder Selbstregulationskompetenzen oft ineffektiver ab als bei direkter Instruktion (Weinert 1997), da ein erheblicher Teil der Arbeitszeit für aufgabenirrelevante Kommunikation verwendet werden muss.

Betrachtet man kooperative Lehr-Lernarrangements aus motivationspsychologischer Perspektive, dann scheinen sie in mehrfacher Hinsicht anderen Sozialformen überlegen zu sein. Ausgehend von der Selbstbestimmungstheorie (Deci/Ryan 1985; Ryan 1995; Grolnick/Deci/Ryan 1997) und der Interessentheorie (Krapp 1998; 1999; Schiefele 1986; Krapp/Prenzel 1992) trägt eine Befriedigung der psychologischen Grundbedürfnisse nach Autonomie-, Kompetenzerleben und sozialer Einbindung dazu bei, dass eine vorhandene intrinsische Motivation aufrecht erhalten bleibt oder selbstbestimmte Formen der Lernmotivation entwickelt werden, so dass sich Lernende zunehmend „aus eigenen Stücken“ mit Inhalten auseinandersetzen. Vor diesem Hintergrund erscheinen kooperative Lehr-Lernarrangements vorteilhafter, weil sie der Orientierung Heranwachsender an Gleichaltrigen Rechnung tragen und Leistungsbewertungsaspekte einer ausgeprägten Aufgabenorientierung Platz machen können, so dass sich leichter Kompetenzerleben einstellen kann. Darüber hinaus sieht das schülerzentrierte Vorgehen ein höheres Maß an Autonomie vor und trägt dazu bei, Einwände und Widerstände weniger als Kontrolle zu empfinden, da sie im Rahmen der egalitären Interaktion zwischen Lernenden geäußert werden. Anwendung des Gelernten und Transfer sollen durch eine sukzessive Erhöhung der Komplexität der im Unterricht zu bearbeitenden Aufgaben gefördert werden. Zur Unterstützung dieser anspruchsvollen Transferleistungen erscheint es notwendig, (1) durch Überlernen für eine Konsolidierung der zuvor aufgebauten Wissensstrukturen zu sorgen, (2) die Gelegenheit zum Erwerb von Wissen in verschiedenen Anwendungsbereichen und aus verschiedenen Perspektiven zu eröffnen, und (3) Schüler durch Vorgabe isomorpher Aufgaben zunächst zu nahen Transferleistungen hinzuführen und so auf einen weiten Transfer vorzubereiten. Diese Aufgaben kommen der kooperativen Hausaufgabenbearbeitung zu.

2.2 Kooperatives problemorientiertes Arbeiten in Hausaufgaben-situationen

Die Voraussetzungen für eine verstärkte Einbeziehung der Eltern in die Hausaufgabenbearbeitung scheinen günstig zu sein. So wünschen sich etwa 70% der Eltern mehr Hilfen zur Ausgestaltung des häuslichen Lernens und wären bereit, täglich bis zu einer

Stunde dafür zu investieren (Krumm 1996). Auch legen angloamerikanische Studien nahe, dass durch eine verstärkte Einbeziehung der Eltern und eine Intensivierung der Eltern-Lehrer-Kooperation die Leistungen von Lernenden verbessert werden können (Chrispeels/Coleman 1996; Griffith 1996). Gleichwohl weisen vorliegende Analysen der elterlichen Hausaufgabenhilfe (z.B. Krohne/Hock 1994) darauf hin, dass wohl gemeinte Bemühungen von Eltern durchaus kontraproduktiv sein können. So wird verständlich, warum Hausaufgaben zwar grundsätzlich als sinnvoll erachtet, gleichzeitig aber oft als langweilig und belastend erlebt werden (Krumm 1998).

Aus theoretischer Sicht sollte eine sinnvolle Gestaltung der Hausaufgaben das Bemühen von Eltern beinhalten, so wenig Hilfe wie möglich und soviel wie nötig zu geben. Die sich daraus ergebende größere Selbstverantwortung der Schüler für die Hausaufgaben sollte sich zusammen mit der in der Hausaufgabensituation zu realisierenden stärkeren Trennung von Lern- und Leistungssituationen positiv auf das Kompetenzerleben auswirken. Da Schüler mehrheitlich eine sehr enge Beziehung zu den Eltern haben (zusf. Wild/Hofer 2002), sollte es Eltern auch grundsätzlich leichter fallen, beim häuslichen Lernen eine vertrauensvolle Atmosphäre herzustellen. Ein konstruktives Umgehen mit Schwächen und Fehlern ist dabei wiederum entscheidend für die Befriedigung des Kompetenzerlebens. Schließlich sollte das Schülerinteresse auch dann steigen, wenn Eltern ihr eigenes Interesse ausdrücken und damit den Inhalten subjektive, für die Lernenden erkennbare Bedeutung verleihen können.

3. Forschungsfragen

Im Mittelpunkt steht die Evaluation zweier verschränkter Interventionen, einer im schulischen Kontext Chemieunterricht (Klasse 7, Thema Säure-Base) und einer im außerschulischen Kontext häusliches Lernen. Es wird vermutet, dass die Effekte der Intervention im Unterricht bei gleichzeitigem Miteinbezug der Eltern deutlicher ausfallen. Der Erfolg der Interventionen wird einerseits festgemacht an der Veränderung stabiler, aus wissenschaftlicher Sicht aber nicht erklärungs-mächtiger Schülervorstellungen, andererseits in der Förderung selbstbestimmter Formen der Lernmotivation. Unter dem Aspekt der multikriterialen Zielerreichung soll also die Förderung von Lernfreude und Interesse nicht auf Kosten der Leistung erreicht werden, sondern vielmehr mit einem höheren Lernzuwachs assoziiert sein.

Vor allem durch die Kombination von Verhaltensbeobachtungen und erlebnisdeskriptiven Daten ermöglicht die Studie zudem eine kleinschrittige, ökologisch valide und nah an den mikrosozialen Interaktionsprozessen orientierte Analyse des Zusammenhangs von lernbezogenen Interaktionsmustern, beobachtbarem Lernverhalten, eintretenden (Miss-)Erfolgen beim Lernen und darauf ggf. folgenden Fremdeinschätzungen der eigenen Kompetenz. Diese Aspekte können dann systematisch in ihrer Beziehung zur erlebten Kompetenz, Autonomie und sozialen Einbindung sowie dem Grad der selbstbestimmten Lernmotivation und dem Lernfortschritt analysiert werden. Dabei kann auch Unterschieden im Lernverhalten und in den lernbezogenen Interaktionsstra-

tegien von Lernenden nachgegangen werden, die diese im Austausch mit verschiedenen Interaktionspartnern (dem Lehrer, dem für die Hausaufgabenbetreuung zuständigen Elternteil sowie den Klassenkameraden) realisieren.

4. Methode

4.1 Untersuchungsdesign und Stichprobe

Um die Wirksamkeit beider Interventionen und deren Wechselwirkungen prüfen zu können, ist ein quasi-experimentelles 2 x 2-Design (vollständige Kombination von zwei Varianten Unterrichtsinterventionen mit zwei Varianten von Hausaufgabeninterventionen) realisiert worden. Eine Einschränkung in der Aussagekraft ergibt sich dadurch, dass die auf den Gruppenvergleich bezogene Analyse von Eltern-Kinder-Interaktionen stets die Mitwirkung eines Elternteils in allen Bedingungen voraussetzt. Aus diesem Grund werden zwei Hausaufgaben-Szenarios mit Elternbeteiligung verglichen, obwohl nicht davon ausgegangen werden kann, dass in der Praxis in jeder Familie eine Hausaufgabenbetreuung durch Eltern stattfindet.

Aus diesem Design resultieren vier, nahezu gleich stark besetzte Probandengruppen: U+/H+ (52 Schüler), U+/H- (54 Schüler), U-/H+ (57 Schüler), U-/H- (53 Schüler). Die Probanden stammen aus jeweils vier Parallelklassen der 7. Jahrgangsstufe zweier Duisburger Gymnasien. Um mit der Person des Lehrers verknüpfte Einflussfaktoren auf die Leistungs- und Motivationsentwicklung kontrollieren zu können, wurden an einem Gymnasium alle vier Varianten in vier Parallelklassen realisiert, in denen derselbe Lehrer unterrichtete. An der zweiten Schule wurden die Parallelklassen von einem Lehrer-ehepaar unterrichtet, das seit Jahren seine Unterrichtskonzepte abstimmt, so dass von interindividuell stark übereinstimmenden unterrichtsbezogenen Vorstellungen und Kompetenzen ausgegangen werden kann. Die Lehrer realisierten die Varianten in wechselnder Abfolge, um Reihenfolgeeffekte zu vermeiden.

Beide Unterrichtsvarianten sind problemorientiert konzipiert, d.h. sie setzen zur Erzeugung kognitiver Konflikte an den Alltagsvorstellungen der Schüler an, wobei in der Variante U+ hinsichtlich der Sozialform eine Gruppenarbeit stattfindet. Das im naturwissenschaftlichen Unterricht übliche experimentelle Vorgehen wird in ein spezielles kooperatives Lehr-Lernsetting (Gruppenrecherche mit Einsatz von Interaktionsboxen mit Geräten, Substanzen, verschiedenen Arbeitsblätter und Informationsmaterialien) eingebunden, um so eine strukturierte und zugleich Eigenaktivität verlangende Lernumgebung zu schaffen. Die Materialien reduzieren durch Vorstrukturierung die kognitive Belastung. Die Dominanz des Lehrers ist verglichen mit dem herkömmlichen Unterricht in beiden Lernumgebungen in sofern zurückgenommen, als die Lernenden auch im Frontalunterricht inhaltliche Problemlöseentscheidungen selbst treffen können und der Lehrer auf ihre Anweisungen hin handelt. Die Fragen der Lehrer sind entsprechend dem problemlösend konzipierten Unterricht offen und die Lehrenden sind bemüht, auf Schülervorstellungen einzugehen.

Bei der Hausaufgabenintervention H+ sollten sich die Eltern zunächst von ihren Kindern den Inhalt der letzten Chemiestunde erklären lassen und dann als Partner bei der Bearbeitung von (nahen) Transferaufgaben fungieren. In der Vergleichsgruppe (H-) sollten die Eltern nur das Erledigen der Aufgaben kontrollieren. Hier wird die Nachbereitung einer Unterrichtsstunde im Sinne der eigenständigen Rekapitulation des Gelernten und der Konsolidierung neuer Wissensstrukturen betont.

Da es sich bei der Untersuchung um ein Quasi-Experiment handelt, wurden einige potentiell relevante Hintergrundvariablen durch Konstanthaltung kontrolliert, andere wurden erfasst um ihren Effekt zumindest im Nachhinein statistisch kontrollieren zu können. Zu den konstant gehaltenen Variablen zählen die Schulzugehörigkeit, die Lerninhalte (Umfang und Schwierigkeitsgrad der Problemstellungen), die Dauer des Unterrichts (als Gesamtlernzeit) sowie Merkmale der Lehrerpersönlichkeit.

Die Prä-Tests wurden vier Wochen vor der Unterrichtsreihe durchgeführt, um Testeffekte zu vermeiden, die Post-Tests erfolgten im direkten Anschluss an die Intervention und die Follow-up-Erhebung ein halbes Jahr später. Alle Unterrichtsstunden wurden videographiert. Die längsschnittliche Videographierung ausgewählter Hausaufgaben-szenarien beschränkt sich auf die Situationen, in denen Eltern als Partner fungieren (H+). So kann geprüft werden, ob die Zusammenarbeit zwischen Jugendlichen und Eltern über die Zeit an Effektivität gewinnt.

Die Ausgangsstichprobe umfasst 215 Schüler (davon 51,2% weiblich). Die Schülerfragebogen füllten 99,1% (Prätest) bzw. 98,1% (Posttest) der Jugendlichen aus, der Chemie-Fachtest wurde von 98,6% (99,1 im Posttest) bearbeitet. Ausfälle ergaben sich nur durch Krankheit einzelner Schüler. Die im Unterricht eingesetzten treatment-begleitenden Fragebögen wurden etwas regelmäßiger (93,5 bis 98,6%) ausgefüllt als die Kurzfragebögen am Ende der Hausaufgabensitzungen (80,0% bis 97,2%). Die Rücklaufquote der Elternfragebögen fällt mit 70,7% in der Prä- und 58,6% in der Posttestung geringer als bei den Jugendlichen aus, ist aber im Vergleich zu Rücklaufquoten in familienpsychologischen Untersuchungen sehr hoch.

Das Durchschnittsalter der Jugendlichen beträgt 12,9 Jahre (N = 212). Unter den Klassenwiederholern (12,8%) sind Jungen stärker vertreten. Im Vergleich zu ihren Mitschülern sind diese etwas älter (13,7 Jahre). In 83,6% der Fälle, in denen Angaben zur Familienzusammensetzung gemacht werden (N = 152), handelt es sich um strukturell intakte Familien, die der oberen (58,3%) bzw. der unteren Mittelschicht (37,5%) angehören. In der Mehrzahl (93,5%) gehen die Väter einer Vollzeitbeschäftigung nach; 26,4% der Mütter sind Hausfrauen, 56,2% teilzeitbeschäftigt und 14% voll berufstätig. Der hohe Anteil von Müttern (79,6% bzw. 82,4% beim Posttest), die den Fragebogen ausfüllen, entspricht anderen Untersuchungen (Schmidt 1984).

4.2 *Erhebungsinstrumente*

In diesem Evaluationsvorhaben werden Daten sowohl durch paper-pencil-Verfahren als auch durch Verhaltensbeobachtungen erhoben. Die Unterrichts- und Hausaufgaben-

aufzeichnungen dienen nicht nur der Prozessanalyse, sondern auch zur Kontrolle der Umsetzung der Intervention.

Zur Erfassung der Schülervorstellungen im Bereich Säure/Base wurden verschiedene Verfahren mit steigender Schwierigkeit eingesetzt: (1) ein Multiple-Select-Test, der auf abrufbares Faktenwissen fokussiert und (2) ein Test mit offenen Transferaufgaben, die das selbstständige Formulieren von Antworten erfordern. Diese beiden Tests wurden neu entwickelt, da es kein validiertes und reliables Testinstrumentarium für den Themenbereich Säure-Base gibt. Diese Tests sind curricular valide und prüfen die Inhalte, die in beiden Interventionsgruppen erarbeitet worden sind. Die interne Konsistenz dieser Tests ist zufriedenstellend.

Um für den Lernfortschritt und die Interessenentwicklung potentiell relevante Persönlichkeitsmerkmale erfassen zu können, wurden etablierte Fragebogenverfahren eingesetzt (Wild u.a. 2001). Fragebögen kamen ferner zum Einsatz, um aus Schülersicht ein breites Spektrum an Merkmalen der Unterrichtsgestaltung (z.B. Lehrerenthusiasmus, konstruktivistisches Vorgehen, repetitives und anspruchsvolles Üben) und der häuslichen Lernumgebung (z.B. häuslicher Anregungsgehalt, elterliche Leistungserwartungen, Art und Umfang der elterlichen Hausaufgabenbetreuung) zu erfassen. Bei der Erhebung des bereichsspezifischen elterlichen Erziehungsverhaltens sowie des Instruktionsverhaltens von Lehrern wurden Skalen eingesetzt, die aus der Perspektive der Selbstbestimmungstheorie zentrale Dimensionen erzieherischen Handelns (autonomieunterstützende Instruktion, direktiv-kontrollierende Instruktion, Strukturierung, emotionale Unterstützung) abbilden. Zu den Schülermerkmalen, die erhoben wurden, zählten demographische Daten wie das Alter und Geschlecht der Kinder, aber auch Angaben zur Schulkarriere und zur sozialen Herkunft der Schüler.

Ergänzt wurden die im Eingangs- und Abschlusstest vorgelegten Fragebogendaten durch Schülerangaben, die am Ende jeder intervenierten Unterrichtsstunde bzw. nach der Erledigung der Hausaufgaben mittels eines Kurzfragebogens erhoben wurden. Diese thematisierten unter anderem die aktualisierte Lernmotivation, die stundenbezogene Erlebnisqualität (affektive Zustände und Flow-Erleben) sowie die Wahrnehmung der zurückliegenden Lehr-Lernsituation.

Aufgrund der zumeist geringen Übereinstimmung von Eltern und Kindern in der Wahrnehmung familialer Sozialisationsbedingungen (z.B. Wild/Hofer 2002) wurde ein Elternfragebogen mit zum Schülerfragenbogen weitgehend analogen Items entwickelt. Ergänzend wurden Skalen zum Familienklima, zur Kooperation von Eltern und Lehrern, zu elterlichen Erziehungszielen und zu den epistemologischen Überzeugungen von Eltern aufgenommen, die sich in anderen Studien bewährt haben. Auch wurde mit einzelnen Items erfasst, ob Eltern aufgrund ihrer schulischen und/oder beruflichen Ausbildung chemisches oder zumindest naturwissenschaftliches Vorwissen mitbringen.

Die Lehrpersonen wurden im Anschluss an die Unterrichtsreihe aufgefordert, mittels zweier Fragebögen die Intervention bezüglich angewandter Arbeitsweisen und Lernstrategien einzuschätzen. Außerdem sollten sie für jeden einzelnen Schüler Fragen zu begabungsrelevanten Verhaltensmerkmalen und zu Veränderungen in Motivation und Verständnis beantworten.

Die internen Konsistenzen der Skalen des Schüler- und Elternfragebogens sind zufriedenstellend (vgl. Wild u.a., 2001).

5. Ergebnisse

Im Rahmen des Quasi-Experiments sind mit den beiden Interventionen zwei zweifach gestufte *unabhängige Variablen* gegeben. Deren genuiner (Haupteffekt) und interaktiver Effekt (Wechselwirkung) auf die Veränderung in der Lernleistungen und der Lernmotivation der Schüler wurde mithilfe von mehrfaktoriellen Varianzanalysen bzw. Kovarianzanalysen mit Messwiederholung geprüft. Das Vorwissen der Schüler und deren Eingangsmotivation wurden statistisch auspartialisiert, sodass untersucht wurde, ob sich statistisch bedeutsame Veränderungen vom Prä- zum Posttest bzw. vom Prä- zum Follow-up-Test aufgrund der Zugehörigkeit der Schüler zu einer experimentellen Gruppe vorhersagen lassen. Der eigentlichen Hypothesentestung vorgeschaltet wurden umfangreiche Voranalysen, bei denen mithilfe univariater Varianzanalysen und Chi²-Tests etwaige Effekte von Störvariablen geprüft wurden (vgl. Abschnitt 5.1). Die nachfolgende Ergebnisdarstellung umfasst Befunde zu den zentralen Hypothesen des Projekts, wobei wir uns aus Raumgründen auf die Wirkung der unterrichtlichen Intervention konzentrieren.

5.1 Rahmenbedingungen und Treatment-Check

Die Vergleichbarkeit der Stichproben in der U+ und U- Bedingungen wurde mit Blick auf viele Variablen untersucht, darunter Geschlecht und Alter der Schüler, Daten zur Schulkarriere und vorherrschenden Art der Hausaufgabenbetreuung und die soziale Herkunft der Schüler (operationalisiert über die Schulbildung, Berufsausbildung und Erwerbstätigkeit beider Elternteile, das Haushaltseinkommen). Da sich in der U+Bedingung mehr Schüler aus sozial privilegierten Elternhäusern finden, wurde der sozioökonomische Status der Schüler bei sämtlichen Gruppenvergleichen als Kovariate berücksichtigt. Darüber hinaus werden Unterschiede in der bisherigen Hausaufgabenpraxis von Schülern der H+ und H- Bedingung in Analysen zum Effekt der Hausaufgabenintervention berücksichtigt.

Für die Umsetzung der Instruktionen bei der Hausaufgabenintervention spricht, dass die Schüler der H+ Variante häufiger ($F_{(1,122)} = 52,93$; $p < .05$) von ihren Eltern bei der Hausaufgabenerledigung unterstützt worden sind (kindperzipiertes Elternverhalten). Im Mittel gaben diese Schüler auch eine geringfügig längere Bearbeitungszeit der Hausaufgaben an (+ 1,67 Minuten bei einer mittleren Bearbeitungszeit von 12,42 Minuten; $F_{(1,135)} = 4,16$; $p < .05$). Mit Blick auf die Umsetzung der Instruktionen an die Lehrer zeigen die videographierten Unterrichtsstunden, dass alle beteiligten Lehrer die von ihnen geforderten Unterrichtsvarianten auch realisiert haben. Wir finden allerdings auch empirische Hinweise darauf, dass es den Lehrern unterschiedlich gut gelungen ist,

die gewünschte Problemorientierung zu realisieren und die Schülervorstellungen konstruktiv aufzugreifen (vgl. Abschnitt 5.3).

5.2 Zum Effekt der unterrichtlichen Interventionen

Beim Vergleich der Prä-Post- und Prä-Follow-up-Testergebnisse sind im Multiple Select-Test und bei den offenen Aufgaben mit Blick auf die Lernleistung klare Zeiteffekte über alle Gruppen hinweg zu beobachten. Signifikante Interaktionseffekte (Gruppe x Zeit) treten nur im Prä-Post-Vergleich der beiden Unterrichtsinterventionen im Zusammenhang mit dem Multiple Select-Test auf. Der Lernzuwachs ist erwartungsgemäß bei Schülern der U+ Variante größer als derjenige der U- Variante (s. Tabelle 1). Kontrastiert man die Leistungen der Schüler im Prä-Test mit denen im Follow-up-Test, dann geht die Schere zwischen den Schülern der U+ und der U- Variante aufgrund eines stärkeren Wissensabfalls der Schüler der U- Bedingung auseinander. Es zeigen sich jetzt auch tendentielle Interaktionseffekte bei den Ergebnissen zu den offenen Aufgaben.

Auch wenn die Unterschiede zwischen den beiden Unterrichtsvarianten nicht stark ausgeprägt sind, werten wir die Ergebnisse als Hinweis auf die Fruchtbarkeit eines problemorientierten kooperativen Zugangs. Die differentiellen Verläufe im Lernzuwachs beider Teilstichproben gehen ausschließlich darauf zurück, dass in der Experimentalgruppe eine 15- bis 20-minütige Gruppenarbeitsphase stattfand. Die besseren Testleistungen der Experimentalgruppe sind auch deshalb besonders positiv zu werten, weil in den Gruppenarbeitsphasen ein Teil der Unterrichtszeit für organisatorische Maßnahmen benötigt wird. Dabei deuten die Aufzeichnungen der Unterrichtsstunden daraufhin, dass die Kleingruppen kaum Zeit auf aufgabenirrelevante Kommunikation verwendet haben, sondern lediglich notwendige organisatorische Absprachen vorgenommen haben, die vor allem durch das experimentelle Arbeiten bedingt sind. Insofern dürfte eine nähere Analyse der Unterrichtsvideos die Hypothese bestätigen, dass die effektive „time on task“ bei der Gruppenarbeitsintervention kleiner ist als im Frontalunterricht – eine Erkenntnis, die konträr zu den erlebten Zeitverhältnissen im Regelunterricht ist. So bleibt aus unserer Sicht festzuhalten, dass Gruppenarbeit mindestens genauso effektiv sein kann wie ein durchgängig lehrergelenkter Unterricht.

5.3 Zur Rolle des Lehrers

Auch wenn durch das Design weitgehend ausgeschlossen werden kann, dass Unterschiede zwischen den Gruppen auf Merkmale der Lehrerpersönlichkeit zurückzuführen sind, finden sich in unseren Daten deutliche Hinweise auf die besondere Bedeutung des Lehrerverhaltens. Hochsignifikante Unterschiede zeigen sich beim Vergleich der Testergebnisse in beiden Fällen (Prä-Post und Prä-Follow-Up) beim Lernzuwachs von Schülern, die – unabhängig von der experimentellen Zugehörigkeit – von verschiedenen Lehrern unterrichtet wurden. Auf der Ebene der Lernmotivation (identifizierte Regula-

Tab. 1: Durchschnittliche Ausprägung (M) und Standardabweichung (S) in den Leistungen von Lernenden vor der Intervention (Prä-Test), danach (Post-Test) und ein halbes Jahr später (Follow-Up-Test) in Abhängigkeit von der Unterrichtsintervention (U+; U-)

	N	U +		U -		Haupteffekt: Unterrichtsintervention		Haupteffekt: Zeit		Interaktion: Unterrichtsintervention x Zeit	
		M	S	M	S	F _(1,N-1)	Eta ²	F _(1,N-1)	Eta ²	F _(1,N-1)	Eta ²
Offene Fragen - Prä	141 ¹	2,6	1,8	2,6	1,7	1,7	1,6	406,6***	3,3 ⁺	3,2 ⁺	267,4***
Offene Fragen - Post		10,6	3,9	9,7	4,2						
Multiple Select - Prä	141 ¹	25,3	7,3	26,7	6,7	.01	.01	.75	.03	.03	.70
Multiple Select - Post		43,1	5,4	41,3	7,1						
Offene Fragen - Prä	119 ¹	2,7	1,8	2,7	1,8	499,5***	0,1	4,0*	259,9***	1,2	3,8 ⁺
Offene Fragen - Follow-Up		9,2	4,1	7,9	4,3						
Multiple Select - Prä	119 ¹	25,6	7,5	26,7	6,8	.78	.00	.03	.69	.01	.03
Multiple Select - Follow-Up		40,7	7,2	38,6	7,1						

Anmerkungen: *** p<.001;
 ** p<.01;
 * p<.05;
 + p<.10;
 1 Stichprobenabnahme aufgrund des Schuljahreswechsels;
 Skalenbereiche: Offene Fragen 0 – 20 Punkte, Multiple Select 0 – 56 Punkte.

Tab. 2: Durchschnittliche Ausprägung (M) und Standardabweichung (S) in der Lernmotivation und den Leistungen von Lernenden vor der Intervention (Prä-Test), danach (Post-Test) und ein halbes Jahr später (Follow-Up-Test) in Abhängigkeit vom Lehrer

	Lehrer A		Lehrer B		Haupteffekt: Lehrer		Haupteffekt: Zeit		Interaktion: Lehrer x Zeit	
	M	S	M	S	F _(1,N-1)	Eta ²	F _(1,N-1)	Eta ²	F _(1,N-1)	Eta ²
Lernmotivation										
Ident.Regulation - Prä	2,6	0,6	2,9	0,5	2,8 ⁺	.02	9,6**	.06	5,8*	.04
Ident.Regulation - Post	2,9	0,7	2,9	0,6						
Lernleistung										
Offene Fragen - Prä	2,3	1,8	2,7	1,6	1,7	.01	558,8***	.79	8,4**	.05
Offene Fragen - Post	11,0	3,7	9,5	4,2						
Multi.Select - Prä	25,2	7,7	26,7	6,2	0,5	.00	458,1***	.76	7,3**	.05
Multi.Select - Post	43,7	6,5	41,1	6,0						
Offene Fragen - Prä	2,3	1,9	2,9	1,6	1,0	.01	296,0***	.71	8,5**	.06
Offene Fragen - Follow-Up	9,4	4,4	7,9	4,0						
Multi.Select - Prä	24,7	7,7	27,2	6,3	0,1	.00	314,3***	.72	12,5***	.09
Multi.Select - Follow-Up	41,4	7,2	38,3	7,0						

Anmerkungen: *** p<.001; ** p<.01; * p<.05; + p<.10; 1 Stichprobenabnahme aufgrund des Schuljahreswechsels; Skalenbereiche: Offene Fragen 0 – 20 Punkte, Multiple Select 0 – 56 Punkte, Identifizierte Regulation 1 – 4 Punkte (Rating).

tion) finden wir signifikante Unterschiede in den Entwicklungsverläufen beim Prä-Post-Vergleich. Ein Prä-Follow-Up-Vergleich ist an dieser Stelle nicht sinnvoll, weil es in dem dazwischenliegenden Zeitraum Lehrerwechsel gegeben hat.

6. Diskussion und Ausblick

Die großen Zeiteffekte über alle Gruppen hinweg, die auch mittelfristig Bestand haben, deuten wir dahingehend, dass der Problemorientierung im Unterricht eine entscheidende Bedeutung zukommt. In beiden Interventionsmaßnahmen werden Experimente nicht nach kleinschrittiger Anleitung durchgeführt, sondern sie müssen der Problemlage angemessen zunächst auch selbstständig geplant werden. Der ausbleibende Einbruch im Wissenstest nach einem halben Jahr spricht dafür, dass tiefergehendes Verständnis erreicht wurde. Die meisten Studien zeigen demgegenüber das Überdauern vorhandener alltagsweltlicher Vorstellungen über Schuljahre (Pfundt/Duit 1994). Auch Ergebnisse im Themenbereich Säure-Base zeigen, dass Schülervorstellungen nach dem Unterricht nahezu unverändert sind und beobachtbare Veränderungen häufig nicht in Richtung auf fachlich richtige Vorstellungen gehen, sondern nur zu einer isolierten Benutzung von Fachbegriffen und zu einer Hybridisierung von schulischem und alltagsweltlichem Wissen führen (Sumfleth/Geisler 2001).

Aus unserer Sicht lässt sich der Effekt der Gruppenarbeit vergrößern, wenn die Gruppenarbeitsphase durch Veränderung der instruktionalen Hilfen weiter optimiert wird. Eine Optimierung des Frontalunterrichts halten wir aufgrund erster Videoanalysen für ausgeschlossen, da alle Lehrer einen qualitativ herausragenden Frontalunterricht realisiert haben. Außerdem zeigen insbesondere die Kleingruppen einen überdurchschnittlichen Lernzuwachs, in denen hoch leistungsfähige Schüler mit entsprechend ausgeprägtem Selbstkonzept zusammenarbeiten. Es zeigt sich auch, dass die durchschnittlich in einer Gruppe zu beobachtende Leistungs- und Motivationsentwicklung auch von der Eingangsmotivation der Schüler abhängt. So ist beispielsweise der Wissenszuwachs dann geringer, wenn die Gruppenmitglieder überwiegend extrinsisch motiviert sind, wobei ein ausgeprägtes Eingangsinteresse der Gruppenmitglieder mit einem Absinken der extrinsischen Motivation im Verlauf der Intervention assoziiert ist. Diese Befunde lassen eine Berücksichtigung der habituellen Leistungs- und Motivationsprofile der Schüler bei der Gruppenarbeit sinnvoll erscheinen. Der hiermit aufgeworfenen Frage, welche Eingangsvoraussetzungen darüber entscheiden, wie beide Unterrichtsvarianten und das Verhalten der Lehrer wahrgenommen werden, lässt sich aufgrund der umfangreichen Daten, die wir über die Schüler haben, gut nachgehen.

Vor allem die bevorstehenden stärker qualitativen Auswertungen der Tests werden weitere wertvolle Hinweise auf Zusammenhänge zwischen Interessens- und Leistungsdaten geben. Neben Auswertungen auf der Ebene aggregierter Daten sollen auch individualspezifische Zusammenhänge geprüft werden, da Längsschnittuntersuchungen, die auf aggregierten Daten basieren, keine Schlüsse auf intraindividuelle Veränderungen zulassen (Asendorpf 2000).

Im Zentrum wird jedoch die qualitative und quantitative Analyse der Videos stehen, die mit den erlebnisdeskriptiven Daten verknüpft werden sollen, um auf diese Weise eine ökologisch valide und nah an den mikrosozialen Interaktionsprozessen orientierte Analyse des Zusammenhangs von lernbezogenen Interaktionsmustern, beobachtbarem Lernverhalten, und darauf ggf. folgenden Selbst- und Fremdeinschätzungen der eigenen Kompetenz durchführen zu können. Erste Beobachtungen zeigen schon jetzt, dass die Lernenden in der Lage zu sein scheinen, die Problemstellung herauszuarbeiten und die vorhandenen Informationen weitgehend zu nutzen. Sie bilden offenkundig auch sehr viele Hypothesen, haben aber Schwierigkeiten diese systematisch zu überprüfen. Die noch ausstehende eingehende Analyse der Videodaten sollte besonders schwierige Entscheidungssituationen offen legen und Hinweise auf von den Lernenden bevorzugte Strategien geben.

Besonders interessant schließlich ist der Vergleich des Interaktionsverhaltens der Schüler, die im Unterricht und zuhause videographiert wurden. Mithilfe dieser Beobachtungen kann geprüft werden, ob sie in der Interaktion mit verschiedenen Interaktionspartnern (d.h. dem Lehrer, den Eltern und den Klassenkameraden) verschiedene Strategien realisieren und ob sie – in Einklang mit strukturgegenständlichen Überlegungen – in der Interaktion mit den Peers ihre fachlichen und argumentativen Kompetenzen stärker einsetzen als in der Interaktion mit Erwachsenen. Unsere Videodaten scheinen zu belegen, dass nicht das Alter entscheidend ist, sondern scheinbare oder tatsächliche fachliche Überlegenheit, die kontraproduktiv für kooperatives Arbeiten sein kann. Es gelingt dann nicht, unterschiedliche Ideen gleichberechtigt intensiv zu diskutieren.

Literatur

- Asendorpf, J. (2000): Idiographische und nomothetische Ansätze in der Psychologie. In: *Zeitschrift für Psychologie* 8, S. 72–90.
- Baumert, J./Lehmann, R./Lehrke, M. u.a. (Hrsg.) (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Baumert, J./Bos, W./Watermann, R. (1998): TIMSS/III – Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J./Klieme, E./Neubrand, M. u.a. (2001): PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen, Leske + Budrich.
- Chrispeels, J./Coleman, P. (1996): Improving schools through better home-school Partnerships. In: *School effectiveness and school improvement* 7, S. 291–296.
- Deci, E.L./Ryan, R.M. (1985): *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Deci, E.L./Ryan, R.M. (1993): Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. In: *Zeitschrift für Pädagogik*, 39, S. 223–238.
- Duit, R./Häußler, P. (1997): Physik und andere naturwissenschaftliche Lernbereiche. In: Weinert, F.E. (Hrsg.): *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Themenbereich D, Serie I). Göttingen: Hogrefe, Bd. 3, S. 427–460.
- Fischer, M./Gräsel, C./Kittel, A. u.a. (1997): Strategien zur Bearbeitung von Diagnoseproblemen in komplexen Lernumgebungen. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 29, S. 62–82.

- Gräber, W. (1992): Untersuchungen zum Schülerinteresse an Chemie und Chemieunterricht. *Chemie in der Schule* 7/8, S. 270–272.
- Gräsel, C. (1997): *Problemorientiertes Lernen*. Göttingen: Hogrefe.
- Griffith, J. (1996): Relation of parental involvement, empowerment and school traits to student academic performance. In: *Journal of Educational Research* 90, S. 33–41.
- Grolnick, W.S./Deci, E.L./Ryan, R.M. (1997): Internalization within the family: The self-determination theory perspective. In: Grusec J.E./Kuczinsky L. (Eds.): *Parenting and children's internalization of values: A handbook of contemporary perspectives*. New York: Wiley, pp. 135–161.
- Helmke, A./Schrader, F.W. (1990): Zur Kompatibilität kognitiver, affektiver und motivationaler Zielkriterien des Schulunterrichts. In: Knopf, M./Schneider, W. (Hrsg.): *Entwicklung*. Göttingen: Hogrefe, S. 201–223.
- Hoffmann, L./Häußler, P./Lehrke, M. (1998): *Die IPN-Interessenstudie Physik*. Kiel: IPN.
- Klieme, E./Clausen, M. (1999): Identifying facets of problem solving in mathematics instruction. Canada: Paper presented at the AERA annual meeting in Montreal.
- Krapp, A. (1998): Entwicklung und Förderung von Interessen im Unterricht. *Psychologie in Erziehung und Unterricht* 45, S. 185–201.
- Krapp, A. (1999): Interest, motivation and learning: An educational-psychological perspective. In: *European Journal of Psychology of Education* 14, S. 23–40.
- Krapp, A./Prenzel, M. (1992): Interesse, Lernen, Leistung. *Neuere Ansätze der pädagogisch-psychologischen Interessenforschung*. Münster: Aschendorff.
- Krappmann, L. (1994): Sozialisation und Entwicklung in der Sozialwelt gleichaltriger Kinder. In: Schneewind, K.A. (Hrsg.): *Psychologie der Erziehung und Sozialisation (Enzyklopädie der Psychologie, Themenbereich D, Serie I)*. Göttingen: Hogrefe, Bd. 1, S. 495–524.
- Krohne, H.W./Hock, M. (1994): *Elterliche Erziehung und Angstentwicklung des Kindes. Untersuchungen über die Entwicklungsbedingungen von Ängstlichkeit und Angstbewältigung*. Bern: Hans Huber.
- Krumm, V. (1996): Über die Vernachlässigung der Eltern durch Lehrer und Erziehungswissenschaft. In: *Zeitschrift für Pädagogik* 34, Beiheft, S. 119–140.
- Krumm, V. (1998): *Elternhaus und Schule*. In: Rost, D. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz, S. 81–85.
- Leutner, D. (1992): *Adaptive Lehrsysteme*. Weinheim: Psychologie Verlags Union.
- Pfundt, H./Duit, R. (1994): Students alternative frameworks and science education. *Alltagsvorstellungen und naturwissenschaftlicher Unterricht*. Kiel: IPN.
- Reinmann-Rothmeier, G./Mandl, H. (1997): Lehren im Erwachsenenalter. In: Weinert, F.E./Mandl, H. (Hrsg.): *Psychologie der Erwachsenenbildung (Enzyklopädie der Psychologie, Themenbereich D, Serie I)*. Göttingen: Hogrefe, Bd. 4, S. 355–404.
- Rogoff, B. (1990): *Apprenticeship in thinking. Cognitive development in social context*. New York: Oxford University Press.
- Ryan, R.M. (1995): Psychological needs and the facilitation of integrative processes. *Journal of Personality* 63, S. 398–427.
- Schiefele, H. (1986): Interesse – Neue Antworten auf ein altes Problem. In: *Zeitschrift für Pädagogik* 32, S. 153–162.
- Schümer, G. (1998): *Mathematikunterricht in Japan – Ein Überblick über den Unterricht in öffentlichen Grund- und Mittelschulen und privaten Ergänzungsschulen*. In: *Unterrichtswissenschaft* 26, S. 195–228.
- Stark, R./Graf, M./Renkl, A. u.a. (1995): Förderung von Handlungskompetenz durch geleitetes Problemlösen und multiple Lernkontexte. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 27, S. 289–312.
- Sumfleth, E. (1987): Über den Zusammenhang zwischen Schulleistung und Gedächtnisstruktur. Eine Untersuchung zu Säure-Base-Theorien. In: *Naturwissenschaften im Unterricht – Physik/Chemie* 21, S. 29–35.

- Sumfleth, E./Geisler, A. (2001): Entwicklung von Schülervorstellungen im Laufe der Schulzeit am Beispiel der Säure-Base-Thematik. In: *chimika didaktika* 27(2), S. 122–157.
- Sumfleth, E./Wild, E. (2001): Schulische und familiäre Bedingungen des Lernens und der Lernmotivation im Fach Chemie: Evaluation eines integrierten Interventionskonzeptes zur Säure-Base-Thematik, Zwischenbericht an die DFG.
- Sweller, J. (1994): Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, S. 295–312.
- Weinert, F.E. (1997): Lerntheorien und Instruktionsmodelle. In: Weinert, F.E. (Hrsg.): *Psychologie des Lernens und der Instruktion (Enzyklopädie der Psychologie, Themenbereich D, Serie I)*. Göttingen: Hogrefe, Bd. 2, S. 1–48.
- Wild, E. (1999): *Elterliche Erziehung und schulische Lernmotivation*. Mannheim, Habilitationsschrift.
- Wild, E./Hofer, M. (2002): Familienbeziehungen in Zeiten sozialen Wandels. In: Walper, S./Pekrun, R. (Hrsg.): *Familie und Entwicklung: Perspektiven der Familienpsychologie*. Göttingen: Hogrefe.
- Wild, E./Gerber, J./Sumfleth, E./Rumann, S./Exeler, J./Buttler, N./Remy, K. (2001): *Dokumentation der Skalen für den Kinder- und Elternfragebogen*. Unveröffentlichtes Manuskript.

Anschrift der Autoren:

Prof. Dr. Elke Sumfleth, Stefan Rumann, Universität-GSH Essen, Institut für Didaktik der Chemie, Schützenbahn 70, 45127 Essen.

Elke Wild, Josef Exeler, Universität Bielefeld, Fakultät für Psychologie, Arbeitseinheit Pädagogische Psychologie, Postfach 10 01 31, 33501 Bielefeld.

Tina Gürtler/Franziska Perels/Bernhard Schmitz/Regina Bruder

Training zur Förderung selbstregulativer Fähigkeiten in Kombination mit Problemlösen in Mathematik¹

1. Einleitung

Die Ergebnisse der PISA-Studie (Programme for International Student Assessment; Baumert u.a. 2001) haben, wie zuvor schon diejenigen der internationalen Vergleichsstudie TIMSS (Third International Mathematics and Science Study, Baumert u.a. 1997) darauf hingewiesen, dass die Fähigkeiten deutscher Schülerinnen und Schüler im internationalen Vergleich als eher mittelmäßig einzustufen sind. PISA vergleicht fachspezifische und fächerübergreifende Basiskompetenzen 15-jähriger Schüler.

Auf der Ebene fachspezifischer Kompetenzen erfasst PISA die Bereiche Lesekompetenz sowie mathematische und naturwissenschaftliche Grundbildung. Ein weiterer integraler Bestandteil von PISA ist die Untersuchung fächerübergreifender Kompetenzen, wie selbstreguliertes Lernen, Kooperation und Kommunikation.

1.1 Mathematische Grundbildung

Bezüglich der mathematischen Grundbildung konnte gezeigt werden, dass deutsche Schülerinnen und Schüler deutlich schlechtere Leistungen zeigen, als die Schüler aller übrigen west- und nordeuropäischen Länder mit Ausnahme Luxemburgs. Es zeigte sich, dass nur eine äußerst kleine Gruppe von 15-jährigen Schülern selbstständig mathematisch argumentieren und reflektieren kann, weniger als die Hälfte der Schüler kann Standardaufgaben des Curriculums mit ausreichender Sicherheit lösen. Bedeutsam ist das Ergebnis, dass ein Viertel der 15-Jährigen als Risikogruppe eingestuft werden muss, deren mathematische Fähigkeit nur eingeschränkt für einen erfolgreichen Abschluss einer Berufsausbildung ausreicht.

Die empirischen Befunde von PISA bestätigen die These, dass ein wesentliches Element für die mathematische Grundbildung Fähigkeiten im Mathematisieren lebensweltbezogener Sachverhalte ist, was im PISA-Konzept als Modellierungsfähigkeit bezeichnet wird. Es zeigte sich, dass in den Ländern, die beim Vergleich bezüglich mathematischer Grundbildung besonders gut abschneiden, Modellierungsprozesse im Mathematikunterricht betont werden. Zur Förderung mathematischen Denkens sind nach

1 Die Studie wurde gefördert durch Mittel der DFG (SCHM 1538/1-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

Blum (2001) eine stärkere Verknüpfung bzw. Reflexion der mathematischen Inhalte, weniger Kalkülorientierung und mehr Eigenaktivität der Schüler sowie die Förderung verschiedener Lösungsmöglichkeiten nötig.

1.2 Selbstreguliertes Lernen

Die PISA-Studie zeigt, dass im schulischen Kontext ein Lernstil zu Erfolg führt, der durch adaptiven Lernstrategieinsatz und die Verwendung motivationaler Strategien gekennzeichnet ist. Da das Wissen über Lernstrategien eine Voraussetzung für deren erfolgreiche Anwendung und damit für bessere Leistungen ist, ergibt sich Förderbedarf für leistungsschwächere Schüler. Betrachtet man dieses Ergebnis vor dem Hintergrund des Selbstregulationsansatzes (vgl. Kraft 1999), so sollte einer Förderung von kognitiven und motivational-volitionalen Lernstrategien zur adaptiven Zielverfolgung im Sinne des selbstregulierten Lernens als fächerübergreifende Kompetenz besondere Bedeutung zukommen, um die Handlungsmöglichkeiten der Schüler zu fördern.

Anknüpfend an die untersuchten Schülergruppen in TIMSS und PISA soll mit dem in der folgenden Studie vorgestellten Trainingsprogramm die Selbstregulationskompetenz von Schülern der achten Gymnasialklasse im Mathematikunterricht verbessert werden. Zusätzlich werden Problemlösestrategien vermittelt, um zum einen die Fähigkeiten der Schüler im Bereich mathematischen Problemlösens zu fördern und zum anderen die Förderung selbstregulativer Kompetenzen zu unterstützen. Als weitere Komponente wurden zur Anleitung der bewussten und reflexiven Steuerung des eigenen Lernverhaltens standardisierte Lerntagebücher eingesetzt, die parallel zum Training das außerschulische Lernen unterstützen sollten. Um die Wirkung der Trainingskomponenten sowohl auf die Selbstregulationskompetenz als auch auf die mathematische Problemlösefähigkeit untersuchen zu können, wurde das Vorgehen einer Interventionsstudie gewählt, die im Gegensatz zu Korrelationsstudien kausale Aussagen ermöglicht. Der Artikel konzentriert sich auf die kurze Darstellung einer Trainingsvariante und die Diskussion ausgewählter Ergebnisse der Evaluation des Trainingsprogramms. Zuvor werden die Grundlagen der zugrunde liegenden theoretischen Annahmen skizziert.

2. Theorie

Im Folgenden werden die theoretischen Rahmenkonzeptionen der für das Trainingsprogramm bedeutsamen Komponenten Selbstregulation, Problemlösen und Monitoring skizziert. Selbstregulation und Problemlösen sollen kombiniert trainiert werden. Das Selbst-Monitoring ist Interventions- und Erhebungsverfahren: es werden standardisierte Lerntagebücher eingesetzt. Besonderer Schwerpunkt ist die Darstellung des Selbstregulationsmodells nach Zimmerman (2000), da dies die Grundlage für die inhaltliche Konzeption des hier vorzustellenden Trainings darstellt.

2.1 Selbstregulation

Das Selbstregulationsmodell nach Zimmerman (2000) beschreibt Selbstregulation als einen Prozess. Es geht davon aus, dass ein selbstregulierter Lerner sich zunächst ein Ziel setzt und aus einem Repertoire von Strategien adaptiv geeignete Vorgehensweisen auswählt. Das Ergebnis dieses Strategieeinsatzes wird mit dem Ziel verglichen und führt entweder zu einer Modifikation der Strategien oder zu einer Modifikation des Ziels. Der Kern des Selbstregulationsansatzes ist demnach die adaptive Zielverfolgung, die wiederum Auswirkungen auf nachfolgende Ziele und Handlungen hat. Schmitz (2001) integriert das Selbstregulationsmodell (Bandura 1991; Zimmerman 2000), das Handlungsphasenmodell (Kuhl 1987) und das Lernprozessmodell (Schmitz/Wiese 1999) zum Prozessmodell der Selbstregulation. Es werden drei Phasen im Lernvorgang unterschieden: die Vorbereitungsphase, die eigentliche Lernphase und eine Phase der Nachbereitung, die hier als präaktionale, aktionale und postaktionale Phase bezeichnet werden. Abbildung 1 stellt das Modell grafisch dar.

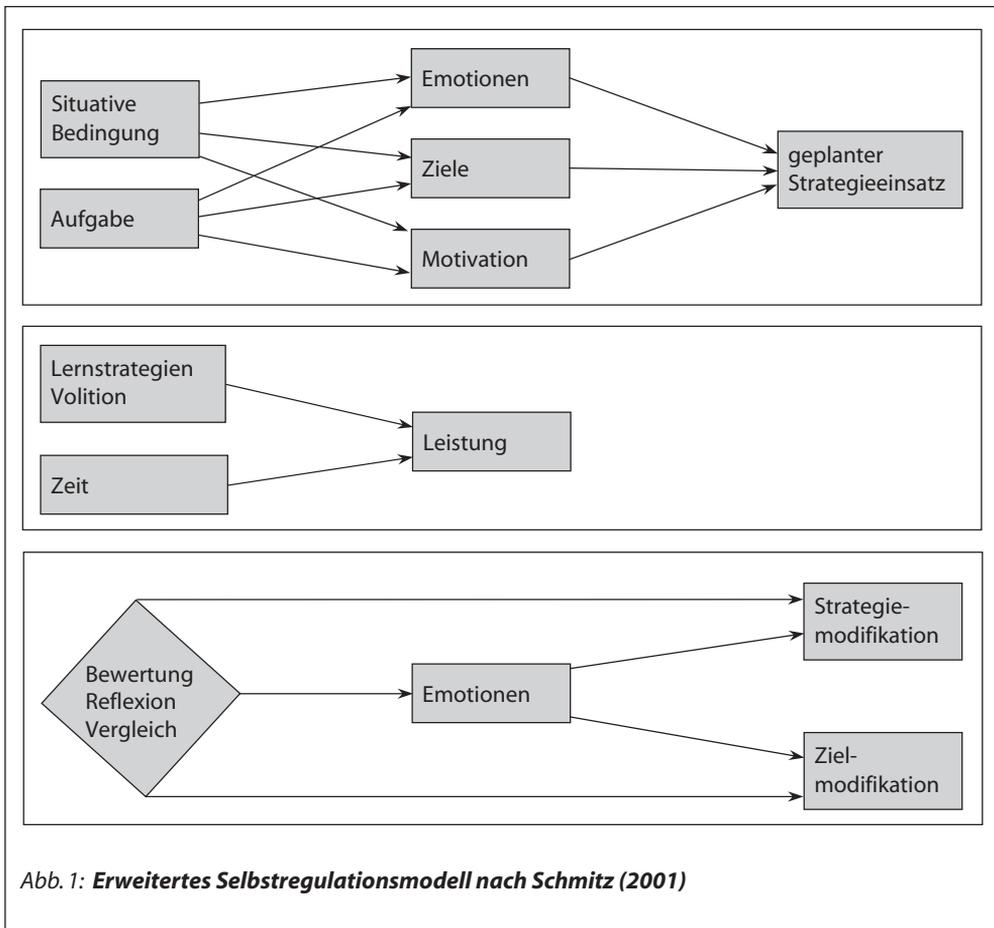


Abb. 1: **Erweitertes Selbstregulationsmodell nach Schmitz (2001)**

Nach diesem Modell beginnt der Lernprozess in der präaktionalen Phase, in der sich der Schüler in einer bestimmten Situation bzw. Lernumgebung vor eine Aufgabe gestellt sieht. Auf Grundlage der Aufgabe und der situativen Bedingung, in der diese bearbeitet werden soll, wird sich der Schüler ein bestimmtes Ziel setzen. Das Zusammenwirken von Zielsetzung und situativen, personalen und aufgabenspezifischen Merkmalen hat einen Einfluss auf Motivation und Emotionen des Schülers. Diese Parameter beeinflussen die Strategien der aktionalen Phase.

In der aktionalen Phase, der eigentlichen Lernphase, sind für den Schüler neben den zur Zielerreichung benötigten Lernstrategien volitionale Strategien von Bedeutung, um bei auftretenden Problemen die Motivation aufrecht zu erhalten und eine Konzentration auf die Aufgabe zu gewährleisten. In dieser Phase ist weiterhin die Lernzeit von Bedeutung, welche neben den qualitativen Parametern wie Lernstrategien und Volition einen quantitativen Indikator für den Lernprozess darstellt. Von besonderer Bedeutung in dieser Phase ist das stetige Überwachen der eigenen Lernhandlung (Monitoring).

In der postaktionalen Phase wird der Schüler das Ergebnis seiner Lernhandlung einschätzen und reflektieren, um Konsequenzen für weitere Lernprozesse zu ziehen. Nach Schmitz/Wiese (1999) resultieren als Folge dieser Bewertung positive bzw. negative Emotionen. Bei einem Nichterreichen des Ziels kommt es zu einer Strategie- oder Zielmodifikation. Dabei spielen die Reflexion über das Ergebnis und der Vergleich mit dem angestrebten Ziel eine bedeutende Rolle.

2.2 Problemlösen

Seit der so genannten kognitiven Wende wird das menschliche Problemlösen als Informationsverarbeitung beschrieben (Newell/Simon 1972). Dabei lassen sich zwei Teilprozesse unterscheiden. Der Problemlöseprozess beginnt mit der Problemrepräsentation, die die Grundlage für die nachfolgende Problemlösung bildet. Der Vorgang des Problemlösens wird als Absuchen eines Problemraumes beschrieben, dessen Ziel es ist, Operatoren zu finden, die vom Anfangs- zum Zielzustand führen. Diese Problemlöseoperatoren können u.a. durch geeignete Instruktion erworben werden (vgl. Anderson 1996).

Bezogen auf das Problemlösen in der Mathematik beschreibt Bruder (2000a) in Anlehnung an das Modell der Verlaufsqualitäten geistiger Handlungen von Lompscher (1972) Merkmale des Denkverlaufs beim Problemlösen mit mathematischen Mitteln durch die folgenden vier Aspekte:

- Reduktion (Vereinfachung der Problemsituation, Verringerung der Komplexität): Bezogen auf eine mathematische Problemaufgabe kann von Reduktion gesprochen werden, wenn ein Lerner in der Lage ist, die wesentlichen Komponenten der Aufgabe zu erkennen und deutlich von unwesentlichen Aufgabenmerkmalen zu unterscheiden.
- Reversibilität (Umkehren von Gedankengängen und Problemlöseschritten): Bei bestimmten Aufgabentypen ist es nötig, vom Ziel der Aufgabe systematisch zum Anfang der Aufgabe (d. h. zum Gegebenen/zur Fragestellung) zurückzuarbeiten. Mög-

liche unterstützende Fragen für Aufgaben dieses Typs sind „Was ist gesucht?; Wie kann ich das Ziel erreichen?“.

- **Aspektbeachtung** (gleichzeitiges Beachten mehrerer Aspekte, oder die Abhängigkeit von Dingen erkennen und gezielt variieren): Bei komplexen Aufgabenstellungen ist es sinnvoll, die Aufgabenstellung in verschiedene Teilaufgaben zu zerlegen, die Aufgabenteile einzeln zu bearbeiten und wieder zusammenzufügen. Dafür ist die Fähigkeit nötig, den Fokus auf bestimmte Teilaspekte zu richten und die Abhängigkeit zwischen den Teilaufgaben zu erkennen und zu beachten.
- **Aspektwechsel** (Wechsel von Annahmen und Kriterien; Umstrukturierung eines Sachverhalts): Viele Problemaufgaben zeichnen sich dadurch aus, dass zu ihrer Lösung eine Umstrukturierung der Aufgabenkomponenten erforderlich ist. Schüler, die in der Lage sind, zwischen den verschiedenen Perspektiven der Aufgabe zu wechseln, werden eher zur Lösung der Aufgabe gelangen.

Nach Bruder/Müller (1990) ist es möglich, Defizite beim Problemlösen zu kompensieren. Dies kann über eine größere Methodenbewusstheit der Schüler und die Aneignung von bestimmten heuristischen Prinzipien, Strategien und Hilfsmitteln erreicht werden. Diese Strategien werden der Problemlösefähigkeit zugeordnet (Bruder 2000b, S. 73) und als „Wirkprinzip heuristischer Bildung“ (Bruder/Stein 1999, S. 13) bezeichnet. Ihre Hypothese lautet: „Wenn sich Schüler Verfahren aneignen, die ein bewusstes Ausführen anspruchsvoller geistiger Operationen unterstützen, so können in Verbindung mit einer entsprechenden Aufgabenauswahl wesentlich bessere Ergebnisse im selbstständigen Problemlösen erzielt werden“ (Bruder 1987, S.108).

Ziel dieses Vorgehens ist es, die Schüler vom Bewusstmachen verschiedener heuristischer Verfahren und Hilfsmittel über differenzierte Einübungsphasen zu einer zunehmend selbstständigen Anwendung dieser Strategien zu befähigen.

2.3 *Selbst-Monitoring*

Unter Selbst-Monitoring wird das Beobachten bzw. Aufzeichnen des eigenen aktuellen Verhaltens verstanden. Die Bedeutung liegt darin, dass allein durch das Beobachten des eigenen Lernverhaltens bereits positive Aspekte ausgelöst werden können. Webber u.a. (1993) berichten auf Basis einer Metaanalyse, dass bereits die Selbstbeobachtung zur Förderung von erwünschtem Verhalten im Unterricht und von sozialen Kompetenzen ausreicht. In eigenen Untersuchungen wurden bereits Studien zum Einsatz von Lerntagebüchern in Verbindung mit Interventionen für verschiedene Zielgruppen durchgeführt (Schüler der fünften, achten, zehnten bis 13. Klasse, Studenten, Auszubildende, Berufsrückkehrerinnen und Berufstätige; Fischer/Heusel 1999; Ginsberg/Schwatlo 1999; Berger 2000; Deißbroth/Geibel 2000; Pflanz/Pöhl 2000; Emmel 2001; Köbler/Wolf 2001), in denen eine unterstützende Wirkung der Lerntagebücher für die Trainingskonstatiert werden konnte. Auch Wild und Remy (in vorliegendem Heft) übernahmen das in vorliegendem Artikel vorgestellte Lerntagebuch in ihren Untersuchungen.

Das hier vorgestellte Monitoring dient somit einerseits der Beobachtung des eigenen Lernverhaltens, andererseits ist es im Modell Zimmermans (2000) eine Voraussetzung für eine Bewertung des eigenen Verhaltens. Um eine angemessene Reaktion auf die Ergebniseinschätzung zu zeigen, muss eine intensive Selbstreflexion vorausgehen, bei der eine möglichst genaue Analyse der Ursachen für das Zustandekommen der Ergebnisse stattfindet.

In der vorliegenden Studie wird Monitoring über den Einsatz standardisierter Lerntagebücher operationalisiert. Die Schüler sollen mithilfe dieses Tagebuches dazu angehalten werden, ihr Lernverhalten zu beobachten, zu reflektieren und zu bewerten.

Die zentrale Fragestellung des in diesem Artikel vorgestellten Trainings betrifft die Förderungsmöglichkeit von Selbstregulations- und Problemlösekompetenz. Durch das Führen von Lerntagebüchern soll der erwartete Effekt noch verstärkt werden.

- Lassen sich Selbstregulations- und Problemlösekompetenzen durch eine kombinierte Vermittlung beider Komponenten fördern?
- Führt das Bearbeiten von standardisierten Lerntagebüchern zu einer Verbesserung ausgewählter Selbstregulations- und Problemlösekompetenzen?

3. Methode

3.1 Versuchsplan

In der Studie wurden verschiedene Varianten von Trainings realisiert, die sich durch unterschiedliche Fokussierung auf Selbstregulations- bzw. Problemlösekompetenzen unterscheiden. Bei dem hier genauer beschriebenen Training wird besonders die kombinierte Vermittlung von Selbstregulation und Problemlösen akzentuiert. Zur Überprüfung der Effekte des Lerntagebuchs wurde von der Hälfte der Schüler über den Trainingszeitraum ein Lerntagebuch bearbeitet.

Die Studie bestand aus zwei Teilen: einem interindividuellen und einem Prozessteil. Das interindividuelle Design war ein dreifaktorieller Versuchsplan mit den Faktoren

- Training Problemlösestrategien (ja/nein)
- Selbstregulationstraining (ja/nein)
- Selbst-Monitoring (ja/nein)

Es erlaubte die Prüfung der Haupteffekte sowie der Interaktion der Faktoren. In Tabelle 1 wird der interindividuelle Versuchsplan wiedergegeben. Daraus lässt sich ersehen, wie die Gruppen den acht Versuchsbedingungen zugeordnet wurden.

Tab. 1: Interindividuelles Design

Gruppe	Selbstregulation	Problemlösen	Monitoring
1	nein	nein	ja
2	nein	nein	nein
3	nein	ja	ja
4	nein	ja	nein
5	ja	nein	ja
6	ja	nein	nein
7	ja	ja	ja
8	ja	ja	nein

Die Gruppen 1, 3, 5 und 7 führten über einen Zeitraum von sieben Wochen ein standardisiertes Tagebuch. Diese Phase wurde unterteilt in zwei Teilphasen: Die erste Woche bildete die Baselinephase, der sechs Trainingswochen folgten und anschließend wieder eine Woche, in der nur Tagebücher geführt wurden. In der Treatmentphase erhielten die Schüler je nach Versuchsbedingung ein kombiniertes Training (Gruppe 5 und 6), ein Training von Problemlösestrategien (Gruppe 3 und 4) oder ein Selbstregulationstraining (Gruppe 7 und 8). Alle Trainings fanden außerhalb des regulären Schulunterrichts statt. Die Gruppen 1 und 2 erhielten kein Training. Dieser Artikel fokussiert auf die Gruppen des kombinierten Trainings.

3.2 Prozedere

Das Training fand in den Schulräumen von drei südhessischen Gymnasien statt und erfolgte einmal wöchentlich in einer Doppelstunde (90 Minuten) nachmittags im Verlaufe von sechs Wochen.

Zu Beginn des Trainingsprogramms fand eine 90-minütige Vorbefragung statt, in der ein Test zur Erfassung der Problemlösekompetenz und ein Fragebogen zur selbstregulatorischen Kompetenz zu bearbeiten waren. Eine Woche nach der letzten Trainingseinheit fand die Nachbefragung statt, die wiederum aus einem Problemlösetest (mit parallelen Aufgaben) und dem Fragebogen zur Erfassung selbstregulatorischer Kompetenz bestand. Zur Stabilitätsmessung der Effekte wurde vier Wochen nach der Nachbefragung wiederum ein Test zur Erfassung von Problemlösefähigkeit und ein Selbstregulationsfragebogen eingesetzt.

3.3 Instrumente

Zur Evaluation der Interventionen wurden längsschnittliche und prozessuale Erhebungsinstrumente eingesetzt. Auf Ebene der Längsschnittdaten diente ein Fragebogen

zur Erfassung der Selbstregulationskompetenz und ein Problemlösetest der Erhebung der jeweiligen Kompetenzen. Zur Erfassung der prozessualen Daten wurde das Lerntagebuch eingesetzt.

Der Selbstregulationsfragebogen wurde aus für die Abbildung des Selbstregulationszyklusses wesentlichen Skalen (Realisierung von Lernintentionen, Aufmerksamkeit, extrinsische Motivationskala, Anstrengung, Planung, Selbstreflexion, Lernumgebung, Umgang mit Fehlern, etc.) und einigen zusätzlichen Variablen (Selbstwirksamkeit, Einstellung zu Hausaufgaben, etc.) konstruiert.

Bezüglich des Problemlösetests wurden zwei Skalen konstruiert: Gesamtergebnis und Gesamtpunktzahl. Gesamtergebnis meint die Summe der richtig gelösten Aufgaben des Tests (maximal 16), Gesamtpunktzahl berücksichtigt die Lösungswege. In dem vorliegenden Artikel werden die Ergebnisse für die beiden Problemlöseparameter sowie das übergeordnete Selbstregulationsmaß dargestellt.

Das Lerntagebuch wurde entsprechend dem erweiterten Selbstregulationsmodell von Schmitz/Wiese (1999) entwickelt. Es wurden Trend- und Interventionsanalysen für solche Variablen durchgeführt, die im Training vermittelt bzw. für die aufgrund der durch das Monitoring angeregten Selbstreflexion Veränderungen erwartet wurden.

3.4 Stichprobe

Es wurden 249 Schüler der Jahrgangsstufe 8 dreier südhessischer Gymnasien untersucht. Die Schüler wurden den aus dem Projektdesign resultierenden Trainingsgruppen zufällig zugeteilt. 58 Schüler erhielten das hier vorgestellte Training, in dem eine Kombination aus Selbstregulation und Problemlösen vermittelt wurde.

3.5 Beschreibung des Trainings

Ziel der Schulungsmaßnahme war die Förderung einer Kombination von Selbstregulations- und Problemlösekompetenzen. Durch das begleitende Lerntagebuch sollte der Transfer relevanter Selbstregulationskomponenten in den Schulalltag, insbesondere hinsichtlich des außerschulischen Lernens unterstützt werden.

Bei der didaktischen Konzeption des Trainings wurde darauf geachtet, dass die Schüler viel Raum erhalten, selbstreguliert zu lernen. Neue Inhalte sollten entweder anhand eines Beispiels eingeführt und anschließend geübt oder entdeckend und aktivierend erlernt werden. Selbstständige Übungsphasen der Schüler folgten auf das Modelllernen. Die Schüler erhielten für ihre Leistungen bei den Hausaufgaben individuelles Feedback. Die wesentlichen Trainingsinhalte wurden teilweise in mehreren Sitzungen wiederholt. Neben aktivierenden Lehr- Lernformen gab es direkte Instruktionen. Durch das Angebot von Kleingruppenarbeit und partnerschaftlichem Bearbeiten von Problemaufgaben sollte der soziale Austausch zwischen den Schülern gefördert werden. Tabelle 2 (S. 230) gibt einen Überblick über die Inhalte des kombinierten Trainings.

Tab. 2: Übersicht über das kombinierte Training

Einheit 1	Einheit 2	Einheit 3	Einheit 4	Einheit 5	Einheit 6
Übersicht Modell	Darstellungsformen	Vorwärtsarbeiten, Rückwärtsarbeiten	Zielsetzung, Selbsteinschätzung und Motivation	Invarianzprinzip	Ergebniseinschätzung: Umgang mit Fehlern
Strategie-reflexion		Übungsphase zu Vorwärtsarbeiten, Rückwärtsarbeiten und Darstellungsformen	Vorwärtsarbeiten, Rückwärtsarbeiten, Invarianzprinzip		
Pause			Pause		Pause
Vorwärtsarbeiten	Pause	Wahlaufgaben	Pause	Volition: lernhinderlichen Gedanken	Pause
	Ziele	Ziele	Motivation		Invarianzprinzip, lernhinderliche Gedanken
Reflexion: Vorwärtsarbeiten und Aufmerksamkeit		Volition			
Zusammenfassung				Strategien zum Umgang mit lernhinderlichen Gedanken	Wissenstest

Aufbauend auf einer Modifikation des Selbstregulationsmodells von Zimmerman (2000) erfolgte beim kombinierten Training die Vermittlung von Selbstregulationsstrategien (vgl. auch Leopold/Leutner in vorliegendem Heft) im Kontext mathematischer Aufgabenstellungen.

So wurden neben den Elementen des Selbstregulationszyklusses (Ziele, Volition, Motivation, Ergebniseinschätzung, Reflexion) mathematische Strategien (Vorwärts- und Rückwärtsarbeiten, Invarianzprinzip und heuristische Hilfsmittel) vermittelt.

Das Ziel der ersten Einheit bestand darin, die Schüler anzuregen, über ihr bisheriges Vorgehen zu reflektieren. Angestrebt wurde dabei, eine aktive Aufgabenanalyse und Strategieplanung im Vergleich zu einem passivem Anwenden eines Lösungsschemas zu präferieren. In diesem Zusammenhang wurde die Notwendigkeit eines zielgerichteten Vorgehens hervorgehoben.

Die zentralen Inhalte der zweiten Einheit bezogen sich auf Zielsetzung sowie auf die Strategien des Vorwärts- und Rückwärtsarbeitens. Die Schüler sollten in dieser Einheit

lernen, dass das Setzen von Zielen ein zentrales Element der Selbstregulation darstellt. Weiterhin sollten die Strategien Vorwärts- und Rückwärtsarbeiten an mathematischen Aufgaben geübt werden.

In der dritten Einheit wurden Zielsetzung und Strategieplanung bzw. -einsatz entsprechend dem zugrunde liegenden Modell vermittelt. Die Schüler sollten dabei geeignete Strategien zur Zielerreichung auswählen. Diese Einheit bestand in einer Vertiefung der Strategien Vorwärts- und Rückwärtsarbeiten sowie der Einführung des Invarianzprinzips. Ferner sollten neben diesen mathematischen Vorgehensweisen volitionale Strategien behandelt werden.

Die vierte Einheit diente der Auffrischung und Wiederholung des für den Selbstregulationszyklus bedeutsamen Themas Ziele. Die Schüler erarbeiteten in Kleingruppen zentrale Strategien zu den Themen Volition und Motivation. Bezogen auf die Problemlösestrategien wurden die Darstellungsformen wiederholt und eingeübt.

Zentrales Ziel der fünften Einheit war es, den Schülern am Beispiel von Textaufgaben bewusst zu machen, dass sich Gedanken und Einstellungen auf die Lernleistung auswirken. Dabei sollten die Schüler im Sinne der Selbstregulation Strategien zum Umgang mit negativen Gedanken kennen lernen.

Die Schüler sollten in der sechsten Einheit die Bedeutung der Fehleranalyse innerhalb der Selbstregulation erkennen und angeleitet werden, Misserfolge auf falsche Zielsetzung, falschen Strategieeinsatz oder zu geringe Anstrengung zurückzuführen. Weiterhin sollten sie befähigt werden, die Inhalte des Trainings bei der Vorbereitung auf Klassenarbeiten anzuwenden.

4. Ergebnisse

Zur Überprüfung der Effekte der Intervention wurden interindividuelle Vergleiche und prozessuale Analysen durchgeführt. Im Folgenden werden zunächst die Ergebnisse des längsschnittlichen Gruppenvergleichs dargestellt. Dazu wurde jeweils eine vierfaktorielle Varianzanalyse mit den Faktoren Zeit, Monitoring, Selbstregulation und Problemlösen für die Oberskala Volition (Selbstregulationsfragebogen) sowie Gesamtergebnis und Gesamtpunktzahl (Problemlösetest) durchgeführt (vgl. Tabelle 3).

Aus Tabelle 3 (S. 232) ist ersichtlich, dass sich bezüglich der Oberskala Volition Effekte für den Monitoring- bzw. den Selbstregulationsfaktor ergaben. Die Gruppen ohne Monitoring blieben in ihren Einschätzungen konstant, während sich die Tagebuchgruppen verbesserten. Die Trainingsgruppen, die selbstregulatorische Strategien vermittelt bekamen (Selbstregulationstraining, kombiniertes Training), berichteten im Nachtest signifikant mehr über selbstregulatorisches Verhalten als die Gruppen, denen diese Strategien nicht vermittelt wurden. Weiterhin wurden die Interaktionen Zeit x Monitoring x Selbstregulation, Zeit x Selbstregulation x Problemlösen sowie Zeit x Monitoring x Selbstregulation x Problemlösen signifikant.

Tab. 3: Übersicht über Signifikanzen (und Effektstärken) für die Skalen Volition, Gesamtergebnis und Gesamtpunktzahl im Vorher-Nachher-Vergleich

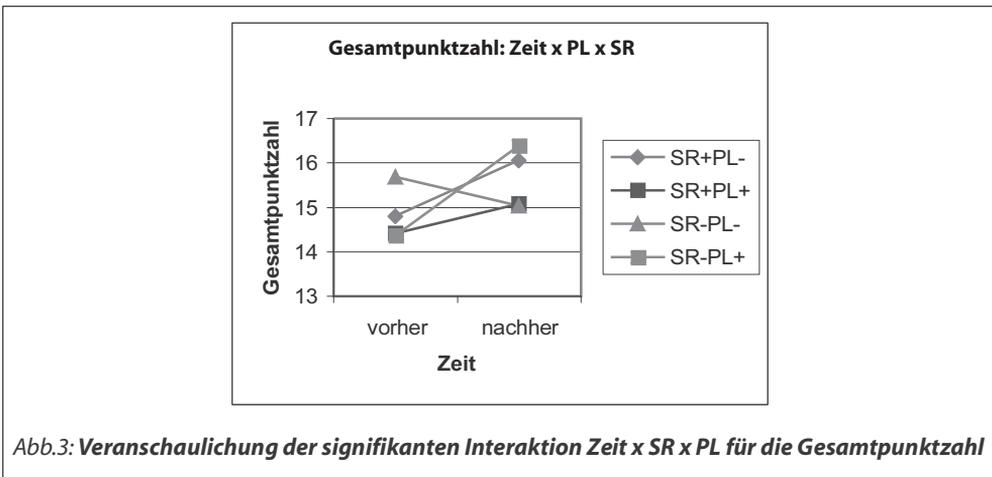
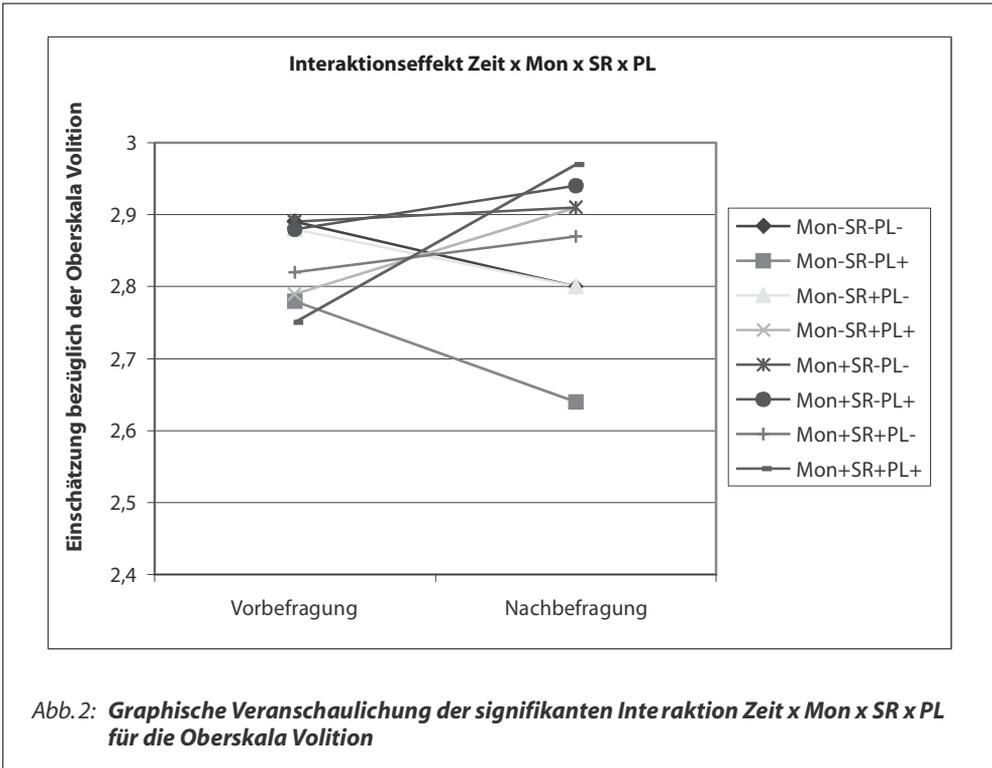
	Oberskala Volition	Gesamtergebnis	Gesamtpunktzahl
Zeit		.00*** (Eta ² = ,05)	.00*** (Eta ² = ,04)
Zeit x Monitoring	.03* (Eta ² = ,10)		.01** (Eta ² = ,03)
Zeit x Selbstregulation	.04* (Eta ² = ,09)		
Zeit x Problemlösen			
Zeit x Monitoring x Selbstregulation	.01** (Eta ² = ,12)		
Zeit x Monitoring x Problemlösen			
Zeit x Selbstregulation x Problemlösen	.10# (Eta ² = ,08)	.06# (Eta ² = ,02)	.02* (Eta ² = ,03)
Zeit x Monitoring x Selbstregulation x Problemlösen	.02* (Eta ² = ,10)		

#: p ≤ .10; *: p ≤ .05; **: p ≤ .01; ***: p ≤ .001

Abbildung 2 veranschaulicht letztere Interaktion, die alle acht Gruppen des Projektdesigns enthält².

Die Abbildung macht deutlich, dass vor allem die kombinierten Trainings (mit und ohne Tagebuch; Effektstärke von $\text{Eta}^2 = .10$) bezüglich der Oberskala Volition den deutlichsten Anstieg zu verzeichnen haben, während die Kontrollgruppe erwartungsgemäß in ihren Einschätzungen gleich bleibt. Die varianzanalytische Auswertung des Problemlösetests ergibt einen Effekt für Zeit x Monitoring und Zeit x Selbstregulation x Problemlösen für die Gesamtpunktzahl. Bezogen auf die Interaktion Zeit x Monitoring zeigt sich eine signifikante Verbesserung der Gesamtpunktzahl bei den Gruppen, die ein Tagebuch bearbeiteten, im Vergleich zu den Gruppen ohne Lerntagebuch. Exemplarisch wird die Interaktion Zeit x Selbstregulation x Problemlösen für die Gesamtpunktzahl dargestellt.

- 2 Mon-SR-PL-: Kontrollgruppe ohne Tagebuch; Mon-SR-PL+: Problemlösetraining ohne Tagebuch; Mon-SR+PL-: Selbstregulationstraining ohne Tagebuch; Mon-SR+PL+: kombiniertes Training ohne Tagebuch; Mon+SR-PL-: Kontrollgruppe mit Tagebuch; Mon+SR-PL+: Problemlösetraining mit Tagebuch; Mon+SR+PL-: Selbstregulationstraining mit Tagebuch; Mon+SR+PL+: kombiniertes Training mit Tagebuch.



Bezüglich des Problemlösetests zeigt das kombinierte Training (SR+PL+) nur einen leichten Anstieg. Zusätzlich zu den längsschnittlichen Analysen wurden anhand der Tagebuchdaten Prozessanalysen in Form von Trend- und Interventionsanalysen durchgeführt. Nachfolgende Tabelle 4 gibt einen Überblick über ausgewählte Trends für alle Gruppen mit Lerntagebuch des Projektdesigns.

Tab. 4: Übersicht über Effekte des Lerntagebuchs für die Gruppen

Abhängige Variablen	Kontrollgruppe	Problem-lösetraining	kombiniertes Training	Selbstregulationstraining
Lernumgebung (Skala)	#+	***+	***+	*+
Planung (Skala)	***+	***+	***+	***+
Internale Ressourcen (Skala)	*+	*+		***+
Lernstrategien (Skala)			**+	***+
Intrinsische Motivation			***+	**-
Reflexion (Skala)		*+	***+	
Lernen aus Fehlern		*+	***+	
Strategieeinsatz		**+	#+	
Überwachung (Skala)		***+	***+	
Hilfe holen		*-		

#: $p \leq .10$; *: $p \leq .05$; **: $p \leq .01$; ***: $p \leq .001$

Es wird deutlich, dass vor allem für das kombinierte Training die Trends für die im Training vermittelten Variablen signifikant in die erwartete Richtung³ weisen. Auch bei den anderen Trainingsgruppen ergeben die Trendanalysen hypothesenkonforme Ergebnisse. Allerdings zeigen sich dort auch Trends, die kontraintuitiv sind oder Kosten der jeweiligen Trainingsart darstellen.

Mithilfe von Trendanalysen ist es möglich, sowohl aggregierte Trends für Trainingsgruppen als auch individuelle Trends für einzelne Studienteilnehmer zu berechnen. Die nachfolgenden Abbildungen zeigen exemplarisch den Trend für die Skala Planung, einmal aggregiert über die Gruppe des kombinierten Trainings und einen individuellen Trend für einen Schüler dieser Gruppe.

Sowohl für die Trainingsgruppe als auch für den einzelnen Schüler ergeben sich signifikante lineare positive Trends für die Skala Planung, d.h. das Planungsverhalten der Schüler nimmt zu. Mithilfe der Einzelfallanalysen war es möglich, jedem Schüler ein individuelles Feedback zu geben.

Um spezifische Effekte der Intervention untersuchen zu können, wurden für entsprechende Variablen und Skalen Interventionsanalysen berechnet. Dabei wurde davon ausgegangen, dass die wöchentliche Vermittlung der für die jeweiligen Trainings spezifischen Themen entsprechende Variablen des Lerntagebuchs beeinflussen sollte. Es sollte zu einem signifikanten Unterschied zwischen den Einschätzungen der Schüler vor der Intervention (Baselinephase) und denen nach der Intervention (Interventionsphase) führen. Den Interventionsanalysen wurde ein autoregressives Modell erster Ordnung (AR1) zugrundegelegt.

3 + bedeutet ansteigender Trend; - abfallender Trend.

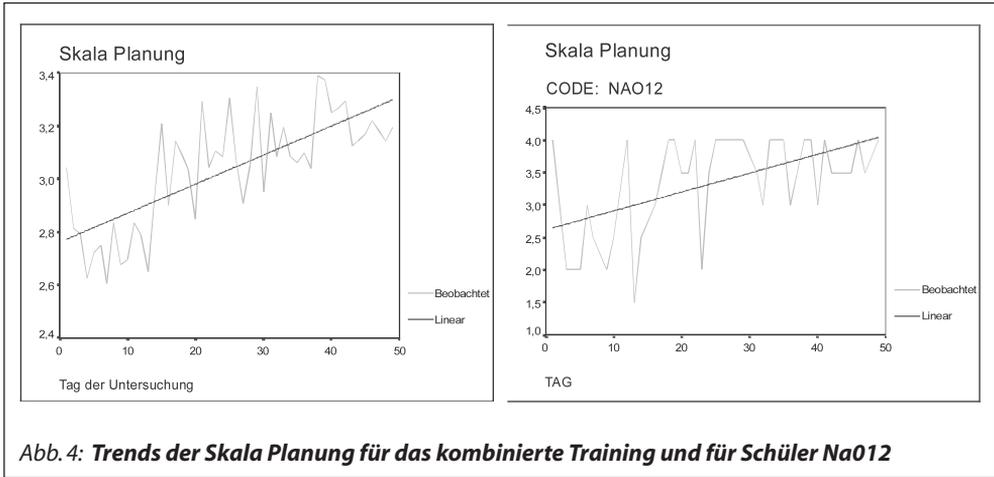


Abb. 4: Trends der Skala Planung für das kombinierte Training und für Schüler Na012

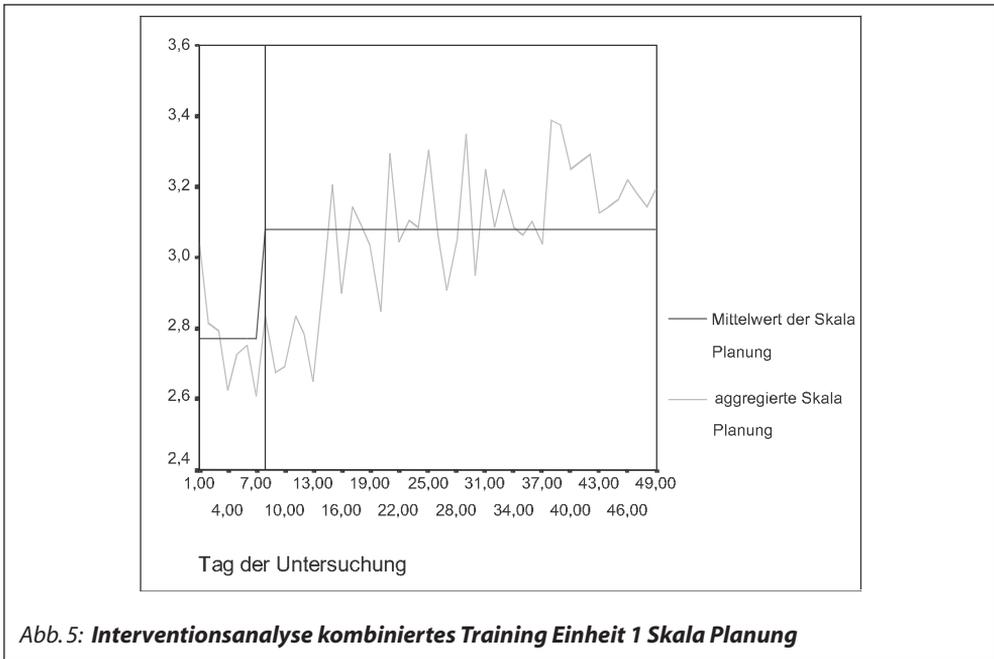


Abb. 5: Interventionsanalyse kombiniertes Training Einheit 1 Skala Planung

Insgesamt ergaben sich für das kombinierte Training bei nahezu allen erwarteten Variablen signifikante Interventionseffekte. Exemplarisch soll nachfolgend eine Interventionsanalyse für das kombinierte Training dargestellt werden. In der ersten Trainingseinheit wurde das Thema Ziele eingeführt und Strategien zur Planung vorgestellt. Die Einschätzung bezüglich der Skala „Planung“ im Lerntagebuch ließ einen signifikanten Unterschied zu Beginn der zweiten Woche der Tagebuchbearbeitung zur einwöchigen Baselinephase erkennen, wie Abbildung 5 zeigt. Wie aus der Abbildung ersichtlich ist nach der Intervention in der ersten Einheit (ab Tag 8) ein stufenförmiger Anstieg bezüglich der Einschätzung auf der Skala „Planung“ zu erkennen.

5. Diskussion

Mit unserer Interventionsstudie konnten wir zeigen, dass es mit unserem Trainingsprogramm möglich ist, Selbstregulation und Problemlösen bei Schülern der achten Gymnasialklasse zu fördern. Es hat sich sowohl in längsschnittlichen als auch in prozessualen Analysen gezeigt, dass Selbstregulationskompetenzen erworben werden können, wenn sie direkt oder in Kombination mit Problemlösestrategien vermittelt werden. Weiterhin konnten wir den Einfluss von Selbstregulation auf Problemlösen nachweisen, wie sich in den Ergebnissen zum Problemlösetest zeigt. Mithilfe einer Stabilitätsuntersuchung konnten wir zeigen, dass die erzielten Effekte längerfristig wirksam sind. Dabei wählten wir den direkten Weg des Nachweises mithilfe einer Interventionsstudie, die aussagekräftiger ist als korrelative Befunde in zahlreichen Querschnitts- oder Längsschnittstudien. Diese Trainingsergebnisse sind insbesondere vor dem Hintergrund der Bedeutung der Selbstregulation für die schulische Entwicklung hervorzuheben.

Die Evaluation zeigt die Vorteile des kombinierten Trainings deutlich auf. Es integriert für das schulische Lernen bedeutsame Komponenten, wie Metakognition, Selbstregulation und mathematische Problemlösestrategien.

Die erzielten Resultate sind insofern von Bedeutung, als die Konzeption des Trainings und seine Durchführung in der achten Jahrgangsstufe aufgrund der schwierigen Altersgruppe eine Herausforderung darstellte. Aus organisatorischen Gründen musste das Training nachmittags im Zusatzunterricht durchgeführt werden. Diese Tatsache in Kombination mit der schwierigen Altersgruppe 13- bis 14-Jähriger stellte hohe Anforderungen an die Trainer, die Schüler zu engagierter Mitarbeit anzuregen. Für die Intervention waren sechs 90-minütige Trainingssitzungen vorgesehen. Dieser Zeitraum ist insofern sehr knapp bemessen, als neben der Vermittlung selbstregulatorischer Strategien und deren Transfer in den Alltag bei den mathematischen Strategien Aufgaben ausgewählt wurden, die bei TIMSS (Baumert u.a. 1997) als besonders schwer eingestuft wurden. So gehörte die folgende Aufgabe bezüglich der Grundlagen der Prozentrechnung bei unserem Problemlösetest zu den leichten Aufgaben, während die identische Aufgabe bei TIMSS als besonders schwierig (680 Punkte) eingestuft wurde: „Der Preis einer Dose Bohnen wird von 60 Pfennig auf 75 Pfennig erhöht. Um wie viel Prozent ist der Preis gestiegen?“

Die Ergebnisse dieser Untersuchung sind insofern einzuschränken, als die untersuchte Stichprobe nicht repräsentativ für die Gesamtheit aller Gymnasiasten Deutschlands ist. Wie bereits angedeutet, wurde die Intervention nur über einen Zeitraum von sechs Wochen durchgeführt. Dies macht deutlich, dass bei einem längeren Training mit größeren Effekten bezüglich der Selbstregulations- und der Problemlösefähigkeit gerechnet werden kann.

6. Ausblick

Die Evaluation der Intervention macht deutlich, dass es möglich ist, sowohl Selbstregulationsstrategien, als auch Problemlösekompetenzen zu fördern bzw. zu vermitteln. In weiterführenden Studien sollte überprüft werden, welche Akzentuierung von Problemlöse- bzw. Selbstregulationsanteilen zu guten Ergebnissen sowohl bezogen auf die Selbstregulationskompetenzen als auch Problemlösefähigkeiten führt. Besonders deutlich sind die Effekte des Lerntagebuchs. Es konnte gezeigt werden, dass es mithilfe eines standardisierten Lerntagebuchs möglich ist, das Lernverhalten der Schüler zu dokumentieren und im Sinne einer stärkeren Reflexion zu verändern. Dies bietet den Vorteil, auch prozessuale Analysen durchführen zu können, die es ermöglichen, individuelle Verläufe zu untersuchen.

Aufgrund der besonderen Bedeutung selbstregulierten Lernens erscheint es sinnvoll, möglichst frühzeitig mit der Förderung selbstregulativer Kompetenzen zu beginnen. Aus diesem Grund ist eine Studie geplant, bei der bereits in der fünften Klasse eine Intervention vorgesehen ist, die über einen längeren Zeitraum erfolgen soll. Auch hierbei wird der Einsatz eines Lerntagebuchs geplant, das die Schüler dazu befähigen soll, sich ihrer Lernhandlungen bewusst zu werden und diese zu überwachen. Parallel dazu ist in einem weiteren Projekt vorgesehen, das Trainingskonzept für verschiedene Klassenstufen zu erweitern und in den normalen Unterricht zu integrieren. Dafür müssen die Lehrer/innen entsprechend aus- bzw. fortgebildet werden. Entsprechende Konzepte sollen ausgearbeitet und in der Lehrerbildung und an einzelnen Schulen evaluiert werden.

Literatur

- Bandura, A. (1991): Self-regulation of motivation through anticipatory and self-reactive mechanisms. In: Dienstbier, R.A. (Ed.): Perspectives on motivation: Nebraska Symposium on motivation (Vol. 38, S. 69–164). Lincoln: University of Nebraska Press.
- Baumert, J./Klieme, E./Neubrand, J./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiß, M. (2001): PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Buderich.
- Baumert, J./Lehmann, R./Lehrke, M./Schmitz, B./Clausen, M./Hosenfeld, I./Köller, O./Neubrand, J. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Buderich.
- Berger, M. (2000): Optimierung von Lernstrategien im Alltagskontext. Eine prozessuale Trainingsstudie zu emotionalen, motivationalen und volitionalen Bedingungen des selbstgesteuerten Lernens. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Blum, W. (2001): Was folgt aus TIMSS für Mathematikunterricht und Mathematiklehrerbildung? In: Klieme, E./Baumert, J. (Hrsg.): TIMSS: Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente. Bonn: Bundesministerium für Bildung und Forschung, S. 75–83.
- Bruder, R. (2000a): Problemlösen im Mathematikunterricht – ein Lernangebot für alle? In: Mathematische Unterrichtspraxis 1, S. 2–11.
- Bruder, R. (2000b): Eine akzentuierte Aufgabenauswahl und Vermitteln heuristischer Erfahrung – Wege zu einem anspruchsvollen Mathematikunterricht für alle. – In: Flade, L./Herget, W. (Hrsg.):

- Mathematik lehren und lernen nach TIMSS – Anregungen für die Sekundarstufen. Berlin: Volk und Wissen, S. 69–78.
- Bruder, R./Müller, H. (1990): Heuristisches Arbeiten im Mathematikunterricht beim komplexen Anwenden mathematischen Wissens und Könnens. In: *Mathematik in der Schule* 28, H. 12, S. 876–886.
- Bruder, R./Stein, G. (1999): Problemlösen im Mathematikunterricht. Unveröffentlichtes Skript zur Vorlesung Wintersemester 1998/1999. Technische Universität Darmstadt.
- Bruder, R. (1987): Wie können Schwierigkeiten im selbstständigen Lösen anspruchsvoller Aufgaben abgebaut werden? In: *Mathematik in der Schule* 25, H. 2/3, S. 107–113.
- Deißroth, S./Geibel, S. (2000): Selbstgesteuertes Lernen in der Ausbildung. Eine vergleichende Studie unterschiedlicher Interventionsmaßnahmen zur Optimierung des Lernverhaltens. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Emmel, S. (2001): Konzeption, Durchführung und prozessanalytische Evaluation eines Trainingsverfahrens zum Zeitmanagement und zur Selbstregulation mit begleitendem Monitoring. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Fischer, R.K./Heusel, N.C. (1999): Prozessuale Erfassung der Effekte eines Trainingsprogramms für Lernstrategien durch Tagebuch. Interventionsstudien an einem Gymnasium. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Ginsberg, E.M./Schwatlo, T. (1999): Lernlust statt Schulfrust. Entwicklung, Durchführung und Evaluation eines Trainings zur Förderung motivationaler, emotionaler und volitionaler Kompetenzen. Eine prozessorientierte Studie in der Sekundarstufe I. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Köbler, S./Wolf, S. (2001): Konzeption, Durchführung und Evaluation eines Trainingsprogramms zur Förderung selbstregulativer Kompetenzen und zur Verbesserung der mathematischen Leistungen. Eine prozessorientierte Studie bei Schülern der 8. Jahrgangsstufe eines Gymnasiums. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Kraft, S. (1999): Selbstgesteuertes Lernen. Problembereiche in Theorie und Praxis. In: *Zeitschrift für Pädagogik*, 45, 6, S. 833–845.
- Kuhl, J. (1987): Action control: The maintenance of motivational states. In: Halisch, F./Kuhl, J. (Hrsg.): *Motivation, Intention, and Volition*. Berlin: Springer, S. 279–291.
- Leopold, C./Leutner, D. (2002): Der Einsatz von Lernstrategien in einer konkreten Lernsituation in Abhängigkeit von der schulischen Jahrgangsstufe. *Zeitschrift für Pädagogik*.
- Lompscher, J. (1972): Theoretische und experimentelle Untersuchungen zur Entwicklung geistiger Fähigkeiten. Berlin: Volk und Wissen.
- Newell, A./Simon, H.A. (1972): *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Pflanz, M./Pöhl, A. (2000): Konzeption, Durchführung und prozessanalytische Evaluation eines Trainingsverfahrens zur Selbststeuerung mit begleitendem Monitoring. Unveröffentlichte Diplomarbeit. Technische Universität Darmstadt.
- Schmitz, B./Wiese, B.S. (1999): Eine Prozessstudie selbstregulierten Lernverhaltens im Kontext aktueller affektiver und motivationaler Faktoren. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 31, H. 3, S. 157–170
- Schmitz, B. (2001): Self-Monitoring zur Unterstützung des Transfers einer Schulung in Selbstregulation für Studierende. In: *Zeitschrift für Pädagogische Psychologie* 15, H. 3/4, S. 179–195.
- Webber, J./Scheuermann, B./McCall, C./Coleman, M. (1993): Research on Self-Monitoring as a Behavior Management Technique in Special Education Classrooms: A Descriptive Review. In: *Remedial and Special Education* 14, H. 2, S. 38–56.
- Wild, E./Remy, K. (2002): Schulische und familiäre Bedingungen des Mathematikinteresses, der Motive und der Zielorientierung von Drittklässlern.
- Zimmerman, B. J. (2000): Self-Efficacy: An Essential Motive to Learn. In: *Contemporary Educational Psychology* 25, S. 82–91.

Anschrift der Autoren:

Dipl. Psych. Tina Gürtler, Technische Universität Darmstadt, Fachbereich 3, Institut für Psychologie, Hochschulstraße 1, 64289 Darmstadt.

Franziska Perels, Dipl. Psych., Technische Universität Darmstadt, Fachbereich 3, Institut für Psychologie, Hochschulstr. 1, 64289 Darmstadt.

Bernhard Schmitz, Prof. Dr., Technische Universität Darmstadt, Fachbereich 3, Institut für Psychologie, Hochschulstr. 1, 64289 Darmstadt.

Regina Bruder, Prof. Dr., Technische Universität Darmstadt, Fachbereich 4, Arbeitsgruppe Fachdidaktik der Mathematik, Schloßgartenstr. 7, 64289 Darmstadt.

Claudia Leopold/Detlev Leutner

Der Einsatz von Lernstrategien in einer konkreten Lernsituation bei Schülern unterschiedlicher Jahrgangsstufen¹

1. Theorie und Fragestellung

Lernstrategien spielen bei der Entwicklung der Fähigkeit zum selbstregulierten Lernen eine zentrale Rolle. Viele Autoren betrachten sie als eine grundlegende Komponente des selbstregulierten Lernens (Boekaerts 1999; Pintrich 2000). Dies liegt unter anderem daran, dass der Lernende erst über den Einsatz von Lernstrategien Einfluss auf seinen Lernprozess und Wissenserwerb nehmen und damit eine Anforderung selbstregulierten Lernens – nämlich die aktive Steuerung des Lernprozesses – realisieren kann.

Bezogen auf das Lernen aus Texten enthalten Lernstrategien konkrete Handlungssequenzen bzw. Handlungsanweisungen, die angeben, *wie* ein Lerner vorgeht bzw. vorgehen soll, um Informationen möglichst gut aufzunehmen und zu verstehen (vgl. Klauer 1988). In Anbetracht der Ergebnisse der PISA-2000-Studie (Deutsches PISA-Konsortium 2001) haben deutsche Schüler hier offensichtlich großen Förderbedarf. So erzielen sie im internationalen Vergleich nur unterdurchschnittliche Ergebnisse, wenn sie aus Texten Informationen ermitteln, interpretieren und in bereits vorhandenes Wissen einordnen bzw. bewerten sollten. Besonders auffallend ist der überdurchschnittlich hohe Anteil von Schülern, deren Lesekompetenz unterhalb der niedrigsten definierten Kompetenzstufe liegt, was darauf hinweist, dass diese Schüler über das einfache Dekodieren des Textmaterials im Sinne einfachen (Vor-)Lesens nicht hinauskommen (Artelt u.a. 2001a). Einen Text „nur“ zu lesen, führt jedoch nicht automatisch dazu, dass der Text auch verstanden wird. Hingegen hängt ein tiefes Verständnis eng mit einer tiefen Verarbeitung der Textinhalte zusammen, was wiederum durch ein komplexes Zusammenspiel von Faktoren wie Intelligenz, Wortschatz, Motivation, Interesse, Vorwissen und Lernstrategien beeinflusst wird. Lernstrategien werden an dieser Stelle vor allem deshalb bedeutsam, weil sie Möglichkeiten für pädagogische Interventionen aufzeigen. Vor diesem Hintergrund geht es im Rahmen des von der DFG geförderten Schwerpunktprogramms Bildungsqualität von Schule (Doll/Prenzel dieses Heft) im Erfurter Projekt zunächst darum, lernstrategische Defizite beim schulischen Lernen aus naturwissenschaftlichen Texten zu untersuchen. Darüber hinaus sollen erfolgversprechende Lernstrategien identifiziert werden, um diese dann gezielt trainieren zu können. Der

1 Die vorliegende Veröffentlichung entstand im Rahmen des DFG-geförderten Forschungsprojektes LE 645/6-1. Andrea Reibert sei gedankt für die Mitarbeit bei der Herstellung der Materialien und der Datenerhebung.

vorliegende Beitrag bezieht sich auf den ersten Teil der Projektfragestellung, nämlich die Angaben der Schüler zu ihrem Lernstrategieeinsatz beim Lernen aus Texten, und die Frage, inwieweit der selbstberichtete Lernstrategieeinsatz mit dem Lernerfolg zusammenhängt.

In den gängigen Lernstrategiefragebögen (z.B. MSLQ: Pintrich u.a. 1993; Pintrich/De Groot 1990; LIST: Wild/Schiefele 1994) lassen sich drei Gruppen von Lernstrategien voneinander abgrenzen. Zur *ersten* Gruppe gehören die kognitiven Strategien. Hierunter fallen solche Strategien, die sich auf eine tiefenorientierte (Elaborationsstrategien, Organisationsstrategien) oder oberflächenorientierte (Wiederholungsstrategien) Verarbeitung des Lernstoffs beziehen. Elaborations- und Organisationsstrategien sind Strategien wie z.B. „Verknüpfung des Gelesenen mit vorhandenem Wissen“, „Nachdenken über Zusammenhänge zwischen Begriffen“ oder „Bildliche Veranschaulichung des Gelesenen“, bei denen eine verständnisbezogene Verarbeitung des Lernstoffs angestrebt wird. Wiederholungsstrategien sind dagegen auf eine oberflächliche, nicht verständnisorientierte Verarbeitung des Lernstoffs gerichtet, z.B. das Auswendiglernen des Lernstoffs.

Zur *zweiten* Gruppe gehören metakognitive Strategien, die im Gegensatz zu den kognitiven Strategien nicht auf eine unmittelbare Verarbeitung des Lernstoffs, sondern auf die Planung, Überwachung und Regulation des Lernprozesses selbst gerichtet sind. Diese Strategien sind den kognitiven Strategien übergeordnet, weil sie sich einerseits auf die Auswahl einer geeigneten kognitiven Strategie beziehen (Pintrich 2000; Zimmerman 1994) und andererseits kontrollieren und überwachen, ob eine bestimmte kognitive Strategie auch qualitativ gut ausgeführt wurde (Klauer 1985; Schreiber 1998). Die *dritte* Gruppe schließlich umfasst ressourcenbezogene Strategien, wie z.B. die Gestaltung des Lernorts, auf die hier nicht näher eingegangen wird.

Obwohl in verschiedenen Modellen selbstregulierten Lernens überzeugend dargelegt wird, dass sich der Einsatz von Lernstrategien auch im Lernergebnis widerspiegeln sollte, zeigen sich oft unklare und inkonsistente Zusammenhänge zwischen dem per Fragebogen erhobenen selbstberichteten Lernstrategieeinsatz und dem tatsächlich erzielten Lernerfolg. Sogar die plausibel klingende Annahme, dass die Verwendung von tiefenorientierten kognitiven Strategien in jedem Fall zu besseren Lernleistungen führen sollte, ließ sich nicht bestätigen (Krapp 1993). In der Forschungsliteratur zeigt sich ein Spektrum, das von negativen Zusammenhängen (z.B. Blickle 1996) über keine Zusammenhänge (z.B. Baumert 1993) bis zu schwachen positiven Zusammenhängen reicht (z.B. Artelt u.a. 2001b; Pintrich/Garcia 1993; Pintrich u.a. 1993).

Gründe für diese zum Teil erwartungswidrigen Korrelationsmuster werden vielfach genannt (Artelt 1999; Krapp 1993; Schreiber 1998; Wild 1996). Sie richten sich einerseits auf die *Operationalisierung des Lernstrategieeinsatzes* und andererseits auf die *Operationalisierung des Lernerfolgs*.

Ein erster Kritikpunkt bezüglich der Operationalisierung des Lernstrategieeinsatzes betrifft die wenig konkrete, handlungsferne Strategiemessung. So wird in den meisten Lernstrategiefragebögen der habituelle Einsatz von Lernstrategien im Sinne einer situations- und lernstoffübergreifenden Einschätzung von Lernaktivitäten erhoben (Krapp 1993). Der Lerner wird beim Ausfüllen des Fragebogens aufgefordert, seinen Lernstra-

tegieeinsatz von konkreten Lernsituationen loszulösen, um zu einer Gesamteinschätzung seines Lernverhaltens zu kommen. Dies erfordert jedoch Abstraktions- und Reflexionsleistungen, die nicht nur jüngeren Schülern Schwierigkeiten bereiten dürften und deshalb zu Verzerrungen zwischen tatsächlicher und per Fragebogen angegebener Lernstrategie Verwendung führen. Ein zweiter Kritikpunkt betrifft die Qualität der Strategieausführung. Die in Fragebogen aufgeführten Strategie-Items erfassen nur, ob bzw. wie oft eine bestimmte Strategie ausgeführt wurde (z.B. das Unterstreichen von wichtigen Informationen). Sie enthalten jedoch keine Information darüber, wie gut jemand eine bestimmte Strategie ausgeführt hat. Wenn ein Lerner z.B. angibt, die Strategie „Zusammenhänge zwischen wichtigen Begriffen herstellen“ angewandt zu haben, kann daraus nicht geschlossen werden, dass er tatsächlich die für das Gesamtverständnis relevanten Begriffe von weniger relevanten Begriffen unterschieden und miteinander verknüpft hat. Bezüglich der Unterstreichungsstrategie zeigt sich ein ähnliches Bild. Empirische Befunde weisen darauf hin, dass diese Strategie gerade dann effektiv ist, wenn nur wenige wirklich wichtige Informationen unterstrichen werden (Rickards/August 1975). Diese eher qualitativen Aspekte des Lernstrategieeinsatzes werden in Lernstrategiefragebögen in aller Regel nicht erfasst. Stattdessen erzielt jemand im Fragebogen genau dann hohe Werte auf einer Strategieskala, wenn er angibt, viele der zu der jeweiligen Skala zusammengefassten Strategien mit mittlerer Intensität oder mittlerer Häufigkeit anzuwenden oder einige wenige Strategien sehr intensiv oder sehr oft auszuführen, z.B. viel zu unterstreichen (Leutner/Leopold 2002b).

Ein weiteres Problem betrifft die Operationalisierung des Lernerfolgs. Es stellt sich die Frage, inwieweit der Einsatz von tiefenorientierten anspruchsvollen Strategien für alltägliche Lernerfolgsmaße wie Prüfungs- oder Schulleistungen überhaupt erforderlich ist. In vielen Fällen oder Situationen mag es vollkommen ausreichend sein, den Lernstoff einfach nur auswendig zu lernen, ohne sich um ein tiefergehendes Verständnis zu bemühen, weil das Lernerfolgskriterium auf das Reproduzieren und weniger auf das Verstehen des Lernstoffs gerichtet ist (Wild 1996). Aus diesem Grund ist es von entscheidender Bedeutung, den Lernstrategieeinsatz in Relation zum jeweiligen Lernerfolgskriterium zu erfassen. Artelt (1999) z.B. hat dies in ihrer Studie beachtet und fand (im Rahmen einer Pfadanalyse) einen Effekt von $\beta = .37$ der mittels Lernbeobachtung und Interview erhobenen Tiefenstrategien auf den Lernerfolg beim Wissenserwerb aus Texten. Der Lernerfolg wurde in dieser Studie nicht über ein relativ globales Leistungskriterium (z.B. Zeugnisnote) bestimmt, sondern errechnete sich aus der Anzahl der nach der Bearbeitung eines Textes korrekt wiedergegebenen Textinhalte.

Einige Untersuchungen deuten im Übrigen darauf hin, dass sich auch per Fragebogen konsistente Zusammenhänge zwischen tiefenorientierten Strategien und der Lernleistung nachweisen lassen, wenn der Lernstrategieeinsatz im Hinblick auf ein spezifisches Lernerfolgsmaß erfasst wird (Pintrich/De Groot 1990; Ainley 1993; Bouffard u.a. 1995). Ainley (1993) z.B. legte Schülern einen Strategiefragebogen vor und instruierte sie anzugeben, in welchem Ausmaß sie jede einzelne Strategie zur Vorbereitung auf eine bestimmte Prüfung verwendet hatten. Die erzielte Prüfungsnote wurde als Lernerfolgsmaß verwendet, und es zeigten sich die erwarteten positiven Korrelationen zwischen

Elaborationsstrategien und der Prüfungsnote ($r = .18$ bis $.38$) bzw. negative oder keine statistisch signifikanten Zusammenhänge zwischen Wiederholungsstrategien und der Prüfungsnote ($r = -.22$ bis $.11$).

Fasst man die genannten Kritikpunkte zusammen, sollten sich auch per Fragebogen konsistente Zusammenhänge zwischen dem Strategieeinsatz und dem Lernerfolg zeigen, wenn (1) der Lernstrategieeinsatz in einer konkreten Lernsituation erhoben wird, (2) die Qualität des Strategieeinsatzes beachtet wird und (3) der Strategieeinsatz in Relation zu einem spezifischen Lernerfolgsmaß erfasst wird. Ziel der hier vorzustellenden Studie war es, diese drei Aspekte im Untersuchungsdesign zu berücksichtigen und die Beziehungen zwischen Lernerfolg und Strategieeinsatz unter dieser Perspektive genauer zu untersuchen.

Als konkrete Lernsituation wählten wir den Wissenserwerb aus Sachtexten. Als spezifisches Lernerfolgsmaß wurde ein kriteriumsorientierter Test zum jeweiligen Text entwickelt. Der Strategiefragebogen wurde in Anlehnung an Klauers Lehrtheorie (Klauer 1985) bzw. bereits bestehende Verfahren (z.B. LIST) so konstruiert bzw. adaptiert, dass er die direkt in der Lernsituation verwendeten Strategien erfassen konnte. Zusätzlich wurden inhaltspezifische tiefenorientierte Strategien formuliert, um qualitative Aspekte des Strategieeinsatzes zu beachten.

Es wurde angenommen, dass sich die in Modellen selbstregulierten Lernens thematisierten Zusammenhänge zwischen Lernstrategien und Lernerfolg (z.B. Pintrich 2000; Zimmerman 1994) in einer konkreten Lernsituation zeigen sollten. Demnach sollten tiefenorientierte bzw. tiefenorientierte inhaltspezifische und metakognitive Strategien positiv mit dem Lernerfolg, oberflächenorientierte Strategien dagegen nicht bzw. negativ mit dem Lernerfolg korrelieren (vgl. Ainley 1993; Artelt 1999; Kardash/Amlund 1991; Pintrich/DeGroot 1990).

Von Interesse war weiterhin, ob sich die Nutzungshäufigkeit von Lernstrategien sowie die Zusammenhänge zwischen Lernstrategieeinsatz und Lernerfolg in Abhängigkeit vom Alter der befragten Schüler verändern. In der Literatur finden sich einige Hinweise, dass das Wissen über Lernstrategien und deren Nutzungshäufigkeit mit dem Alter und der Lernerfahrung zunehmen (Baumert/Köller 1996; Leutner/Leopold 2001; Myers/Paris 1978; Zimmerman/Martinez-Pons 1990), dass sich allerdings erst im frühen Erwachsenenalter ein differenziertes Repertoire an Lernstrategien herausbildet (Baumert 1993). Um dies zu untersuchen, wurden in der hier vorzustellenden Studie Schüler der Jahrgangsstufen 5, 7, 9 und 11 einbezogen. Erwartet wurde, dass die Nutzungshäufigkeit von tiefenorientierten- und metakognitiven Strategien mit dem Alter der Schüler zunimmt und die Nutzung von eher wenig effektiven oberflächenorientierten Strategien abnimmt. Aufgrund der zunehmenden Ausdifferenzierung eines flexibel einsetzbaren Strategierepertoires sollten auch die Zusammenhänge zwischen dem Lernerfolg und dem Strategieeinsatz mit ansteigendem Alter deutlicher werden. Aus entwicklungspsychologischer Perspektive wird in diesem Zusammenhang vor allem der Begriff „Nutzungsdefizit“ bedeutsam (vgl. Artelt 2000). Demnach geben vor allem jüngere Kinder zwar an, eine spezifische Strategie zu verwenden, führen sie allerdings nur ineffizient aus und können somit nicht vom Strategieeinsatz profitieren. Das würde aber bedeuten,

dass der Strategieeinsatz bei dieser Altersgruppe nur einen sehr geringen oder gar keinen Effekt auf den Lernerfolg hätte.

Zusammengefasst sollten in der hier vorzustellenden Studie folgende Fragestellungen bearbeitet werden: (1) Verändert sich die Nutzungshäufigkeit der Lernstrategien in Abhängigkeit vom Alter der untersuchten Schüler? (2) Zeigen sich konsistente Zusammenhänge zwischen Lernstrategien und dem Lernerfolg innerhalb der untersuchten Jahrgangsstufen und verändern sich diese Zusammenhänge in Abhängigkeit vom Alter der untersuchten Schüler?

2. Methode

2.1 Versuchspersonen

In die Untersuchung wurden insgesamt 318 Schüler aus zwei Realschulen und zwei Gymnasien einbezogen. Die Verteilung der Schüler auf die beiden Schultypen ist aus Tabelle 1 zu entnehmen.

Tab 1: Zusammensetzung der Stichprobe

		Mädchen	Jungen	gesamt
Klasse 5	Realschule	20	18	38
	Gymnasium	24	26	50
Klasse 7	Realschule	26	18	44
	Gymnasium	25	28	53
Klasse 9	Realschule	22	24	46
	Gymnasium	17	33	50
Klasse 11	Gymnasium	12	24	37 ¹
gesamt		146	171	318

Anmerkungen: ¹ Eine Person machte keine Angaben über ihr Geschlecht.

2.2 Material

Das für diese Veröffentlichung relevante Material umfasst naturwissenschaftliche Sachtexte, lehrzielorientierte Wissenstests, den Subtest 1 (Wortschatz) des KFT (Heller/Perleth 2000) sowie einen selbst entwickelten Lernstrategiefragebogen.

Schüler der 5. und 7. Jahrgangsstufe erhielten einen Sachtext zum Thema „Schallwellen – oder wie man im Dunkeln sehen kann“ (1164 Wörter). Schüler der 9. und 11. Jahrgangsstufe erhielten einen Text zum Thema Wasser: „H₂O – der pure Stoff“ (1681 Wörter). Zu jedem der beiden Texte wurde ein kriteriumsorientierter Wissenstest mit 16 Multiple-Choice-Fragen entwickelt (nach Klauer, 1987). Bei der Testkonstruktion

wurden vor allem solche Fragen berücksichtigt, die das Verständnis der im Text beschriebenen Sachverhalte abfragen, d.h. Fragen, deren Antworten z.B. nicht wörtlich dem Text entnommen werden konnten. Die interne Konsistenz der beiden Tests betrug $\alpha = .79$ (Schall) und $\alpha = .88$ (Wasser). Der KFT mit der Subskala Wortschatz wurde verwendet, weil er eine Schlüsselvariable für das Leseverständnis ist (Rost 1998).

Bei der Entwicklung des Strategiefragebogens wurde auf Klauers Lehrtheorie zurückgegriffen. Im Sinne von Klauer (1985) wurden Strategien erfragt, die den Lehrfunktionen Motivation, Informationsaufnahme, Informationsverarbeitung, Speicherung sowie Steuerung und Kontrolle zugeordnet werden können, wobei die Klauersche Lehrfunktion Anwendung und Transfer nicht berücksichtigt werden konnte. Die Strategie-Items wurden zum Teil selbst konstruiert und zum Teil aus bereits vorhandenen Verfahren zur Erfassung des habituellen Lernverhaltens (Lompscher 1996; Wild/Schiefele 1994; Pintrich/DeGroot 1990) so umformuliert, dass sie das strategische Vorgehen in einer konkreten Textbearbeitungssituation erfassen. Dies erschien sinnvoll, weil das in den üblichen Verfahren erfasste, von konkreten Situationen abstrahierte Lernverhalten wenig Auskunft darüber gibt, welche Strategien beim Lernen aus einem konkreten Text eingesetzt werden und erfolgreich sind. Beispiele für solche auf eine konkrete Lernsituation bezogenen Strategie-Items sind „Beim Lesen habe ich mir Erklärungen für die beschriebenen Sachverhalte überlegt“ oder „Ich habe überlegt, ob ich alle wichtigen Informationen aus dem Text aufgenommen habe“. Auf einer vierstufigen Skala (nein/eher nein/eher ja/ja) sollten die Schüler bewerten, inwieweit sie die jeweils vorgegebene Strategie beim Bearbeiten des Texts ausgeführt hatten. Die Items wurden bewusst einfach formuliert, um die jüngeren Schüler nicht zu überfordern. Wie anhand der Beispiele zu erkennen ist, wählten wir für jedes Strategie-Item eine „ich habe“-Formulierung, um den Schülern den direkten Bezug zur Lernsituation vor Augen zu halten.

Anhand von Skalenanalysen wurde geprüft, ob die einzelnen Strategie-Items zu den theoretisch angenommenen Skalen zusammengefasst werden konnten. Drei Strategieskalen ließen sich relativ klar bestimmen: Strategien zur Informationsverarbeitung (tiefenorientierte Strategien), Strategien zur oberflächlichen Speicherung (oberflächenorientierte Strategien) und eine metakognitive Strategieskala. Nur diese drei einigermaßen gut identifizierbaren Skalen, die sich auf eine begrenzte Auswahl der im Fragebogen enthaltenen Strategie-Items beziehen, wurden weiter verwendet, um die Nutzungshäufigkeit von Lernstrategien und deren Zusammenhang mit dem Lernerfolg zu untersuchen.

Die Skala der tiefenorientierten Strategien setzt sich aus neun Strategie-Items zusammen ($\alpha = .67$) und umfasst Einzelstrategien, die im Sinne der Klauerschen Lehrfunktion „Informationsverarbeitung“ darauf ausgerichtet sind, Beziehungen zwischen Begriffen herzustellen, Begriffe zu ordnen, neues Wissen mit dem Vorwissen zu verknüpfen und bildliche Vorstellungen zu generieren. Sie streben eine tiefe Verarbeitung des Lernstoffs an und entsprechen den in anderen Verfahren betitelten Elaborations- und Organisationsstrategien (Wild/Schiefele 1994) bzw. Tiefenstrategien (Artelt 1999; Marton/Säljö 1984).

Bezüglich der Klauerschen Lehrfunktion „Speichern und Abrufen“ konnte eine Skala gebildet werden, die vier Strategie-Items enthält ($\alpha = .60$). Sie wird als Skala oberflächenorientierter Strategien bezeichnet, weil die einbezogenen Strategien nicht auf das Verstehen, sondern auf das Auswendiglernen der Lerninhalte fokussiert sind, z.B. „Ich habe mich bemüht, einzelne Textstellen auswendig zu lernen, indem ich die Worte immer wieder wiederholt habe“.

Entsprechend der Klauerschen Lehrfunktion „Steuerung und Kontrolle umfasst die dritte, metakognitive Skala zehn Monitoring- und Regulationsstrategie-Items ($\alpha = .65$). Monitoring-Strategien betreffen die ständige Überwachung und Überprüfung des eigenen Lernfortschritts oder Verständnisses, z.B. „Ich habe mich in Gedanken selbst getestet, ob ich das Gelesene auch Verstanden habe“. Regulationsstrategien sind eng mit den Monitoring-Strategien verknüpft. Sie sind gewissermaßen die Reaktion auf die vorhergegangene Einschätzung des Verständnisses und passen die Lerntätigkeit an die Aufgabenforderungen an, z.B. „Inhalte, die ich mir nicht merken konnte, habe ich aufgeschrieben“.

Um den qualitativen Aspekt des Strategieeinsatzes zu beachten, waren für zwei tiefenorientierte Strategien, „Beziehungen zwischen Begriffen herstellen“ und „bildliche Vorstellungen zum Lernstoff generieren“, jeweils neun inhaltspezifische Strategie-Items entwickelt und in den Fragebogen aufgenommen worden. Die Idee dahinter bezieht sich darauf, dass es durchaus einen qualitativen Unterschied in der Strategieausführung macht, ob ein Schüler die wirklich relevanten Begriffe eines Textinhalts identifiziert und anschließend miteinander in Beziehung setzt oder ob er über irrelevante, oberflächliche Informationen nachdenkt. Ebenso sollte es bei der bildlichen Vorstellung des Lernstoffs einen Effekt auf das Verständnis haben, was genau sich ein Lerner vorstellt. Einerseits kann ein Schüler z.B. eine recht oberflächliche Vorstellung zum H_2O -Molekül entwickeln, in der er sich nur das Sauerstoffatom und die beiden Wasserstoffatome vorstellt. Andererseits kann er aber auch eine ausdifferenzierte Vorstellung generieren, in der er sich nicht nur den Aufbau des H_2O -Moleküls, sondern auch die im Text beschriebene Anordnung der Außenelektronen veranschaulicht.

Bei der Konstruktion der inhaltspezifischen Strategie-Items wurden sowohl für den Schall- als auch für den Wassertext drei Inhaltsbereiche identifiziert, und zu jedem dieser Bereiche wurden relevante Begriffe bzw. Vorstellungen bestimmt. In jeweils drei Strategie-Items pro Inhaltsbereich wurde einerseits erfasst, inwieweit ein Schüler die betreffenden inhaltspezifischen Begriffe miteinander verknüpft hat, z.B. „Hast du darüber nachgedacht, in welcher Beziehung die Begriffe ‚Dipol‘ und ‚Wasserstoffbrückenbindung‘ stehen?“, und andererseits, inwieweit er konkrete, inhaltspezifische bildliche Vorstellungen generiert hat, z.B. „Hast du dir bildlich vorgestellt, wie die Außenelektronen im H_2O -Molekül angeordnet sind?“. Auf der schon beschriebenen vierstufigen Skala hatte jeder Schüler zu bewerten, in welchem Ausmaß das jeweilige Strategie-Item für seine Vorgehensweise beim Lernen zutreffend war. Angenommen wurde, dass dieses Format aufgrund der Inhaltsspezifität leichter zu beurteilen sein sollte und sich im Vergleich zu den allgemein formulierten Strategien eher dem annähern sollte, was eine Person beim Bearbeiten des Texts tatsächlich an kognitiver Informationsverarbeitung ge-

leistet hat. Die Ergebnisse der Reliabilitätsanalyse weisen für beide Texte und Strategieskalen zufriedenstellende bis hohe interne Konsistenzen auf (Schall – Begriffe verknüpfen: $\alpha = .73$; Schall – bildliche Vorstellung: $\alpha = .77$; Wasser – Begriffe verknüpfen: $\alpha = .78$; Wasser – bildliche Vorstellung, $\alpha = .80$).

2.3 Vorgehensweise

Die gesamte Untersuchung wurde als Gruppentest im Rahmen des regulären Schulunterrichts durchgeführt und erstreckte sich über ca. zwei Schulstunden. Zunächst erhielten die Schüler einen naturwissenschaftlichen Sachtext mit der Aufgabenstellung, den Text so zu lesen bzw. zu bearbeiten, dass sie ihn möglichst gut verstehen und viel Wissen aus dem Text erwerben. Sie wurden darauf hingewiesen, das später ein Test zum Text folgt. Für Notizen etc. erhielt jeder Schüler zusätzlich zum Text ein weißes Blatt, das nach 25 Minuten zusammen mit dem Text wieder abgegeben wurde. Direkt nach der Bearbeitung des Texts wurden die Schüler instruiert, im Strategiefragebogen anzugeben, wie sie bei der Textbearbeitung vorgegangen waren bzw. welche Strategien sie eingesetzt hatten, um die Textinhalte zu verstehen. Um Lese- und Verständnisproblemen vorzubeugen, wurde den jüngsten Versuchspersonen (Fünftklässler) jedes Strategie-Item einzeln vorgelesen, das sie danach im Fragebogen entsprechend ihrem Vorgehen selbstständig bewerten sollten. Für das Ausfüllen des Strategiefragebogens wurde kein Zeitlimit vorgegeben; die meisten Schüler waren jedoch nach ca. 20 Minuten fertig. Danach erhielten alle Schüler die ihrer jeweiligen Jahrgangsstufe entsprechende Version des Subtest 1 des kognitiven Fähigkeitstests (KFT).

Zum Schluss bearbeiteten die Schüler einen lehrzielorientierten Test zu den Inhalten des Sachtextes (10 min).

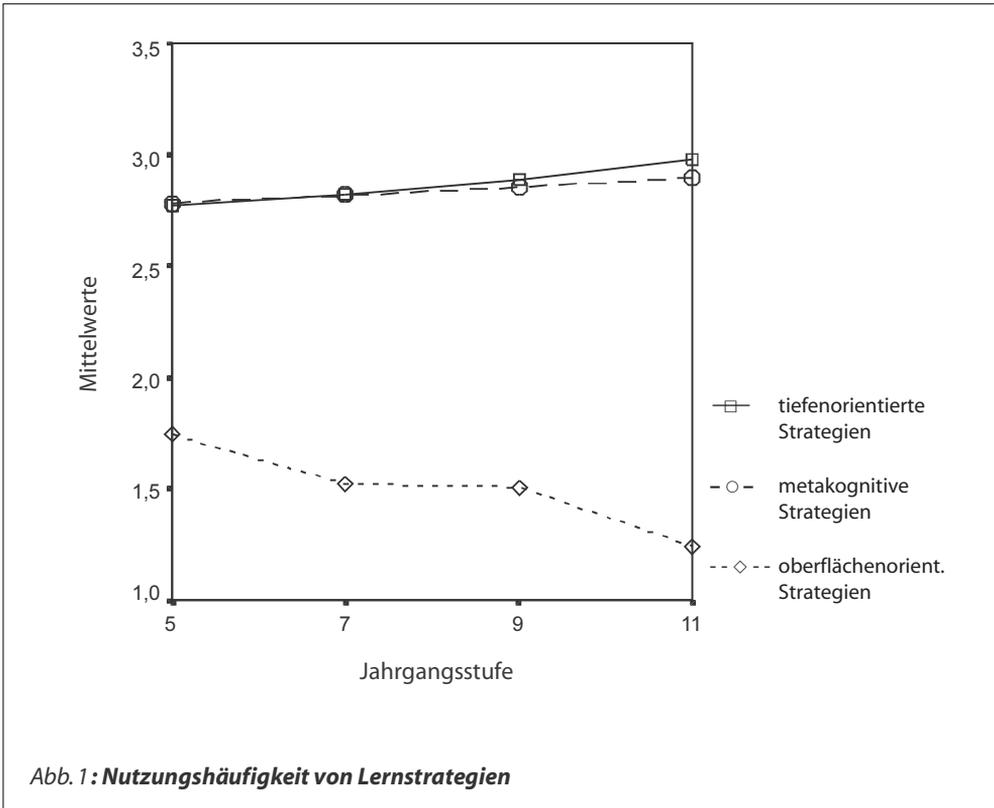
3. Ergebnisse

3.1 Nutzungshäufigkeit von Lernstrategien in Abhängigkeit vom Alter

Die erste Fragestellung bezieht sich auf die Nutzungshäufigkeit der Strategien in Abhängigkeit vom Alter der untersuchten Schüler. Es wurde erwartet, dass die Verwendung von tiefenorientierten und metakognitiven Strategien mit ansteigendem Alter zunehmen sollte. Im Gegensatz dazu sollte die Verwendung von oberflächenorientierten Strategien abnehmen.

Wie in Abbildung 1 dargestellt ist, zeigt sich für alle drei Variablen ein fast linearer Trend: Mit zunehmender Klassenstufe nimmt der selbstberichtete Einsatz von tiefenorientierten Strategien ($M = 2.77$ [$s = 0.56$]; 2.80 [0.49]; 2.90 [0.56]; 3.00 [0.41] für die Klassenstufen 5, 7, 9, 11) und metakognitiven Strategien ($M = 2.78$ [$s = 0.44$]; 2.82 [0.48]; 2.85 [0.50]; 2.90 [0.46]) zu, während der selbstberichtete Einsatz von oberflächenorientierten Strategien ($M = 1.75$ [$s = 0.57$]; 1.53 [0.54]; 1.50 [0.53]; 1.24 [0.31])

abnimmt. Für tiefenorientierte Strategien, $F_{(1,314)} = 5.96$, $p = .008$, und oberflächenorientierte Strategien, $F_{(1,314)} = 24.49$, $p < .001$, zeigt sich ein statistisch signifikanter linearer Trend. Der in den Mittelwerten der metakognitiven Strategien erkennbare Trend ist zwar nicht linear, $F_{(1,314)} = 1.66$, n.s., immerhin aber – als schwächere Annahme – monoton ansteigend (Jonckere-Terpstra-Test; in SPSS als nonparametrischer Test verfügbar), $Z_{J-T} = 1.71$, $p = .044$. Hier und im Folgenden gilt, dass alle Signifikanztests auf dem Niveau $\alpha = .05$ gerechnet werden.



Bezüglich der beiden inhaltsspezifischen Strategien – „Verknüpfen von Begriffen“ und „bildliche Vorstellung“ – wurde ebenfalls angenommen, dass ihre Verwendung mit dem Alter der befragten Schüler zunehmen sollte. Da diese Strategien für zwei unterschiedliche Texte entwickelt wurden, greifen sie allerdings unterschiedliche Inhalte, d.h. Begriffe und bildliche Vorstellungen auf, die einen Vergleich zwischen den Texten erschweren. Zudem könnte einer der beiden Texte im Gegensatz zum anderen eher geeignet sein, bildliche Veranschaulichungen anzuregen. Aus diesem Grund werden im Folgenden nur die Jahrgangsstufen miteinander verglichen, die denselben Text bearbeitet hatten. Die Mittelwertunterschiede in der Strategienutzung wurden mittels t-Test für die jeweiligen Jahrgangsstufen überprüft.

Für die Fünft- und Siebtklässler (Schall-Text) weisen die Ergebnisse des t-Tests signifikante Unterschiede für die Strategie „Begriffe verknüpfen“, $t_{(182)} = -1.99$, $p = .048$ ($d = 0.29$), jedoch nicht für die Strategie „bildliche Vorstellungen generieren“ auf, $t_{(183)} = 0.498$, n.s. Im Vergleich zu den Fünftklässlern dachten die Siebtklässler ($M = 3.09$, $s = 0.48$) somit etwas häufiger über Verknüpfungen zwischen wichtigen Begriffen nach als die Fünftklässler ($M = 2.94$, $s = 0.52$), während bei der Verwendung bildlicher Vorstellungen kein Unterschied sichtbar wurde (Fünftklässler: $M = 2.85$, $s = 0.59$; Siebtklässler: $M = 2.80$, $s = 0.64$).

Für die Neunt- und Elftklässler zeigte sich das entgegengesetzte Bild. Die t-Werte bezüglich der Strategie „Bildliche Vorstellung“ sind statistisch signifikant, $t_{(128)} = -3.74$, $p < .001$ ($d = 0.74$): Elftklässler ($M = 3.07$, $s = 0.65$) veranschaulichten sich demnach die gelesenen Textinformationen deutlich häufiger als die Neuntklässler ($M = 2.57$, $s = 0.68$). Allerdings zeigt sich kein statistisch signifikanter Unterschied beim Einsatz der Strategie „Begriffe verknüpfen“, $t_{(130)} = -1.47$, n.s. (Neuntklässler: $M = 2.72$, $s = 0.60$; Elftklässler: $M = 2.89$, $s = 0.53$).

3.2 Korrelationen zwischen Lernstrategieinsatz und Lernerfolg innerhalb der untersuchten Jahrgangsstufen

In der zweiten Fragestellung sollte geprüft werden, ob sich die erwarteten korrelativen Zusammenhänge zwischen den Strategieskalen einerseits und zwischen Strategieinsatz und Lernerfolg andererseits zeigen. Im Einzelnen wurde erwartet, dass die Korrelationen zwischen Lernerfolg einerseits und tiefenorientierten, metakognitiven Strategien sowie inhaltspezifischen Strategien andererseits mit dem Alter der Schüler stärker positiv bzw. bei oberflächenorientierten Strategien stärker negativ werden sollten. Außerdem sollten tiefenorientierte und metakognitive Strategien sowie inhaltspezifische Strategien hoch untereinander korreliert sein, jedoch nicht bzw. negativ mit den Oberflächenstrategien korrelieren.

Betrachtet man die Korrelationen innerhalb der Jahrgangsstufen, so findet sich bei den Fünftklässlern (siehe Tabelle 2) hinsichtlich des Lernerfolgs nur eine einzige statistisch signifikante Korrelation, und zwar bei der inhaltspezifischen Strategie „Begriffe verknüpfen“ ($r = .25$). Erwartungsgemäß sind allerdings die Zusammenhänge zwischen den beiden inhaltspezifischen Strategien ($r = .51$) sowie zwischen inhaltspezifischen und tiefenorientierten ($r = .54$ bzw. $.45$) relativ hoch. Besonders auffällig ist die Korrelation von $r = .73$ zwischen metakognitiven und tiefenorientierten Strategien: Fünftklässler, die sich – nach eigenen Angaben – beim Lernen auf eine tiefe Verarbeitung des Lernstoffs konzentrierten, gaben somit mit großer Wahrscheinlichkeit ebenfalls an, ihren Lernprozess beim Lesen zu überwachen und zu kontrollieren. Im Gegensatz dazu zeigen sich aber auch nicht erwartete positive Zusammenhänge zwischen oberflächen- und tiefenorientierten sowie zwischen oberflächenorientierten und metakognitiven Strategien.

Tab 2: Korrelationen zwischen Lernerfolg und Strategieskalen (untere Dreiecksmatrix: Schüler der Jahrgangsstufe 5; obere Dreiecksmatrix: Schüler der Jahrgangsstufe 7)

Variable	1	2	3	4	5	6
1. Lernerfolg	–	.29**	.23*	–.11	.27*	.14
2. tiefenorientierte Strategien	.01	–	.59**	.00	.51**	.41**
3. metakognitive Strategien	–.05	.73**	–	.14	.38**	.41**
4. oberflächenorientierte Strategien	–.09	.23*	.23*	–	.10	–.01
5. Begriffe verknüpfen	.25*	.54**	.52**	.08	–	.52**
6. Bildliche Vorstellung	.01	.45**	.55**	.13	.51**	–

Anmerkungen: Die Anzahl der Pbn variiert bei den Fünftklässlern (Siebtklässlern) zwischen 87 (92) und 88 (97).

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Bei der Bearbeitung des Textes verwendeten die in der untersuchten Stichprobe jüngsten Schüler demzufolge – nach eigenen Angaben – nicht nur mehr Oberflächenstrategien als ältere Schüler (vgl. Abbildung 1), sondern sie scheinen auch nicht zwischen einer oberflächlichen und tiefen Verarbeitung differenziert zu haben. Betrachtet man allerdings die inhaltspezifischen Strategien, zeigt sich ein anderes Bild. Diese Strategien korrelieren nicht mit Oberflächenstrategien, was darauf hinweist, dass die Inhaltsspezifität dieser Strategien eine realistischere Einschätzung des Lernverhaltens unterstützt.

Im Vergleich zu den Fünftklässlern treten bei den Siebtklässlern mehr statistisch signifikante Korrelationen zwischen dem Lernstrategieeinsatz (tiefenorientierte Strategien: $r = .29$; metakognitive Strategien: $r = .23$; „Begriffe verknüpfen“: $r = .27$) und dem Lernerfolg auf (siehe Tabelle 2). Das heißt: Siebtklässler, die angaben, Lernstoff beim Lesen tief zu verarbeiten, ihren eigenen Lernprozess zu kontrollieren und zu regulieren und darüber nachzudenken, wie bestimmte relevante Begriffe zusammenhängen, erzielten im Durchschnitt bessere Ergebnisse im Wissenstest als Schüler, die dies nicht oder in geringerem Ausmaß taten. Die Zusammenhänge zwischen den Strategieskalen fallen hier wie erwartet aus, sind jedoch etwas schwächer als bei Schülern der 5. Jahrgangsstufe. Während tiefenorientierte und metakognitive Strategien bei den Fünftklässlern zu $r = .73$ korrelieren, sinkt die Korrelation bei den Siebtklässlern auf $r = .59$. Dieses Sinken der Korrelation ist statistisch signifikant, $t_{(173)} = 1.65$, $p = .05$. Im Hinblick auf den Einsatz von oberflächenorientierten Strategien bestätigen sich die Erwartungen. Die bei den Fünftklässlern noch statistisch signifikanten Korrelationen zwischen oberflächen- und tiefenorientierten bzw. metakognitiven Strategien verschwinden bei den Siebtklässlern, was auf eine zunehmende Ausdifferenzierung der Strategiebereiche hinweist ($r < .14$).

Tab. 3: Korrelationen zwischen Lernerfolg und Strategieskalen (untere Dreiecksmatrix: Schüler der Jahrgangsstufe 9; obere Dreiecksmatrix: Schüler der Jahrgangsstufe 11)

Variable	1	2	3	4	5	6
1. Lernerfolg	–	.44**	.15	–.56**	.59**	.56**
2. tiefenorientierte Strategien	.39**	–	.70**	–.39*	.41*	.68**
3. Metakognitive Strategien	.31**	.70**	–	–.21	.40*	.47*
4. oberflächenorientierte Strategien	–.14	.08	.10	–	–.32	–.46**
5. Begriffe verknüpfen	.32**	.72**	.57**	.08	–	.54**
6. Bildliche Vorstellung	.40**	.60**	.58**	–.10	.60**	–

Anmerkungen: Die Anzahl der Pbn variiert bei den Neuntklässlern (Elftklässlern) zwischen 92 (36) und 96 (37).

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Bei den Neuntklässlern zeigen sich nicht nur höhere Korrelationen zwischen den Strategieskalen und dem Lernerfolg ($r > .31$), sondern diese Korrelationen betreffen darüber hinaus auch alle vier anspruchsvollen Strategieskalen (siehe Tabelle 3). Besonders auffällig ist der relativ hohe Zusammenhang von $r = .40$ zwischen dem Lernerfolg und der bildlichen Veranschaulichung spezifischer Lerninhalte. Zu beobachten ist auch, dass sich für die Beziehung zwischen Lernerfolg und Oberflächenstrategien eine negative Korrelation andeutet ($r = -.14$), die allerdings noch nicht statistisch signifikant ausfällt. Wie erwartet korrelieren tiefenorientierte und metakognitive Strategien sowie die beiden inhaltspezifischen Strategien relativ hoch miteinander ($.56 < r < .72$). Keine statistisch signifikanten Korrelationen sind dagegen zwischen oberflächenorientierten und den übrigen Strategiebereichen zu verzeichnen ($|r| < .08$).

Wie in Tabelle 3 zu erkennen ist, sind die korrelativen Zusammenhänge zwischen dem Lernerfolg und der Strategieanwendung bei den Elftklässlern deutlich höher als bei den Neunt- oder Siebtklässlern. Das trifft sowohl für die erwarteten positiven Korrelationen (tiefenorientierte Strategien: $r = .44$; inhaltspezifische Strategien: $r = .59$ bzw. $.57$) als auch für die erwarteten negativen Korrelation (oberflächenorientierte Strategien: $r = -.56$) zu. Eine Ausnahme bilden allerdings die metakognitiven Strategien: Hier ist die Beziehung mit dem Lernerfolg mit $r = .15$ nicht statistisch signifikant. In ähnlicher Weise treten bei den Elftklässlern im Vergleich zu den Neuntklässlern etwas geringere Zusammenhänge zwischen inhaltspezifischen und metakognitiven Strategien auf ($r = .40$ bzw. $.47$). Statistisch signifikante negative Korrelationen sind schließlich zwischen oberflächenorientierten einerseits und tiefenorientierten Strategien sowie inhaltspezifischen Strategien andererseits zu verzeichnen ($r < -.32$).

3.3 *Verändern sich die Zusammenhänge zwischen Lernstrategieinsatz und Lernerfolg in Abhängigkeit vom Alter der untersuchten Schüler?*

Um die Beziehung zwischen Lernerfolg und Strategienutzung über alle untersuchten Klassenstufen zu verfolgen, ist der Verlauf der Korrelationskoeffizienten in Abbildung 2 und 3 grafisch dargestellt.

Es ist zu beobachten, dass die Korrelationen zwischen den Strategieskalen und dem Lernerfolg in Abhängigkeit von der Jahrgangsstufe variieren. Die Korrelationsunterschiede zwischen den vier Jahrgangsstufen sind für tiefenorientierte, $\text{Chi}^2(3) = 9.27$, $p = .026$, oberflächenorientierte Strategien, $\text{Chi}^2(3) = 8.23$, $p = .041$, sowie für die Strategie „bildliche Vorstellung“, $\text{Chi}^2(3) = 13.49$, $p = .004$, statistisch signifikant. Für metakognitive Strategien werden die Korrelationsunterschiede signifikant, wenn nur die Jahrgangsstufen 5, 7 und 9 in die Analyse einbezogen werden, $\text{Chi}^2(2) = 6.53$, $p = .038$. Umgekehrt ist es bei der inhaltsspezifischen Strategie „Begriffe verknüpfen“ der Fall. Die Korrelationen zwischen den Jahrgangsstufen 5, 7 und 9 unterscheiden sich nicht, $\text{Chi}^2(2) = 0.30$, n.s., während der Unterschied zwischen 9. und 11. Jahrgangsstufe signifikant ist, $t_{(126)} = 1.69$, $p = .047$.

Wie den Abbildungen 2 und 3 zu entnehmen ist, werden die korrelativen Zusammenhänge zwischen Lernerfolg und tiefenorientierten Strategien sowie zwischen Lernerfolg und den beiden inhaltsspezifischen Strategien mit ansteigender Jahrgangsstufe immer klarer. Bei den inhaltsspezifischen Strategien ist dies besonders deutlich. Während sie in der 5. Jahrgangsstufe bei $r = .25$ (Begriffe verknüpfen) bzw. $r = .01$ (bildliche Vorstellung) liegen, steigen sie in der 11. Jahrgangsstufe auf $.59$ bzw. $.57$ an. Betrachtet man die metakognitiven Strategien, zeigt sich zunächst ein kontinuierlicher Anstieg der Korrelationskoeffizienten bis zur 9. Jahrgangsstufe ($r = .31$), bei den Elftklässlern ist jedoch ein Abfall zu verzeichnen ($r = .15$). Ein zunächst schwacher, ab der 9. Jahrgangsstufe bis zur 11. Jahrgangsstufe jedoch stärkerer rückläufiger Verlauf der Korrelation zeigt sich bei den oberflächenorientierten Strategien. Mit zunehmendem Alter sind die Schüler demzufolge nicht nur weniger geneigt, Wiederholungsstrategien beim Lernen zu verwenden, sondern es zeigen sich auch die erwarteten höheren negativen Korrelationen mit dem Lernerfolg.

Um zu prüfen, ob die korrelativen Zusammenhänge und ihre Veränderung mit steigender Jahrgangsstufe auch bei Auspartialisierung des Wortschatzes bestehen, wurden für jede Jahrgangsstufe Partialkorrelationen zwischen dem Lernerfolg und dem Strategieinsatz berechnet. Obwohl die Korrelationen, wie nicht anders zu erwarten, bei Auspartialisierung des Wortschatzes geringer sind, bleiben diejenigen Korrelationen statistisch signifikant, die vor Auspartialisierung des Wortschatzes auch schon signifikant waren. Eine Ausnahme bilden lediglich die Fünftklässler, bei denen der Zusammenhang zwischen der Strategie „Begriffe verknüpfen“ und dem Lernerfolg mit $r = .13$ die Signifikanzgrenze nicht mehr erreicht. Im Gegensatz dazu haben sich die Korrelationskoeffizienten bei den Siebtklässlern kaum geändert bzw. sind bei der inhaltsspezifischen Strategie „Begriffe verknüpfen“ sogar leicht angestiegen ($r = .30$). Insgesamt gleicht der Verlauf der Partialkorrelationskoeffizienten den in Abbildung 2 und 3 dargestellten Ver-

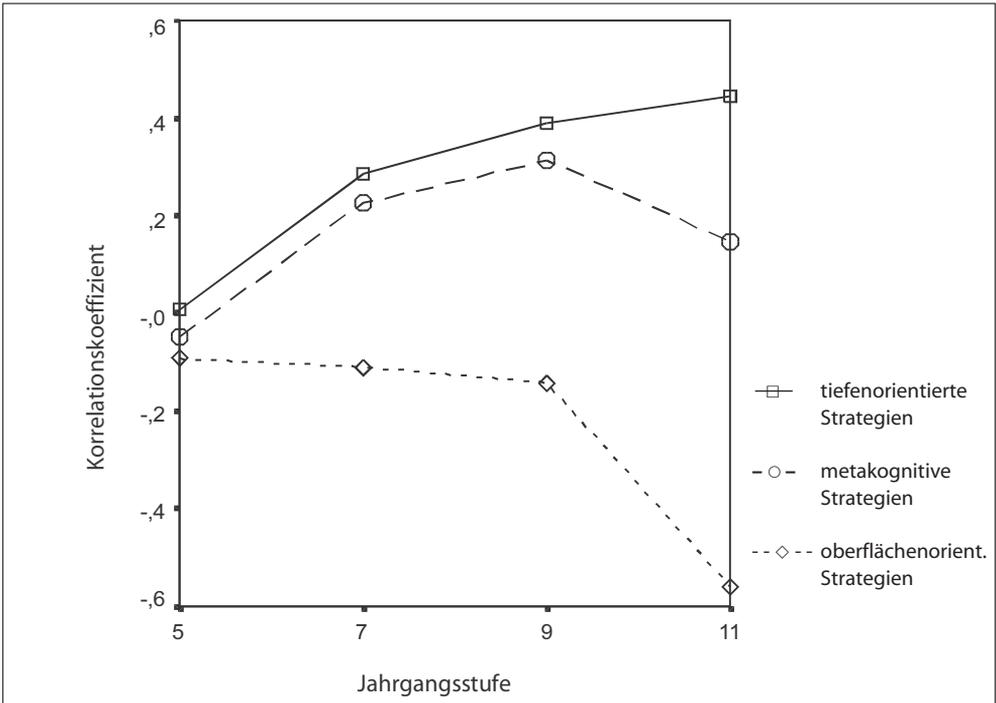


Abb. 2: **Korrelation zwischen Lernstrategien und Lernerfolg (nicht inhaltsspezifisch)**

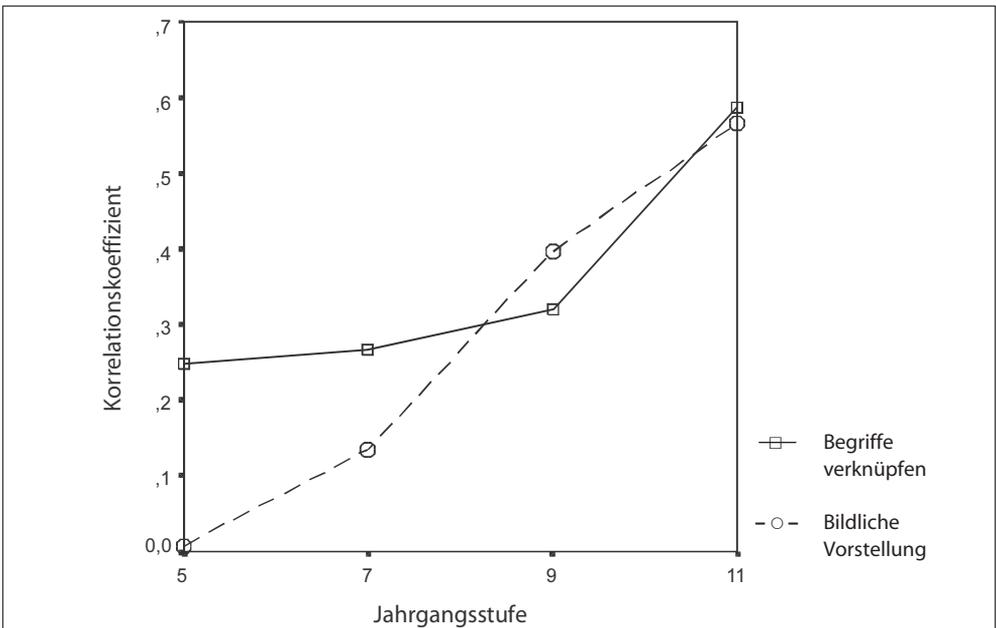


Abb. 3: **Korrelation zwischen Lernstrategien und Lernerfolg (inhaltsspezifisch)**

laufen, befindet sich jedoch auf einem niedrigeren Niveau. Bei den Elftklässlern zeigen sich tendenziell höhere Korrelationen zwischen dem Strategieeinsatz und dem Lernerfolg als bei den Neuntklässlern, während die Korrelationen der Neuntklässler die der Siebtklässler und die der Siebtklässler wiederum die der Fünftklässler übertreffen.

4. Diskussion

In der Literatur wird vielfach angenommen, dass sich geübte selbstregulierte Lerner von weniger geübten vor allem darin unterscheiden, dass sie tiefenorientierte Strategien einsetzen, um einen Lernstoff auf einem verständnisbezogenen Niveau zu verarbeiten, und dass sie darüber hinaus metakognitive Strategien einsetzen, um Verständnisprobleme zunächst wahrnehmen und anschließend gezielt abbauen zu können (Pintrich 2000; Zimmerman/Martinez-Pons 1990). Im Widerspruch dazu stehen aber empirische Befunde, die oft nur schwache Zusammenhänge zwischen dem Einsatz von Lernstrategien und dem Lernerfolg berichten (z.B. Pintrich u.a. 1993; Baumert 1993; Baumert/Köller 1996). Angenommen wurde, dass dies einerseits an der situationsfernen, habituellen Lernstrategieerfassung liegt und dass andererseits qualitative Aspekte des Lernstrategieeinsatzes in den gängigen Lernstrategiefragebögen zu wenig beachtet werden (Artelt 1999; Leutner/Leopold 2002a; Schreiber 1998).

Ebenso wurde angenommen, dass sich auch ein klarerer altersabhängiger Verlauf der Strategienutzung zeigen sollte, wenn Lernstrategien situationsspezifisch erhoben werden. Die Ergebnisse der Untersuchung bestätigen dies. Für den Einsatz von tiefenorientierten und metakognitiven Strategien konnten statistisch signifikante positive Trends und für den Einsatz von oberflächenorientierten Strategien konnte ein statistisch signifikanter negativer Trend festgestellt werden: Mit zunehmendem Alter berichten Schüler eine häufigere Verwendung von metakognitiven und kognitiven Strategien und eine seltenere Verwendung von Oberflächenstrategien – ein Ergebnis, das im Einklang mit anderen Befunden zum Wissen über bzw. zur Anwendung von Lernstrategien steht (Leutner/Leopold 2001; Myers/Paris 1978; Zimmerman/Martinez-Pons 1990). Mit zunehmendem Alter bzw. zunehmender Lernerfahrung scheinen viele Schüler von einer oberflächlicheren zu einer tieferen Verarbeitungsweise zu wechseln, was gleichzeitig ein Hinweis auf die Entwicklung eines differenzierteren Strategierepertoires ist (Baumert 1993). Bezüglich der inhaltsspezifischen Strategien dachten Siebtklässler im Vergleich zu Fünftklässlern häufiger über Beziehungen zwischen Begriffen nach, während Elftklässler gegenüber Neuntklässlern eher bildliche Vorstellungen zu den gelesenen Inhalten generierten. Diese Unterschiede zeigen zumindest für jeweils eine inhaltsspezifische Strategie den altersabhängigen Strategieverlauf auf; sie können jedoch auch darin begründet sein, dass die Textinhalte in unterschiedlicher Weise zur Veranschaulichung geeignet waren.

Bezüglich der eingangs gestellten Frage nach dem Zusammenhang zwischen Strategieeinsatz und Lernerfolg, d.h. dem aus dem Text erworbenen Wissen auf dem lehrziel-taxonomischen Niveau des Verstehens, zeigten sich Korrelationsmuster, die mit zunehmendem Alter immer konsistenter werden. Während bei den Fünftklässlern oberflä-

chenorientierte mit tiefenorientierten Strategien positiv korreliert sind, ist das bei den Siebtklässlern nicht mehr der Fall. Bei den Neuntklässlern deutet sich eine negative Korrelation an, die bei den Elftklässlern schließlich statistisch signifikant wird. Für die Fünftklässler scheint demnach eine oberflächliche Verarbeitung des Lernstoffs nicht unbedingt im Widerspruch zu einer tieferen Verarbeitung zu stehen, was bei Elftklässlern jedoch der Fall ist. Dies wäre ein Hinweis auf bei jüngeren Schülern schwächer ausgebildetes metakognitives und prozedurales Wissen (Schneider/Pressley 1989), das sich auch in einer höheren Nutzungshäufigkeit von oberflächenorientierten und in einer geringeren Nutzungshäufigkeit von tiefenorientierten und metakognitiven Strategien zeigt. Jüngere Schüler scheinen im Vergleich zu älteren nicht genau zu wissen, welche Strategien in einer konkreten Lernsituation erfolgversprechend sind – ein Ergebnis, das sich auch in den nicht statistisch signifikanten Korrelationen zwischen Strategieeinsatz und Lernerfolg widerspiegelt und auf einen unflexiblen Lernstrategieeinsatz hinweist (Baumert 1993).

Auch dann, wenn Fünftklässler angeben, Tiefenstrategien beim Lernen zu verwenden, werden keine Zusammenhänge mit dem Lernergebnis sichtbar, während dies bei Siebtklässlern, Neuntklässlern und Elftklässlern der Fall ist. Dies kann einerseits daran liegen, dass Fünftklässler ihren Strategieeinsatz aufgrund schwächer ausgebildeter metamemorialer Bewusstheit (Schneider/Pressley 1989) auch in einer konkreten Lernsituation nicht angemessen beurteilen können, oder dass sie andererseits diese Strategien qualitativ nicht gut genug eingesetzt haben. Das eine wäre ein Metagedächtnisdefizit, das andere ein Nutzungsdefizit, wobei die vorliegenden Daten leider keine Möglichkeit eröffnen, zwischen den beiden Defiziten zweifelsfrei zu unterscheiden.

Das zuvor genannte Qualitätsargument würde im Übrigen an Überzeugungskraft gewinnen, wenn in Folgestudien gezeigt werden könnte, dass der altersbedingte Anstieg der Korrelationen nur dann auftritt, wenn die gedankliche Verknüpfung *relevanter*, nicht aber, wenn die Verknüpfung *irrelevanter* Begriffe erfragt wird. Darüber hinaus wäre es für Folgestudien von Interesse, einen Indikator für die Schwierigkeit der in der Lernsituation zu bearbeitenden Sachtexte verfügbar zu haben, um durch eine geeignete Textauswahl auszuschließen, dass Tiefenstrategien und metakognitive Strategien ganz einfach deswegen nicht zum Einsatz gebracht werden, weil der Text sehr leicht oder zu schwierig zu lesen ist (vgl. Weinert 1984; Artelt 2000). Schließlich könnte in Folgestudien auch untersucht werden, ob die korrelativen Zusammenhänge bzw. Trends noch deutlicher werden, wenn zur Messung des Lernerfolgs neben verständnisprüfenden Multiple-Choice-Items auch offenere Aufgabenstellungen verwendet werden, bei denen z.B. eigene Beispiele zu finden sind oder etwas in eigenen Worten zu erklären ist.

Bezüglich der hier vorliegenden Studie ist jedoch festzuhalten, dass die beschriebenen Zusammenhänge zwischen Lernstrategienutzung und Lernerfolg nicht nur bei inhaltspezifischen Tiefenstrategien, sondern auch bei inhaltsunspezifischen allgemeinen Tiefenstrategien und zum Teil auch bei metakognitiven Strategien sichtbar sind, was die in verschiedenen Studien (Artelt 1999; Kardash/Amlund 1995) thematisierte Bedeutung einer situativen Erfassung von Lernstrategien bestätigt.

Obwohl mittels Korrelationen keine Aussagen über Kausalitäten getroffen werden können, weisen die Ergebnisse zu den inhaltspezifisch erfassten Lernstrategien darauf hin, dass neben relativ stabilen kognitiven Personeneigenschaften, wie z.B. Wortschatzkenntnisse, auch leichter veränderbare Denkprozesse einen deutlichen Einfluss auf das Lernergebnis haben können. Allerdings weisen die hier berichteten Ergebnisse darauf hin, dass es vor allem jüngeren Lernern (Fünft- und Siebtklässlern) Schwierigkeiten bereitet, die Qualität ihrer Strategieausführung realistisch einzuschätzen. Bei der Vermittlung von Wissen über Lernstrategien und beim Training des Lernstrategieeinsatzes sollten deshalb über die eigentlichen Lernstrategien hinaus metakognitive Überwachungs- und Regulationsstrategien mit trainiert werden, um eine qualitätsbezogene Ausführung der Strategie bewusst zu machen und zu fördern. Anhand einer Reihe von Trainingsexperimenten konnte gezeigt werden, dass dieser Ansatz eines kombinierten Lernstrategie Trainings nicht nur durchführbar ist, sondern auch zu verbessertem Wissenserwerb beim Lesen von Sachtexten führt (Leutner/Barthel/Schreiber 2001; Leutner/Leopold 2002b; Schreiber 1998). Wie die Ergebnisse von Gürtler/Perels/Schmitz/Bruder (dieses Heft) zeigen, trifft dies nicht nur für das Lernen aus Texten, sondern in etwas veränderter Form auch für das mathematische Problemlösen zu: Ein Training, das die Vermittlung von Problemlösestrategien mit Selbstregulationsstrategien kombinierte erwies sich im Hinblick auf den Lernerfolg erfolgreicher als ein Training, in dem jeweils nur einer der beiden Strategiebereiche trainiert wurde. Angesichts dieser Ergebnisse erscheint es sinnvoll, einen derartigen Interventionsansatz in der Lernstrategieforschung weiter zu verfolgen. Das heißt, man trainiert ausgewählte Lernstrategien so, dass die Lernenden zum einen Wissen über die Strategien erwerben und zum anderen lernen, den Einsatz der Strategien selbst so zu regulieren, dass die mit den Strategien verfolgten Ziele tatsächlich auch erreicht werden.

Literatur

- Ainley, M.D. (1993): Styles of engagement with learning: multidimensional assessment of their relationship with strategy use and school achievement. In: *Journal of Educational Psychology* 85, S. 395–405.
- Artelt, C. (1999): Lernstrategien und Lernerfolg – eine handlungsnahe Studie. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 31, S. 86–96.
- Artelt, C./Demmrich, A./Baumert, J. (2001b): Selbstreguliertes Lernen. In: Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske und Budrich, S. 271–298.
- Artelt, C./Stanat, P./Schneider, W./Schiefele, U. (2001a): Lesekompetenz: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske und Budrich, S. 69–137.
- Baumert, J. (1993): Lernstrategien, motivationale Orientierung und Selbstwirksamkeitsüberzeugungen im Kontext schulischen Lernens. In: *Unterrichtswissenschaft* 4, S. 327–354.
- Baumert, J./Köller, O. (1996): Lernstrategien und schulische Leistungen. In: Möller, J./Köller, O. (Hrsg.), *Emotionen, Kognitionen und Schulleistung*. Weinheim: Beltz, S. 137–154.
- Blickle, G. (1996): Personality traits, learning strategies, and performance. In: *European Journal of Personality* 48, S. 266–279.

- Boekaerts, M. (1999): Self-regulated learning: Where we are today. In: *International Journal of Educational Research* 31, S. 445–457.
- Bouffard, T./Boisvert, J./Vezeau, C./Larouche, C. (1995). The impact of goal orientation on self-regulation and performance among college students. In: *British Journal of Educational Psychology* 65, S. 317–329.
- Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske und Budrich.
- Heller, K.A./Perleth, C. (2000): Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Beltz.
- Kardash, D.M./Amlund, J.T. (1991): Self-reported learning strategies and learning from expository text. In: *Contemporary Educational Psychology* 16, S. 117–138.
- Klauer, K.J. (1985): Framework for a theory of teaching. In: *Teaching and Teacher Education* 1, S. 5–17.
- Klauer, K.J. (1987): Kriteriumsorientierte Tests. Göttingen: Hogrefe.
- Klauer, K.J. (1988): Teaching for learning-to-learn: A critical appraisal with some proposals. In: *Instructional Science*, 17, S. 351–367.
- Krapp, A. (1993): Lernstrategien: Konzepte, Methoden und Befunde. In: *Unterrichtswissenschaft* 21, S. 291–311.
- Leutner, D./Barthel, A./Schreiber, B. (2001): Studierende können lernen, sich selbst zum Lernen zu motivieren. Ein Trainingsexperiment. In: *Zeitschrift für Pädagogische Psychologie*, 15, 155–167.
- Leutner, D./Leopold, C. (2001): Einsatz von Lernstrategien bei Schülern der 7. und 11. Klasse (Vortrag, 8. Fachtagung Pädagogische Psychologie, 17.–19.9.01). Landau.
- Leutner, D./Leopold, C. (2002a): Lehr-lerntheoretische Grundlagen selbstregulierten Lernens. In: Witthaus, U./Wittwer, W./Espe, C. (Hrsg.), *Selbstgesteuertes Lernen. Theoretische und praktische Zugänge*. Bielefeld: Bertelsmann.
- Leutner, D./Leopold, C. (2002b): Selbstreguliertes Lernen mit Textmarkierungsstrategien. Ein Trainingsexperiment (eingereicht zur Begutachtung).
- Lompscher, J. (1996): Erfassung von Lernstrategien auf der Reflexionsebene. In: *Empirische Pädagogik* 10, S. 245–275.
- Marton, F./Säljö, R. (1984): Approaches to learning. In: Marton, F./Hounsell, D./Entwistle, N. (Hrsg.), *The experience of learning*. Edinburgh: Scottish Academic Press, S. 36–55.
- Myers, I.M./Paris, S.G. (1978): Children's metacognitive knowledge about reading. In: *Journal of Educational Psychology* 70, S. 680–690.
- Pintrich, P.R. (2000): The role of goal orientation in self-regulated learning. In: Boekaerts, P./Pintrich, R./Zeidner, M. (Hrsg.), *Handbook of self-regulated learning*. San Diego: Academic Press, S. 451–502.
- Pintrich, P.R./DeGroot, E.V. (1990): Motivational and self-regulated learning components of classroom academic performance. In: *Journal of Educational Psychology* 82, S. 33–40.
- Pintrich, P.R./Garcia, T. (1993): Intraindividual differences in students' motivation and self-regulated learning. In: *Zeitschrift für Pädagogische Psychologie* 7, S. 99–107.
- Pintrich, P.R./Smith, D.A.F./Garcia, T./McKeachie, W.J. (1993): Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). In: *Educational and Psychological Measurement* 53, S. 801–813.
- Rickards, J.P./August, G.J. (1975): Generative underlining strategies in prose recall. In: *Journal of Educational Psychology* 67, S. 860–865.
- Rost, D. H. (1998): Leseverständnis. In: Rost, D.H. (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Weinheim: PVU, S. 334–339.
- Schneider, W./Pressley, M. (1989): *Memory development between 2 and 20*. New York: Springer.
- Schreiber, B. (1998): *Selbstreguliertes Lernen*. Münster: Waxmann.
- Weinert, F.E. (1984): Metakognition und Motivation als Determinanten der Lerneffektivität. Einführung und Überblick. In: Weinert, F.E./Kluwe, R.H. (Hrsg.), *Metakognition, Motivation und Lernen*. Stuttgart: Kohlhammer, S. 9–21.

- Wild, K.-P. (1996): Beziehungen zwischen Belohnungsstrukturen der Hochschule, motivationalen Orientierungen der Studierenden und individuellen Lernstrategien beim Wissenserwerb. In: Lompscher, J./Mandl, H. (Hrsg.), *Lehr- und Lernprobleme im Studium*. Bern: Huber, S. 54–69.
- Wild, K.-P./Schiefele, U. (1994): Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. In: *Zeitschrift für Differenzielle und Diagnostische Psychologie* 15, S. 185–200.
- Zimmerman, B.J. (1994): Dimensions of academic self-regulation: A conceptual framework for education. In: Schunk, D.H./Zimmerman, B.J. (Hrsg.), *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale: Erlbaum, S. 3–24.
- Zimmerman, B.J./Martinez-Pons, M. (1990): Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. In: *Journal of Educational Psychology* 82, S. 51–59.

Anschrift der Autoren:

Dipl.-Päd. Claudia Leopold, Lehrstuhl für Lehr-Lernpsychologie, FB2, Universität Essen, Postfach, 45117 Essen.

Prof. Dr. Detlef Leutner, Zentrum für Lehr/Lern- und Bildungsforschung, Universität Erfurt, Postfach 900221, 99105 Erfurt.

Alexander Renkl/Silke Schworm

Lernen, mit Lösungsbeispielen zu lehren

1. Einleitung

Wie alle Projekte des DFG-Schwerpunktprogramms BIQUA strebt auch das vorliegende Projekt an, mathematisch-naturwissenschaftlichen Unterricht zu verbessern. Dieses Ziel kann über unterschiedliche Zugangsweisen verfolgt werden. An welchen Stellen, mit welcher Logik und Zielsetzung wir ansetzen, wird im Folgenden aufgezeigt.

Die prinzipielle Idee besteht darin, die Lernart „Lernen aus Lösungsbeispielen“, die experimentell sehr gut untersucht ist und die sich vielfach bewährt hat, Lehrenden näher zu bringen. Dies soll insbesondere über den Einsatz einer computer-basierten Lernumgebung erfolgen, in der Wissen über die Gestaltung von Lösungsbeispielen und deren produktive Verarbeitung durch Schüler mittels Selbsterklärungen vermittelt wird. Zudem sollen Feldstudien zum Transfer des im Lernprogramm erworbenen Wissens durchgeführt werden.

Um zu zeigen, wie sich dieses Projekt in das BIQUA-Schwerpunktprogramm einzuordnen ist, nehmen wir auf das mehrbenenanalytische Modell von Doll und Prenzel (in diesem Heft) Bezug. Unser Projekt setzt in erster Linie an den Lehrermerkmalen, speziell der Lehr-Expertise, an. Lehrenden soll Wissen über eine effektive Lernart, Lernen aus Lösungsbeispielen, vermittelt werden, das sie dann auch im Unterricht umsetzen können. Diese Umsetzung wird in Feldstudien untersucht. Insofern werden auch Unterrichtsprozesse mit einbezogen (siehe Doll/Prenzel in diesem Heft). Es wird explorativ analysiert, inwieweit bestimmte Kontextmerkmale des Unterrichts im Klassenzimmer (z.B. das Vorwissensniveau der Schüler, „Lerngewohnheiten“ der Schüler, Unterrichtsstil des Lehrers) die Anwendung der im Lernkontext erworbenen Kenntnisse behindern.

Auch der im mehrbenenanalytischen Modell (siehe Doll/Prenzel in diesem Heft) genannte Bereich der Lehr-Lernmaterialien wird durch das vorliegende Projekt berührt. Eine Grundvoraussetzung für effektives Lernen aus Lösungsbeispielen ist, dass gut gestaltete Beispiele bzw. gute Lernmaterialien im Unterricht verwendet werden. Die Lehrenden sollen im Lernprogramm zweierlei lernen: Sie sollen aus vorhandenen Lehrbüchern, Arbeitsblättern etc. geeignete Lösungsbeispiele auswählen können und sie sollen – wenn keine geeigneten Vorlagen zu finden sind – in der Lage sein, selbst günstige Lösungsbeispiele zu erstellen. Damit ist es auch ein Anliegen des vorliegenden Projekts, die Qualität der im Unterricht eingesetzten Materialien zu verbessern.

Als letzter wichtiger Bereich, an dem dieses Projekt ansetzt, sind die Lernstrategien der Schüler zu nennen. Dabei wird allerdings „nur“ eine spezielle Lernstrategie geför-

dert, nämlich Selbsterklärungen (zum effektiven Einsatz von Lernstrategien siehe Leopold/Leutner, in diesem Heft, bzw. zu Strategietraining siehe Gürtler u.a. in diesem Heft). Diese bezeichnen – zumindest im Kontext mathematisierter Inhaltsbereiche – Bemühungen von Lernenden, sich die Logik von Lösungsschritten klar zu machen. Lehrende sollen lernen, wie diese zu fördern sind. Wenn Lehrer bei ihren Schülern Selbsterklärungen fördern, ist dies für alle Inhaltsbereiche relevant, in denen auch Lösungsbeispiele verwendet werden können; dies sind insbesondere die Mathematik und die Physik. Zudem können Selbsterklärungen auch in produktiver Weise mit Problemlöseaktivitäten kombiniert werden (z.B. Neuman/Schwarz 1998).

Betrachtet man die Bereiche, auf die das vorliegende Projekt abzielt, so wird deutlich, dass keine radikale Veränderung der gängigen Unterrichtspraxis angestrebt wird. Eine „revolutionäre“ Zugangsweise würde radikale Veränderungen der gängigen Unterrichtspraxis anstreben (z.B. Einführung situierten Lernens oder problemorientierten Unterrichts) und dabei weitreichende Verbesserungen bei den Schülern anstreben etwa in den Facetten mathematisch-naturwissenschaftlicher Kompetenz, in Problemlösestrategien, in epistemologischen Überzeugungen, in Lernstrategien und in Motivation sowie Interesse (siehe mehrebenenanalytisches Modell von Doll/Prenzel in diesem Heft). Erstens besteht aber ein nicht ganz unerhebliches Problem darin, diese grundlegenden Veränderungen des Unterrichts zunächst einmal zu bewirken und dann aufrechtzuerhalten. Dies erfordert, wie entsprechende Erfahrungen zeigen, meist einen enormen Aufwand, der flächendeckend kaum betrieben werden kann (vgl. z.B. Cognition and Technology Group at Vanderbilt 1997). Zweitens fehlt es an stringenter empirischer Evidenz für die höhere Wirksamkeit dieser Unterrichtsmodelle.

Die hier gewählte „bescheidenere“, aber wohl realistischere Vorgehensweise setzt darauf, dass Lernen aus Lösungsbeispielen als ein Unterrichtsmodul vermehrt eingesetzt wird. Es sollen das fachbezogene Verständnis und die Lernstrategie „Selbsterklärung“ gefördert werden. Diese beiden Zielaspekte sind den Zielkriterien „Facetten der mathematisch-naturwissenschaftlichen Kompetenz“ und „Merkmale selbstregulierten Lernens im mathematisch-naturwissenschaftlichen Unterricht“ von Doll und Prenzel (in diesem Heft) zuzuordnen.

Zwei essenzielle Vorteile unserer Herangehensweise bestehen in folgenden Punkten: (1) Das Lernen aus Lösungsbeispielen ist empirisch sehr gut untersucht und seine Effektivität wurde vielfach belegt (für einen Überblick: Atkinson u.a. 2000). (2) Lernen aus Lösungsbeispielen ist relativ einfach im Unterricht zu implementieren, auch unter den zurzeit in Deutschland vorgegebenen Rahmenbedingungen schulischen Lernens.

Im Folgenden wird erläutert, was wir unter Lernen aus Lösungsbeispielen genau verstehen, warum es eine wichtige Lernart ist, und wie es effektiv zu gestalten ist. Die generelle Fragenstellung des vorliegenden Projekts sowie die spezifischen Fragen einer ersten experimentellen Studie werden anschließend vorgestellt. In dieser ersten Untersuchung wurden verschiedene Versionen eines computerbasierten Lernprogramms für Lehrkräfte zum Thema „Gestaltung von Lösungsbeispielen“ miteinander verglichen. Schließlich wird die Studie mit ihren Ergebnissen und Implikationen beschrieben.

2. Lernen aus Lösungsbeispielen

Lösungsbeispiele bestehen aus einer Problemstellung, Lösungsschritten und der endgültigen Lösung selbst. Derartige Beispiele findet man typischerweise in Mathematik- und Physikbüchern. Sie folgen meist der Erklärung oder Erarbeitung eines Prinzips oder eines Gesetzes. Das Lösungsbeispiel soll aufzeigen, wie das eingeführte Prinzip oder Gesetz angewandt werden kann. Nach dem Beispiel folgen meist Aufgaben, die von den Lernenden selbst gelöst werden sollen.

Unter Lernen aus Lösungsbeispielen versteht man aber nicht die üblicherweise kurze Lernphase zwischen der Behandlung eines Prinzips/Gesetzes o.Ä. und dem Lösen von Rechenaufgaben, in der meist nur ein Lösungsbeispiel behandelt wird. Es werden vielmehr die Lernprozesse und -effekte betrachtet, die auftreten, wenn statt nur eines Lösungsbeispiels mehrere verwendet werden und die Phase des Lösungsbeispielstudiums also verlängert wird. In der Regel ist ein derartiges Lernen aus Lösungsbeispielen dem üblichen Vorgehen (ein Lösungsbeispiel, dann Aufgaben) überlegen. Dies konnte eine Vielzahl von Untersuchungen zeigen (vgl. z.B. Sweller/van Merriënboer/Paas 1998). Dabei kann aus Lösungsbeispielen die „einfache“ Anwendung von Algorithmen oder Ähnlichem erlernt und darüber hinaus ein Verständnis der Lösungslogik erreicht werden, was Transferleistungen fördert.

Eine Erklärung für die Effektivität des Lernens aus Lösungsbeispielen besteht darin, dass bei einem „verfrühten“ Lernen durch Problemlösen viele kognitive Ressourcen durch Problemlöseaktivitäten gebunden werden. Dadurch steht weniger mentale Kapazität für den eigentlichen Lernprozess zur Verfügung, z.B. bei der Erarbeitung von Verständnis oder der Konstruktion abstrakter Schemata. Beim Lernen aus Lösungsbeispielen dagegen entfällt die Anforderung des Problemlösens. Es stehen für den eigentlichen Lernprozess ausreichend kognitive Ressourcen zur Verfügung (Weber u.a. 2001). Natürlich ist der Einsatz des Lernens aus Lösungsbeispielen kein Garant für hohen Lernerfolg. Wie viel im konkreten Fall aus Lösungsbeispielen gelernt wird, hängt insbesondere von den beiden folgenden Aspekten ab (Atkinson u.a. 2000): (1) Selbsterklärungen und instruktionale Erklärungen und (2) Gestaltung der Lösungsbeispiele. Diese beiden Faktoren werden in den folgenden Abschnitten diskutiert.

2.1 Selbsterklärungen und instruktionale Erklärungen

Lernende profitieren nur dann substanziell vom Lernen aus Lösungsbeispielen, wenn sie sich klarmachen, welches Rationale hinter den präsentierten Lösungsschritten steht. Dieses sollten sie sich optimalerweise selbst erklären und dadurch verständlich machen (Chi u.a. 1989; Renkl 1997; Renkl u.a. 1998). Es ist insbesondere förderlich, wenn Lernende sich den Sinn/Zweck einzelner Operationen vergegenwärtigen. Dies heißt, dass sie sich bewusst machen, welches (Zwischen-)Ziel mit einer Operation erreicht wird und/oder welches Prinzip bzw. welche Regel einer Operation zu Grunde liegt.

Es hat sich gezeigt, dass Selbsterklärungen von großer Bedeutung sind, während instruktionale Erklärungen (z.B. eines Lehrers, Tutors oder Lernprogramms) oftmals weniger wirksam sind, als ihre weite Verbreitung erwarten lässt (z.B. Stark u.a. 2001). Es ist vielfach günstiger, Selbsterklärungen anzuregen, als instruktionale Erklärungen zu geben (z.B. Chi 1996; zu einer ausführlichen Diskussion der Gründe für die oft mangelnde Effektivität von instruktionalen Erklärungen siehe Renkl 2002). Gleichwohl ist festzuhalten, dass ein alleiniges „Bauen“ auf Selbsterklärungsaktivitäten auch etliche Nachteile hat. Es kommt beispielsweise vor, dass die Lernenden sich bestimmte Schritte nicht erklären können oder dass sie falsche Selbsterklärungen geben (Renkl 2002).

Vor diesen Hintergrund entwickelte Renkl (2002) ein Set instruktionaler Prinzipien, mit denen sichergestellt werden sollte, dass die spontane Selbsterklärungsaktivität in produktiver Weise durch instruktionale Erklärungen unterstützt wird. Zwei zentrale Prinzipien sind dabei die Priorität der Selbsterklärungsaktivität (instruktionale Erklärungen lediglich als „Notfallhilfe“) und der Abruf der instruktionalen Erklärungen auf Lerneranforderung (d.h. keine von den Lernenden unerwünschte Vorgabe). Empirisch zeigte sich, dass derartige instruktionale Erklärungen den durchschnittlichen Lernerfolg steigerten (einschränkend muss jedoch festgehalten werden, dass dieser Effekt in erster Linie auf eine Untergruppe von Lernenden zurückging, die ein effektives Nutzungsmuster der instruktionalen Erklärungen zeigte).

Instruktionale Erklärungen der von Renkl (2002) eingesetzten Art können in Bezug auf Selbsterklärungen aber nicht nur positive, ergänzende und korrigierende Funktionen haben. Aus der Rückmeldefunktion von instruktionalen Erklärungen zu Selbsterklärungen kann auch ein Problem auftreten, das seit Langem in der Feedback-Forschung bekannt ist (z.B. Kulhavy 1977). Wenn Feedback, das die richtige Antwort enthält, leicht verfügbar ist, reduzieren Lernende typischerweise ihre Anstrengung, selbst auf die Antwort zu kommen. Sie neigen dazu, die richtige Lösung nachzusehen, statt sie selbst zu bestimmen, was den Lernerfolg beeinträchtigt. Dies bedeutet, dass die Bereitstellung instruktionaler Erklärungen die Selbsterklärungsaktivität reduzieren kann (im Sinne von „Ich kann es mir ja erklären lassen“).

Als Fazit ist festzuhalten, dass die förderliche Rolle von Selbsterklärungen für das Lernen aus Lösungsbeispielen als gesichert angesehen werden kann. Instruktionale Erklärungen können, müssen das Lernen aber nicht unterstützen.

2.2 *Gestaltung der Lösungsbeispiele*

Selbstverständlich ist es von Bedeutung, wie die einzelnen Lösungsbeispiele gestaltet sind (*Intra-Beispiel-Merkmale*). Im Kontext einer ersten Studie des vorliegenden Projekts wurde vor allem aufgegriffen, dass Lösungsbeispiele, die grafische und textuelle Informationen enthalten (z.B. in der Geometrie oder der Physik), vielfach ungünstig gestaltet sind (z.B. Tarmizi/Sweller 1988). Wenn ein Lösungsbeispiel aus mehreren Informationsquellen besteht, wird ein Abgleich notwendig, der kognitive Kapazität in Anspruch nimmt. Sie steht dann für den Lernprozess nicht mehr zur Verfügung. Dieser ne-

gative Effekt lässt sich vermeiden, indem die entsprechenden Lösungsbeispiele in ein *integriertes Format* überführt werden, d.h. Informationen aus verschiedenen Quellen gewissermaßen in eine Quelle integriert werden (Mwangi/Sweller 1998; Ward/Sweller 1990).

Neben der Gestaltung einzelner Lösungsbeispiele ist es für den Lernerfolg auch von Bedeutung, in welcher Form verschiedene Beispiele miteinander kombiniert werden (*Inter-Beispiel-Merkmale*). In der bereits erwähnten ersten Studie wird der Aspekt der *strukturbetonenden Beispielsequenz* fokussiert. Es handelt sich dabei um eine Sequenz von Beispielen, die insbesondere aufzeigt, dass bei Aufgaben trotz ähnlicher Oberflächenmerkmale (z.B. Zahlen, Gegenstände, grafische Elemente oder Variablenbezeichnungen) eine unterschiedliche Problemstruktur (Lösungslogik) zugrunde liegen kann. Damit wird das Problem aufgegriffen, dass Lernende die Ähnlichkeit von Aufgaben häufig nicht danach beurteilen, inwieweit ihnen die selben strukturellen Merkmale zu Grunde liegen, sondern hinsichtlich Übereinstimmungen zwischen den klar ersichtlichen Oberflächenmerkmalen. Als Konsequenz verwenden sie häufig einen falschen Lösungsweg, den sie von einer oberflächlich ähnlichen, aber strukturell unterschiedlichen Aufgabe kennen. Bei der Behandlung verschiedener, aufeinander bezogener Problemtypen (z.B. proportionale und antiproportionale Dreisatzaufgaben) sollen mehrere Lösungsbeispiele so zusammengestellt werden, dass lösungsrelevante, strukturelle Merkmale hervortreten. Dies wird erreicht, indem erstens in verschiedenen Beispielen eines jeweiligen Problemtyps unterschiedliche inhaltliche Einbettungen verwendet werden. Zweitens werden dann dieselben bzw. sehr ähnliche inhaltliche Einbettungen für einen anderen Problemtyp verwendet. Damit kann gezeigt werden, dass trotz ähnlicher Oberflächenmerkmale ganz unterschiedliche Strukturen zu Grunde liegen können (Quilici/Mayer 1996).

Natürlich gibt es noch zahlreiche andere Beispielmerkmale, die die Effektivität des Lernens aus Lösungsbeispielen beeinflussen. In einer ersten empirischen Studie dieses Projekts bzw. in dem Modul eines Computerlernprogramms, das in dieser Studie zum Einsatz kam, wurden jedoch vorerst exemplarisch nur das integrierte Format und die strukturbetonende Beispielsequenz als Lösungsbeispielmerkmale betrachtet.

3. Generelle Fragestellungen und Arbeiten des Projekts

Wie bereits erläutert, bezieht sich die Hauptfragestellung des Projekts darauf, wie man (künftigen) Lehrkräften Wissen über die effektive Gestaltung von Lösungsbeispielen vermitteln kann (zur Kompetenzerweiterung von Lehrern vgl. auch Fischler u.a. in diesem Heft). Damit könnten gut gesicherte Befunde aus der psychologischen Lehr-Lernforschung für die alltägliche Schulpraxis fruchtbar gemacht werden. Damit wäre ein substanzieller Beitrag zur Verbesserung der Qualität des mathematisch-naturwissenschaftlichen Unterrichts geleistet.

Um dieses Ziel zu erreichen wird eine computer-basierte Lernumgebung erstellt, in der Lehrende Wissen darüber erwerben können, wie Lösungsbeispiele zu gestalten sind

und wie Schüler dazu angehalten werden können, die Lösungsbeispiele so zu verarbeiten, dass ihr Verständnis vertieft wird. Experimentelle Arbeiten dienen dazu, die Lernumgebung zu optimieren. Feldstudien gehen der Frage nach, inwieweit es Lehrenden gelingt, das über Lernen aus Lösungsbeispielen erworbene Wissen auch im Klassenzimmer einzusetzen. Auf der Basis der Feldstudienbefunde werden ggf. zusätzlich instruktionale Maßnahmen zur Transferförderung ergriffen.

4. Ausgesuchte Fragestellungen eines ersten Experiments

Im Rahmen dieses Artikels können nur exemplarisch Fragestellungen einer empirischen Arbeit und entsprechende Befunde dargestellt werden. Es wurde ein erstes Modul eines computer-basierten Lernprogramms erstellt, in dem Lehrende Wissen über das integriertes Format und über strukturbetonende Beispielsequenzen erwerben konnten. Das Programm wurde verschiedenen Gruppen von Lehramtstudierenden vorgegeben (didaktisch versus fachwissenschaftlich orientiertes Studium), um zu erkunden, inwieweit es für unterschiedliche Zielgruppen geeignet ist.

In diesem Programmmodul werden Beispiele für günstig und ungünstig gestaltete Lösungsbeispiele vorgestellt. Die Lehrenden lernen also aus Beispielen von Lösungsbeispielen. Damit wird ein „pädagogischer Doppeldecker“ realisiert, d.h. der Lerninhalt stimmt mit der Lernmethode überein. Dieses Prinzip hat sich für Lehrerbildung als ein entscheidender Faktor erwiesen (Borko/Putnam 1996).

Wie bereits ausgeführt, ist beim Lernen aus Lösungsbeispielen die Selbsterklärungsaktivität der Lernenden von großer Bedeutung. Neben der Förderung von Selbsterklärungsaktivität können, um den Lernprozess zu unterstützen, instruktionale Erklärungen implementiert werden (Renkl 2002). Bei den Beispielen von (un-) günstig gestalteten Lösungsbeispielen in unserer computerbasierten Lernumgebung handelt es sich aus der Perspektive der Lehrenden jedoch nicht um Lösungsbeispiele im eigentlichen Sinn. Die Lehrkräfte sollen Wissen über die Gestaltung von Lösungsbeispielen erwerben. Für die Lösungsbeispielgestaltung können jedoch keine Lösungsschritte angegeben werden. Wir nennen Beispiele, bei denen keine Lösungsschritte vorhanden sind, gelöste Beispielprobleme. Es war zu untersuchen, ob bei diesem Beispieltyp Aufforderungen zur Selbsterklärung und die Bereitstellung von instruktionalen Erklärungen ebenso lernförderlich sind, wie sie es erwiesenermaßen bei Lösungsbeispielen (mit Problemstellung, Lösungsschritten und endgültiger Lösung) sind.

Spezifisch werden in dieser Arbeit die Befunde zu folgenden Forschungsfragen dargestellt:

- 1) Wirken sich, wie erwartet, Aufforderungen zur Selbsterklärung positiv auf die Lernleistung aus?
- 2) Gibt es einen positiven Effekt der instruktionalen Erklärungen auf die Lernleistung?
- 3) Reduziert die Verfügbarkeit instruktionaler Erklärungen die Selbsterklärungsaktivität und schlägt sich dies in der Lernleistung nieder?

5. Methode

5.1 Stichprobe und Design

Es nahmen 80 Lehramtstudierende an einem 2x2-faktoriellen Experiment teil (pro Zelle $N = 20$): Faktor 1: Aufforderungen zur Selbsterklärung (mit und ohne), Faktor 2: instruktionale Erklärungen (mit und ohne). 47 Probanden waren Studierende der Pädagogischen Hochschule, 33 absolvierten ein Universitätsstudium. Während an der Pädagogischen Hochschule ein Schwerpunkt auf den pädagogisch-didaktischen Teilen der Lehrerausbildung gelegt wird, liegt beim Universitätsstudium der angehenden Gymnasiallehrer der Schwerpunkt im fachwissenschaftlichen Bereich. Bei der Analyse der in Reaktion auf die Selbsterklärungsaufforderungen vorgenommenen schriftlichen Elaborationen fehlten sechs Lernende (jeweils drei pro Zelle), bei denen auf Grund technischer Probleme keine entsprechenden Daten vorlagen.

5.2 Die Lernumgebung

Das Programmmodul enthielt eine kurze Einführung in das Lernen mit Lösungsbeispielen. Anschließend wurden Beispiele für Lösungsbeispiele bzw. für Sets von Lösungsbeispielen vorgegeben. Inhaltlich kamen Lösungsbeispiele aus der Geometrie und aus mathematisierten Bereichen der Physik zum Einsatz. Durch die Verwendung unterschiedlicher Inhaltsgebiete, aus denen die Lösungsbeispiele entnommen wurden, sollte die Transferierbarkeit der erworbenen Kenntnisse über verschiedene Bereiche der Mathematik und Physik hinweg erhöht werden.

Abbildung 1 zeigt eine exemplarische Seite aus dem Lernprogramm, wie es in der Gruppe mit Aufforderungen zur Selbsterklärung und zusätzlichen instruktionalen Erklärungen eingesetzt wurde. Die Bildschirmseite zeigt ein gelöstes Beispielproblem zum integrierten Format. Das gelöste Beispielproblem besteht aus zwei Lösungsbeispielen. Das linke Beispiel wird in einem fraktionierten, das rechte Beispiel in einem integrierten Format dargeboten. Beide Beispiele enthalten dieselbe Informationsmenge. Im fraktioniert formatierten Beispiel erfordert die Zuordnung von Information aus Rechnung, Text und Grafik ein hohes Maß an kognitiven Ressourcen. Diese Aufmerksamkeit steht nicht mehr für das Verständnis der Strahlensatzaufgabe zur Verfügung.

Die *Aufforderungen zur Selbsterklärung* bestanden aus Aufforderungen, schriftlich in entsprechende Textfelder Begründungen dafür einzugeben, warum bestimmte Lösungsbeispiele günstig oder auch ungünstig sind (siehe Abb. 1). Die *instruktionalen Erklärungen*, die von den Lernenden bei Bedarf aufgerufen werden konnten, waren gleichsam Antworten auf die Aufforderungen zur Selbsterklärung. Sie waren über die Schaltfläche, die das Portrait eines fiktiven, im Programm eingeführten Expertenlehrers zeigte, aufzurufen (siehe Abb. 1). Entsprechend einschlägigen Befunden zur Multimedia-Forschung (z.B. Mayer 1997) wurden die instruktionalen Erklärungen akustisch dargeboten.

Die Zeichnung zeigt die sogenannte Schattenmethode zur Bestimmung der Höhe eines Baums. Verglichen wird hier die Schattenlänge des Baums mit der eines Wanderers. Bestimme die Höhe des Baums.

Fraktioniertes Format



Der Schatten des Baums (c+d) ist 25m lang. Der Wanderer (a) wirft bei einer Größe von 1,80m einen Schatten (c) von 1,50m.

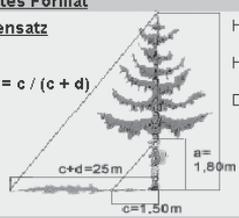
1 Strahlensatz:
 $a / b = c / d$
 oder
 $a / (a + b) = c / (c + d)$

$H = a * (c + d) / c$
 $H = 1,80 * 25 / 1,5$
 $H = 30$

Der Baum ist 30 m hoch.

Integriertes Format

1. Strahlensatz
 $a / b = c / d$
 oder
 $a / (a + b) = c / (c + d)$



$H = a * (c + d) / c$
 $H = 1,80 * 25 / 1,50$
 Der Baum ist 30 m hoch.

Bei welchem Beispiel fällt es Michaela leichter, die gegebenen und gesuchten Strecken in der Abbildung zu identifizieren? Warum?

Bei welchem Beispiel kann Michaela besser über die Abbildung ein Verständnis über den Zusammenhang der Größen a, b, c und d bekommen? Warum?




Abb. 1: Exemplarische Seite eines gelösten Beispielproblems mit Aufforderung zur Selbsterklärung und Möglichkeit, über eine Schaltfläche eine instruktionale Erklärung aufzurufen.

5.3 Versuchsablauf

Die Untersuchung erfolgte in Einzelsitzungen, die etwa drei Stunden in Anspruch nahmen. Das Lernprogramm begann mit einem Lehrtext zu den grundlegenden Prinzipien des Einsatzes, der Gestaltung und der Kombination von Lösungsbeispielen im Unterricht. Anschließend wurden den Lehramtstudierenden im Lernprogramm günstige und ungünstige Lösungsbeispiele dargeboten. Merkmale des Lernprozesses wurden bei den Gruppen mit Aufforderungen zur Selbsterklärung über das Aufzeichnen der schriftlichen Eingaben erfasst. Ein Nachtest erfasste den Lernerfolg.

5.4 Materialien und Instrumente

Die Auswertung der schriftlichen Selbsterklärungen

Bei der Bearbeitung des Lernprogramms wurden die Lernenden der Bedingungen, die Aufforderungen zur Selbsterklärung beinhalteten, insgesamt dreizehnmal aufgefordert, ihre Gedanken in einem dafür vorgesehenes Textfeld zu notieren. Die schriftlichen

Selbsterklärungen wurden nach einem spezifisch entwickelten Kodiersystem ausgewertet (zu Details siehe Schworm/Renkl 2002). In dieser Arbeit wird lediglich ein Gesamtwert für die Elaborationsaktivität verwendet.

Nachtest: Erfassung der Lernleistung

Lehrende stehen vor der Anforderung, aus Lehrmaterial (Schulbüchern, Arbeitsblättern etc.) Geeignetes auszusuchen. Falls keine guten Materialien vorhanden sind, sollten Lehrende diese selbst erstellen können. Deshalb bestand der Nachtest aus zwei Arten von Aufgaben. Erstens mussten aus vorgegebenen Lösungsbeispielen geeignete ausgewählt (integriertes Format) bzw. zusammengestellt (strukturbetonende Beispielsequenz) werden. Zweitens hatten die Teilnehmenden eine strukturbetonende Beispielsequenz zu generieren, wobei die einzelnen Beispiele im integrierten Format zu gestalten waren; die entsprechende Leistung wurde nach vordefinierten Kriterien bewertet.

6. Ergebnisse

Insgesamt konnten im Nachtest 34 Punkte erreicht werden, das Gesamtmittel lag bei 23,20 ($s = 5,67$). Es ergaben sich damit weder Boden- noch Deckeneffekte. Eine qualitative Inspektion der Nachtestdaten zeigte, dass die meisten Teilnehmenden die prinzipiellen Aspekte effektiver Beispielgestaltung verstanden hatten. Schwierigkeiten gab es bei der präzisen Umsetzung, insbesondere wenn Lösungsbeispiele selbst generiert werden mussten.

Die beiden Unterbereiche des Nachtests wurden zu einem Gesamtwert zusammengefasst, da sich vergleichbare Ergebnismuster zeigten und die Werte zudem signifikant positiv korrelierten ($r = .31$; $p < .05$).

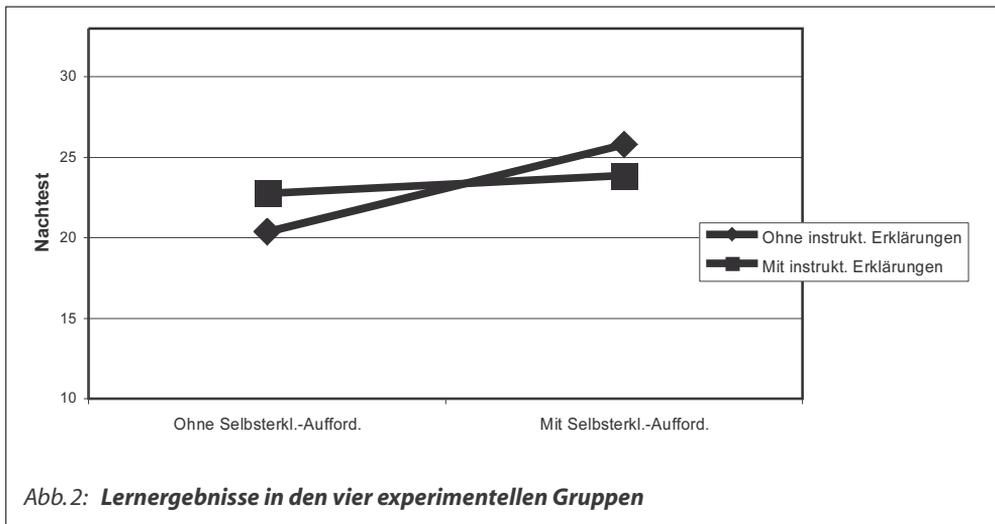
Tab. 1: Mittelwerte (Standardabweichungen) des Nachtests und der schriftlichen Elaborationen in den vier experimentellen Gruppen

	Nachtest	Schriftliche Elaborationen
Ohne Selbsterklärungsaufford./ ohne instruktionale Erklärungen	20,36 (5,46)	–
Mit Selbsterklärungsaufford./ ohne instruktionale Erklärungen	25,80 (4,44)	19,82 (4,51)
Ohne Selbsterklärungsaufford./ mit instruktionale Erklärungen	22,75 (5,30)	–
Mit Selbsterklärungsaufford./ mit instruktionale Erklärungen	23,86 (6,29)	16,88 (4,51)

Tabelle 1 zeigt die Mittelwerte und Standardabweichungen des Nachtests und der schriftlichen Elaborationen. Da die Studierenden der Universität bedeutsam besser abschnitten als die Studierenden der Pädagogischen Hochschule, wurde die Variable

(„Hochschule: Fachwissenschaftlicher Schwerpunkt versus Didaktischer Schwerpunkt“) als Kovariate einbezogen. Tests auf Steigungsunterschiede zwischen Kovariate und Lernerfolg in den einzelnen experimentellen Gruppen ergaben keine bedeutsamen Unterschiede (alle $F_s < 1$), sodass die entsprechende Voraussetzung für die Kovarianzanalyse gegeben war.

Für die Lernleistung ergab sich ein signifikanter Haupteffekt „Aufforderung zur Selbsterklärung“ ($F_{(1,75)} = 8,68$; $p < .05$) von mittlerer bis starker praktischer Signifikanz ($\eta^2 = 0.11$) und ein signifikanter Interaktionseffekt ($F_{(1,75)} = 4,91$; $p < .05$) von mittlerer praktischer Signifikanz ($\eta^2 = 0.06$). Es zeigte sich kein Haupteffekt „Bereitstellen von instruktionalen Erklärungen“ ($F_{(1,75)} = 0,37$; $p > 0.1$; $\eta^2 < .01$). Abbildung 2 illustriert den Interaktionseffekt.



Die Gruppe ohne Aufforderungen zur Selbsterklärung und ohne instruktionale Erklärungen schnitt erwartungsgemäß am schlechtesten ab. Die zusätzliche Bereitstellung instruktionaler Erklärungen führte hier zu verbesserten Lernleistungen. Am besten schnitt allerdings die Gruppe mit Aufforderungen zur Selbsterklärung, aber ohne instruktionale Erklärungen ab. Im Falle von Aufforderungen zur Selbsterklärung war die zusätzliche Möglichkeit, instruktionale Erklärungen aufzurufen, für den Lernerfolg abträglich.

Wie bei den theoretischen Vorüberlegungen bereits erwähnt, könnte die Möglichkeit, instruktionale Erklärungen aufzurufen, die Selbsterklärungsbemühungen reduziert haben („man kann sich ja die Erklärungen holen“). Tatsächlich ergab sich in der Gruppe mit Selbsterklärungsaufforderung und instruktionaler Erklärung gegenüber der Gruppe, die lediglich Selbsterklärungsaufforderungen bekam, eine niedrigere Selbsterklärungsaktivität: durchschnittlich knapp 17 versus knapp 20 Elaborationen. Dieser Unterschied war statistisch signifikant ($t_{(32)} = 1,72$; $p < .05$; einseitig) und von mittlerer praktischer Signifikanz ($\eta^2 = 0.08$).

7. Diskussion

Die vorliegenden Befunde lassen sich wie folgt zusammenfassen: Für den Lernerfolg ist es am besten, die Lernenden zur Selbsterklärung aufzufordern. Instruktionale Erklärungen, zumindest wenn sie in Kombination mit den Selbsterklärungsaufforderungen gegeben werden, sind hier eher abträglich. Dies liegt daran, dass sie in diesem Fall die Selbsterklärungsaktivität und damit den Lernerfolg reduzieren. Interessanterweise war also eine Lernumgebung am günstigsten, die ausschließlich auf eine Förderung von Eigenaktivität in der Verarbeitung des Lernmaterials setzte.

Die vorliegenden Befunde werfen Fragen bzgl. des Stellenwerts instruktionaler Erklärungen auf. Da sie beim *Nicht-Vorhandensein von Selbsterklärungsaufforderungen* den Lernerfolg förderten, waren sie prinzipiell durchaus sinnvoll gestaltet und hatten das Potenzial, Lernen zu unterstützen. Die gefundene Effektivität der instruktionalen Erklärungen stimmt in diesem Fall mit den entsprechenden Befunden der Studie von Renkl (2002) überein, in der keine Selbsterklärungsaufforderungen eingesetzt wurden. In der vorliegenden Studie zeigte sich dagegen, dass *in Kombination mit Aufforderungen zur Selbsterklärung* die zusätzliche Bereitstellung instruktionaler Erklärungen abträglich war. Ob dies immer so ist oder ob doch günstige Kombinationen von Selbsterklärungsaufforderungen und instruktionalen Erklärungen gefunden werden können, ist in weiteren Studien zu klären (z.B. ein allmähliches Übergehen von instruktionalen Erklärungen zu Selbsterklärungen).

Schließlich ist als offene Frage herauszustellen, inwieweit sich bei der Arbeit mit dem vorliegenden Lernprogramm (bzw. mit einer um weitere Gestaltungsmerkmale erweiterten Version) „nur“ Binneneffekte (Lernleistung in einem Nachtest) ergeben, oder ob das erworbene Wissen auch ins Klassenzimmer transferiert werden kann. Zu diesem Zweck sind Feldstudien geplant, bei denen einzelne Lehrer gebeten werden, das im Programm Erlernte im Unterricht zu implementieren. Der entsprechende Unterricht soll dabei beobachtet und die Lehrer sollen zu Schwierigkeiten bei der Umsetzung interviewt werden. Die dabei gewonnenen Erkenntnisse sollen wiederum helfen, das Lernprogramm zu optimieren bzw. zusätzliche Transferfördermaßnahmen zu ergreifen.

Weiterhin wird im Projekt die Thematik der Selbsterklärungen von Schülern angegangen werden. Es wird Lehrenden vermittelt, wie die Schüler in produktiver Weise angehalten werden können, sich Lösungsbeispiele selbst so zu erklären, dass daraus „Verstehen“ entsteht.

Literatur

- Atkinson, R.K./Derry, S.J./Renkl, A./Wortham, D.W. (2000): Learning from examples: Instructional principles from the worked examples research. In: Review of Educational Research 70, S. 181–214.
- Borko, H./Putnam, R.T. (1996): Learning to teach. In: Berliner, D.C./Calfee, R.C. (Hrsg.) Handbook of Educational Psychology. New York: Macmillan, S. 673–708.
- Chi, M.T.H. (1996): Constructing self-explanations and scaffolded explanations in tutoring. In: Applied Cognitive Psychology 10, S. 33–49.

- Chi, M.T./Bassok, M./Lewis, M.W./Reimann, P./Glaser, R. (1989): Self-explanations: How students study and use examples in learning to solve problems. In: *Cognitive Science* 13, S. 145–182.
- Cognition and Technology Group at Vanderbilt. (1997): *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development*. Hillsdale, NJ: Erlbaum.
- Doll, J./Prenzel, M. (2002). Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. In: *Zeitschrift für Pädagogik* (in diesem Band).
- Fischler, H./Zedler, P./Schröder, H.-J./Tonhäuser, C. (2002). Unterrichtsskripts und Lehrerexpertise: Bedingungen ihrer Modifikation. In: *Zeitschrift für Pädagogik* (in diesem Band).
- Gürtler, T./Perels, F./Schmitz, B./Bruder, R. (2002). Training zur Förderung selbstregulativer Fähigkeiten in Kombination mit Problemlösen in Mathematik. In: *Zeitschrift für Pädagogik* (in diesem Heft).
- Kulhavy, R.W. (1977): Feedback in written instruction. In: *Review of Educational Research* 47, S. 211–232.
- Leopold C./Leutner, D. (2002): Der Einsatz von Lernstrategien in einer konkreten Lernsituation bei Schülern unterschiedlicher Jahrgangsstufen. In: *Zeitschrift für Pädagogik* (in diesem Heft).
- Mayer, R.E. (1997): Multimedia learning: Are we asking the right questions. In: *Educational Psychologist* 32, S. 1–19.
- Mwangi, W./Sweller, J. (1998): Learning to solve compare word problems: The effect of example format and generating self-explanations. In: *Cognition and Instruction* 16, S. 173–199.
- Neuman, Y./Schwarz, B. (1998): Is self-explanation while solving problems helpful? The case of analogical problem solving. In: *British Journal of Educational Psychology* 68, S. 15–24.
- Quilici, J.L./Mayer, R.E. (1996): Role of examples in how students learn to categorize statistics word problems. In: *Journal of Educational Psychology* 88, S. 144–161.
- Renkl, A. (1997): Learning from worked-out examples: A study on individual differences. In: *Cognitive Science* 21, S. 1–29.
- Renkl, A. (2002): Learning from worked-out examples: Instructional explanations supplement self-explanations. In: *Learning & Instruction* 12, S. 149–176.
- Renkl, A./Stark, R./Gruber, H./Mandl, H. (1998): Learning from worked-out examples: The effects of example variability and elicited self-explanations. In: *Contemporary Educational Psychology* 23, S. 90–108.
- Schworm, S./Renkl, A. (2002): Lernen effektive Lösungsbeispiele zu erstellen: Ein Experiment zu einer computer-basierten Lernumgebung für Lehrende. In: *Unterrichtswissenschaft* 30, 7–26.
- Stark, R./Gruber, H./Mandl, H./Hinkofer, L. (2001): Wege zur Optimierung eines beispilsbasierten Instruktionsansatzes: Der Einfluss multipler Perspektiven und instruktionaler Erklärungen auf den Erwerb von Handlungskompetenz. In: *Unterrichtswissenschaft* 29, S. 26–37.
- Sweller, J./van Merriënboer, J.J.G./Paas, F.G.W.C. (1998): Cognitive architecture and instructional design. In: *Educational Psychology Review* 10, S. 251–296.
- Tarmizi, R. A./Sweller, J. (1988): Guidance during mathematical problem solving. In: *Journal of Educational Psychology* 80, S. 424–436.
- Ward, M./Sweller, J. (1990): Structuring effective worked examples. In: *Cognition and Instruction* 7, S. 1–39.
- Weber, S./Renkl, A./Gruber, H./Schweizer, K./Hillebrand, M./Wittenzeller, M./Lerche, T. (2001): Cognitive Load beim Lernen aus Lösungsbeispielen. Vortrag auf der „61. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF)“ in Schwäbisch-Gmünd, September 2001.

Anschrift der Autoren:

Prof. Dr. Alexander Renkl, Psychologisches Institut, Abteilung Pädagogische Psychologie, Universität Freiburg, Engelbergerstr. 41, 79085 Freiburg.

Dipl.-Psych. Silke Schworm, Psychologisches Institut, Abteilung Pädagogische Psychologie, Universität Freiburg, Engelbergerstr. 41, 79085 Freiburg.

Teil IV:

Diagnose und Förderung von Interessen und Lernmotivation

Förderung des Interesses und der Motivation von Schülerinnen
und Schülern für mathematisch-naturwissenschaftliche Fächer:
Zum Einfluss schulischer und familiärer Lehr-Lernumgebungen

Elke Wild

Einleitung

Internationale Vergleichsstudien wie TIMSS (z.B. Baumert/Bos/Lehmann 2000) und PISA (PISA-Konsortium 2001) sollen insbesondere dadurch, dass sie eine detaillierte Bestandsaufnahme der durchschnittlichen Leistungsfähigkeit von Schülerpopulationen geben, Hinweise auf die Qualität von Bildungssystemen liefern. So richtet sich auch die in der Öffentlichkeit und Bildungspolitik geführte Diskussion vornehmlich auf die Frage, wie die Kompetenzen deutscher Schülerinnen und Schüler gefördert werden können. In diesem Kontext wird die Lernmotivation der Schüler bereits indirekt über ihre lern- und leistungsrelevante Funktion bedeutsam (zusf. Wild/Hofer/Pekrun 2001; Krapp 1999; Schiefele 1996). Weil mit der Lernmotivation zwar keine hinreichende aber eben doch notwendige Bedingung für qualitativ hochwertige Leistungen angesprochen ist, und weil in Deutschland der Anteil der Schüler, die dem Lesen nichts Positives abgewinnen können, größer ist als in sämtlichen anderen teilnehmenden Staaten (vgl. PISA-Konsortium 2001), können die ernüchternden Befunde zur Leistungsfähigkeit deutscher Schülerinnen und Schüler auch als Hinweis auf problematische Motivlagen interpretiert werden.

Diese motivationalen Defizite bekommen noch größeres Gewicht, wenn in Rechnung gestellt wird, dass im schulischen Kontext angesichts des schnellen Anwachsens und der ebenso rasch sinkenden Halbwertszeit von Wissen lediglich grundlegende kognitive Fertigkeiten vermittelt werden können und Heranwachsende dieser und nachkommender Generationen gezwungen sein werden, sich auch nach Abschluss ihrer Schulzeit autodidaktisch weiterzubilden. Wenn somit eine grundsätzliche Lernbereitschaft in modernen Informationsgesellschaften wie der unseren als eine Zielgröße pädagogischer Bemühungen betrachtet werden muss, dann stimmen Befunde bedenklich, wonach mit dem Absinken der Lernmotivation im Verlauf der Schulzeit zwar ein in vielen Ländern, vor allem in Deutschland, zu beobachtendes Phänomen angesprochen ist (Eccles u.a. 1993; Pekrun 1993; Helmke 1993; Fend 1997; Wild/Hofer 2001).

Bei der Suche nach den Ursachen für die hierzulande stark abfallende Lernbereitschaft und der Analyse der lern- und leistungsbezogenen Konsequenzen erscheint es fruchtbar, konzeptuelle Differenzierungen aufzugreifen, die vor allem in den letzten Jahren innerhalb der pädagogisch-psychologischen Motivationsforschung herausgearbeitet wurden (Wild/Hofer/Pekrun 2001; Rheinberg 2002; Pintrich/Schunk 1996; Schiefele 1996). Gerade diese gleichzeitige Betrachtung qualitativ verschiedener Formen der Lernmotivation kennzeichnet die drei in BIQUA angesiedelten Projekte mit motivationspsychologischem Schwerpunkt. Wegen dieser Besonderheit, aber auch wegen der Überschneidungen und Unterschiede zwischen den Projekten (vgl. Tabelle 1) versprechen sie über den vorliegenden Forschungsstand hinausgehende, praxisrelevante Informationen zu liefern.

Tab. 1: **Parallelen und Unterschiede zwischen den BIQUA-Projekten mit motivationspsychologischem Hintergrund**

	Potsdamer Projekt	Münsteraner Projekt	Bielefelder Projekt
Theoretischer Hintergrund	Erweitertes kognitives Motivationsmodell	Interessen- und Selbstbestimmungstheorie, Einstellungsänderungsmodell	Interessen- und Selbstbestimmungstheorie, Zieltheorien, familienpsychologische Ansätze
Zentrale Konstrukte (abhängige Variablen)	Erwartungen/Nutzenüberzeugungen, Selbstregulationskomponenten, Tätigkeits- und Folgenanreize, Leistungen (Noten)	Interessen, Desinteressen, Abneigungen, Einstellungen zu Schule/ zum naturwissenschaftlichen Unterricht	Mathematikbezogene Interessen und Abneigungen, motivationale Orientierungen sowie lernbezogene Ziele
Zentrale Kontextfaktoren (unabhängige Variablen)	Ausgewählte Lehrermerkmale (Bezugsnormorientierung) und Merkmale des Familiensystems (Anregungsklima, Interesse)	Selbstbestimmungstheoretisch relevante Bedingungen im schulischen und außerschulischen Kontext	Lernbezogene Einstellungen, Erziehungspraktiken und Instruktionsstrategien von Eltern, selbstbestimmungstheoretisch relevante Aspekte des kindperzipierten Lehrerhaltens
Domäne	Mathematik	Naturwissenschaften	Mathematik
Angestrebte Altersgruppe	Heranwachsende (5. bis 9. Klasse), Eltern, Lehrer	Heranwachsende (5–12 Jahre), Eltern, Erzieher, Lehrende	Heranwachsende (3. bis 7. Klasse), Eltern, Lehrende
Design	Kohortengestaffelte Längsschnittstudie	Kohortengestaffelte Längsschnittstudie	Längsschnittstudie
Methoden	Fragebögen, Schulnoten	Interviews, Fragebögen	Fragebögen, Tests, Interviews, Verhaltensbeobachtungen
Übergeordnete Projektziele	Analyse des Zusammenhangs von Motivkomponenten und Leistungen, Identifizierung schulischer und außerschulischer Einflussfaktoren, Entwicklung eines Kompendiums von motivationsbezogenen Fördermaßnahmen	Analyse der Genese von Interessen und Nicht-Interessen von Heranwachsenden in verschiedenen Altersstufen, Identifizierung relevanter Kontextbedingungen	Analyse der Genese von mathematikbezogenen Einstellungen und Motiven, Analyse von Veränderungen in den motivrelevanten Kontextbedingungen beim Übergang in die Sekundarstufe, Entwicklung eines motivbezogenen Elterntrainings und Konzepts der Elternarbeit von Lehrern

Ordnet man die drei Projekte entlang ihrer theoretischen Einordnung, dann steht in dem *Potsdamer Projekt* (vgl. Rheinberg/Wendland in diesem Band) der leistungsmotivations-theoretische Ansatz im Mittelpunkt, der in der deutschsprachigen Motivationsforschung lange Zeit eine dominierende Rolle spielte und in vielfältiger Weise weiterentwickelt wurde. Um die aus dieser theoretischen Perspektive relevanten Motivationskomponenten erfassen zu können, wurde mit dem „*Postdamer Motivations Inventar (PMI)*“ zunächst ein bei Sekundarstufenschülern einsetzbares Erhebungsinstrument entwickelt. Wie Rheinberg und Wendland in ihrem Beitrag ausführen, leisten mehrere PMI-M Faktoren zusammengenommen einen *signifikanten (je nach Klassenstufe zwischen knapp 5 und 21% betragenden) Beitrag zur Aufklärung der Varianz in den Schulleistungen von Schülern*, selbst wenn der Effekt der vorausgegangenen Note kontrolliert wird. Besonders hervorhebenswert ist, dass die verschiedenen Motivationskomponenten zwar unterschiedlichen Verlaufsmustern folgen, sich neben *diversen „Verschlechterungstypen“* aber keine Komponenten finden lassen, die einem Aufwärtstrend folgen. Insofern unterstreichen die Befunde die Notwendigkeit einer – auch im Rahmen des Potsdamer Projekts angestrebten – Entwicklung von (förmlichen) Fördermaßnahmen, die je nach diagnostiziertem Problemprofil einsetzbar sind.

Welche besonderen (weil emotional getönten und wertbesetzten) *Person-Gegenstands-Bezüge sowie schulbezogenen Einstellungen sich bei Kindern im Kindergarten sowie in der Grundschule und Sekundarstufe I* finden lassen, steht im Zentrum des *Münsteraner Projekts* (Upmeier zu Belzen u.a., in diesem Band). Um kurz- und langfristige Entwicklungen in den Interessen und Nicht-Interessen Heranwachsender beschreiben und erklären zu können, wurden halbstandardisierte Interviews mit Kindern und Eltern sowie Erzieherinnen und Lehrpersonen durchgeführt, die mit dem Textanalyseprogramm WinMAX ausgewertet wurden. Erste Ergebnisse, die auf den Daten von knapp 100 Fällen basieren, liefern vor allem eine *differenzierte Deskription von Interessenverläufen*. So steigt mit zunehmenden Alter nicht nur der Anteil der Kinder mit definierten Interessen. Vielmehr lassen sich auch *qualitative Verschiebungen* dahingehend nachzeichnen, dass immer mehr Kinder über qualitativ höherwertige und stabile Interessen berichten und mit dem Eintritt ins Schulleben vermehrt Abneigungen herausgebildet werden. Nach *subjektiv als bedeutsam für ihre Interessenentwicklung erlebten Bedingungen* gefragt berichten Kinder kaum von Lehrpersonen, die ihr Interesse für Biologie geweckt hätten, wohl aber von Lehrern, die zur Ausbildung von abneigenden Haltungen beigetragen hätten. Im Kontrast dazu wird der Einfluss der Gleichaltrigen und vor allem der Eltern durchgängig als sehr positiv beschrieben.

Diese Befunde konvergieren teilweise mit den Ergebnissen des *Bielefelder Matheprojekts*, in dem (unter anderem) Grundschüler gebeten wurden, die Ausprägung vorgegebener (theoretisch als motivrelevant erachteter) Merkmale der Eltern-Kind- und der Lehrer-Schüler-Beziehung einzuschätzen. Auch hier zeichnet sich die Bedeutung der *Familie für die Herausbildung von naturwissenschaftlichen Einstellungen, Motiven und Zielen* ab. Die ebenfalls aus Schülersicht erfassten Verhaltensweisen von Lehrern stehen aber in ähnlich engem Zusammenhang zu den mathematikbezogenen Einstellungen, Motiven und Zielorientierungen von Schülern. Aus theoretischer Sicht ist interessant,

dass sich sowohl auf der Basis der inhaltsanalytischen Auswertung von Interviewdaten (Münster) als auch auf Grundlage quantitativer Analyse von Fragebogendaten (Bielefeld) schulische und familiäre Kontextbedingungen als motivrelevant identifizieren lassen, die *in Einklang mit zentralen Aussagen der Interessen- und Selbstbestimmungstheorie* stehen. Die längsschnittlichen Analysen, die im Bielefelder Projekt anstehen, werden zeigen, wie es um die *Verbreitung dieser motivfördernden und -hemmenden Erziehungs- und Instruktionsstrategien von Eltern und deren Veränderung im Laufe der Schulzeit* steht.

Mit Blick auf das übergeordnete Erkenntnisinteresse des BIQUA-Schwerpunkts bleibt festzuhalten, dass die drei im nachfolgenden Abschnitt skizzierten Projekte – nicht zuletzt aufgrund ihrer Vernetzung untereinander und mit Projekten aus anderen Schwerpunktbereichen – in den nächsten Jahren umfängliche und zugleich äußerst differenzierte Erkenntnisse über die Genese verschiedener Formen der Lernmotivation und deren Beeinflussung durch schulische und außerschulische Kontextbedingungen liefern dürften.

Literatur

- Baumert, J./Bos, W./Watermann, R. (1998): TIMSS/III – Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J./Bos, W./Watermann, R. (2000): TIMSS/III – Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schulbahn. Band 1 + 2. Opladen: Leske + Budrich.
- Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000. Opladen: Leske + Budrich.
- Eccles, J.S./Midgley, C./Wigfield, A./Buchanan, C.M./Reumann, D./Flanagan, C./MacIver, D. (1993): Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and families. In: *American Psychologist*, 48, Heft 2, S. 90–101.
- Fend, H. (1997): Der Umgang mit Schule in der Adoleszenz – Aufbau und Verlust von Lernmotivation, Selbstachtung und Empathie (Bd. 4). Bern: Huber.
- Krapp, A. (1999): Interest, motivation and learning: An educational-psychological perspective. In: *European Journal of Psychology of Education*, 4, Heft 1, S. 23–40.
- Pekrun, R. (1993): Entwicklung von schulischer Aufgabenmotivation in der Sekundarstufe: Ein erwartungswert-theoretischer Ansatz. In: *Zeitschrift für Pädagogische Psychologie*, 7, Heft 2/3, S. 87–97.
- Pintrich, P.R./Schunk, D.H. (1996): Motivation in education. Theory, research, and applications. Englewood Cliffs, NJ.
- Rheinberg, F. (2002): Motivation. Stuttgart: Kohlhammer.
- Schiefele, U. (1996): Motivation und Lernen mit Texten. Göttingen: Hogrefe.
- Wild, E./Hofer, M./Pekrun, R. (2001): Psychologie des Lerner. In: Krapp, A./Weidenmann, B. (Hrsg.): *Pädagogische Psychologie*. Weinheim: Beltz, S. 207–270.
- Wild, E./Hofer, M. (2000): Elterliche Erziehung und die Veränderung motivationaler Orientierungen in der gymnasialen Oberstufe und der Berufsschule. In: Schiefele, U./Wild, K.-P. (Hrsg.): *Interesse und Lernmotivation*. Münster: Waxmann, S. 31–52.
- Wild, E./Remy, K. (2002): Affektive und motivationale Folgen elterlicher Einstellungs- und Verhaltensweisen bei den Mathematik-Hausaufgaben. In: *Unterrichtswissenschaft*, 30, Heft 1, S. 27–51.

Anschrift der Autorin:

Prof. Dr. Elke Wild, Universität Bielefeld, Abteilung Psychologie, AE 09 Pädagogische Psychologie, Postfach 100131, 33501 Bielefeld.

Elke Wild/Katharina Remy

Quantität und Qualität der elterlichen Hausaufgabenbetreuung von Drittklässlern in Mathematik¹

1. Einleitung

Wurden bereits in der TIMS-Studie (vgl. Baumert/Lehmann u.a. 1997; Baumert/Bos/Lehmann 2000) die Defizite deutscher Schülerinnen und Schüler im Umgang mit komplexen mathematischen Problemen deutlich, konkretisieren die jüngst veröffentlichten Ergebnisse der PISA-Studie (vgl. Deutsches PISA-Konsortium 2001) die spezifischen Schwächen des deutschen Bildungssystems. Diese können in vier Punkten zusammengefasst werden:

- Deutsche Schülerinnen und Schüler bekommen insbesondere dann Schwierigkeiten, wenn es um das Reflektieren und Bewerten von Texten geht, wenn also eine selbstständige Auseinandersetzung mit Lerninhalten gefragt ist.
- Der in Deutschland extrem geringe Anteil der Schülerinnen und Schüler, der zum Vergnügen liest, weist zusammen mit Befunden aus früheren Studien darauf hin, dass die selbstständige Auseinandersetzung mit Lerninhalten nicht zuletzt an der fehlenden Lernmotivation scheitert.
- Gerade die leistungsschwächeren Schülerinnen und Schüler werden im Rahmen des deutschen Bildungssystems, welches eine frühzeitige Differenzierung vorsieht, nur unzureichend gefördert. Dies unterstreicht die besondere Bedeutung adaptiver Instruktion und belegt die unzulängliche Realisation im Unterricht.
- Schichtspezifische Ungleichheiten werden hierzulande weit weniger kompensiert als in anderen Nationen. Dies kann nicht zuletzt als Hinweis auf das Fehlen von Maßnahmen interpretiert werden, die sowohl über eine kompensatorische Förderung als auch durch Einbeziehung der Familien auf eine Verbesserung der Bildungschancen sozial benachteiligter Schülerinnen und Schüler abzielen.

Das vorliegende Projekt setzt an diesen vier Problembereichen insofern an, als angenommen wird, dass eine Verbesserung der Bildungsqualität von Schule über eine gezielte Förderung der *Bereitschaft zum selbstgesteuerten Lernen* zu erreichen ist. Weiterhin wird postuliert, dass eine solche Förderung nicht nur an der Ausgestaltung der Lehrer-Schüler-Beziehung ansetzen darf, sondern auch die Qualität des häuslichen Lernens sowie das Zusammenspiel von Elternhaus und Schule in den Blick nehmen muss.

Mit dieser thematischen Schwerpunktsetzung nimmt das Projekt, bezogen auf das in Abb. 4 (Doll/Prenzel in diesem Band) skizzierte Rahmenmodell vier Systeme in den

1 Die Studie wurde gefördert durch Mittel der DFG (WI 1607/1-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

Blick. Neben dem System Eltern/Familie und den Unterrichtsprozessen (hier insbesondere die Qualität der Lehrer-Schüler-Beziehung) sind dies die individuellen Voraussetzungen und Verarbeitungsprozesse, die in ihrem Einfluss auf den Lernzuwachs, die generelle Einstellung zu Schule und Lernen sowie die Fähigkeit zum selbstregulierten Lernen untersucht werden.

2. Theoretischer Hintergrund

Aus dem Spektrum der Formen der Motivation, die dem Lernverhalten von Schülerinnen und Schülern zugrunde liegen können (zusf. Wild/Hofer/Pekrun 2001; Pekrun u.a., in diesem Band), werden im Bielefelder Mathematikprojekt solche ins Zentrum gerückt, die mit dem Gefühl der Selbstbestimmung einhergehen und als motivationale Basis für die selbstregulierte Auseinandersetzung mit mathematischen Problemen gelten können. Hierzu zählt zunächst das (personale) *Mathematikinteresse*, welches im Sinne der Person-Gegenstands-Theorie des Interesses (Krapp 1998, 1999) als spezifische überdauernde Person-Gegenstands-Beziehung definiert ist, die insbesondere durch gefühlsbezogene Valenzen gekennzeichnet ist – die Beschäftigung mit dem Gegenstand macht Spaß. Für weitere Formen selbstbestimmter Motivation ist dagegen kennzeichnend, dass sie mit dem beim Lernen erzielten Kompetenzzuwachs zusammenhängen – Bildung wird als wertvoll erachtet, Lernfortschritte gehen mit einer positiveren Selbstbewertung einher. Die Selbstbestimmungstheorie (zusf. Ryan/Deci 2001) spricht in diesem Zusammenhang von einer „*identifizierten Motivation*“, die insofern als eine extrinsische Form der Motivation verstanden werden kann, als dass sie sich aus der Valenz des Ergebnisses der Lernhandlung ableitet. Aufgrund der freiwilligen Selbstbindung der Person an den Wert von Bildung und des mit dem Lernen einhergehenden Gefühls von Autonomie kann sie aber zugleich als selbstbestimmte Form der Motivation gelten. Dies trifft auch auf die *Aufgaben- oder Lernzielorientierung* von Schülern zu, die meist als Gegenpol zur Ich- oder Leistungsorientierung gefasst wird und dadurch definiert ist, dass bei Lernern mit ausgeprägter Lernzielorientierung der Lernprozess im Vordergrund steht und dass sich nur bei der Bearbeitung von herausfordernden Aufgaben Zufriedenheit einstellt (zusf. Wild/Hofer/Pekrun 2001).

In dem Bemühen, die (aktual- und ontogenetischen) Bedingungen (der Veränderung) des Interesses und anderer Formen selbstbestimmter Lernmotivation aufzuklären, wird im Bielefelder Mathematikprojekt auf die bereits erwähnte Selbstbestimmungstheorie und die Interessentheorie zurückgegriffen. Schülerinnen und Schüler lernen demnach nur dann aus Spaß oder innerer Überzeugung und sind mit dem erzielten Lernfortschritt erst dann zufrieden, wenn *schulische und außerschulische Lehr-Lernarrangements* so gestaltet sind, dass Schüler ihre psychologischen Bedürfnisse befriedigen können (vgl. auch den stage-environment-fit-Ansatz von Eccles u.a. 1993; Roesner/Eccles 1998). Legt man die drei in der Selbstbestimmungstheorie genannten Grundbedürfnisse nach Autonomie- und Kompetenzerleben und nach sozialer Einbindung zugrunde, dann gilt es, Schülerinnen und Schülern in schulischen wie außerschulischen

lischen Kontexten die Möglichkeit zu geben, sich als autonom, kompetent und wertgeschätzt zu erleben.

Da in den letzten drei Jahrzehnten eine Vielzahl empirischer Studien zur Prüfung selbstbestimmungstheoretischer Annahmen durchgeführt wurden, die sich auf den schulischen Kontext konzentrierten – dieser Bereich wird durch das Kooperationsprojekt Essen-Bielefeld weiter beforscht (Sumfleth u.a., in diesem Band) – richteten sich unsere Bemühungen im Projekt zunächst auf die Klärung von konzeptuellen Fragen zum elterlichen Schulengagement und die Identifizierung motivrelevanter Erziehungspraktiken und Instruktionsstrategien von Eltern beim häuslichen Lernen. Aufbauend auf familienpsychologischen Arbeiten und Studien zur Selbstbestimmungstheorie wurde vorgeschlagen, zwischen bereichsübergreifenden und -spezifischen Erziehungspraktiken von Eltern zu unterscheiden und auf der bereichsspezifischen (lernbezogenen) Ebene von vier Dimensionen elterlichen Verhaltens in häuslichen Lehr-Lernarrangements auszugehen (vgl. Wild 1999; 2001a; Wild/Remy 2001). Neben der *autonomieunterstützenden Instruktion* von Eltern (die sich in den Schüler einbeziehenden und aktivierenden Impulsen zeigt) und der *direktiv-kontrollierenden Instruktion* (die in einem stark lenkenden und am Lernergebnis orientierten Lehrverhalten zum Ausdruck kommt) wird der Aspekt der *Struktur* (im Sinne des konsistenten Vertretens lernbezogener Werte und Standards) und der *elterlichen Responsivität* genannt (der sich in Bemühungen der Eltern spiegelt, ihr Kind emotional im Umgang mit Lernschwierigkeiten und herausfordernden Aufgaben zu begleiten).

Da andernorts (vgl. Wild/Remy 2001; Wild 2001b) bereits dargelegt wurde, dass die skizzierten Dimensionen des Elternverhaltens ebenso wie korrespondierende Dimensionen des Verhaltens von Lehrern zur Erklärung von interindividuellen Unterschieden in der Lernmotivation von Schülern herangezogen werden können, und da derzeit noch keine längsschnittlichen Analysen durchgeführt werden können (s. nachfolgender Abschnitt), sollen in diesem Beitrag deskriptive Befunde zur vorherrschenden Hausaufgabenpraxis vorgestellt werden, die Aufschluss über den Bedarf an verschiedenen Formen der Elternberatung geben.

Da Fragen des Zusammenspiels von Familie und Schule sowohl in der Pädagogik als auch in der Unterrichtsforschung weitgehend vernachlässigt wurden (zusf. Wild 2001b), lässt sich die Qualität oder Problematik familialer Lehr-Lernprozesse hierzulande kaum abschätzen. Recht gut belegt ist lediglich, dass Lehrer, Eltern und Schüler den Hausaufgaben einen hohen Wert beimessen (Nilshon 1998) und dass das außerschulische Lernen in den letzten Jahren und Jahrzehnten – nicht zuletzt infolge der gestiegenen Bildungsaspirationen von Eltern und den Schülern selbst – quantitativ stark angestiegen ist. Nachhilfe gehört somit zum Alltag vieler Schüler und deren Eltern, die in etwa der Hälfte der Fälle mit ihrem Kind selbst üben. Das außerschulische Lernen stellt vielmehr auch einen bedeutenden Wirtschaftsfaktor dar – Kramer und Werner (1998) gehen von knapp 4,5 Milliarden DM aus, die jährlich in privaten Nachhilfeunterricht investiert werden, wobei kommerzielle Nachhilfeinstitutionen keinen Qualitätskontrollen unterliegen und anzunehmen ist, dass Nachhilfe oft von unzureichend qualifizierten Kräften durchgeführt wird.

Über den *Nutzen* speziell der Hausaufgaben ist in den letzten Jahren lange und heftig debattiert worden (zusf. Nilshon 1998). Die Vermutung, dass nicht selten gut gemeinte aber didaktisch ungünstige Bemühungen der Eltern dazu beitragen, dass die mit dem außerschulischen Lernen verknüpften didaktischen und erzieherischen Ziele vereitelt oder sogar unerwünschte Arbeitshaltungen hervorgerufen werden, stützt sich zum einen auf Experimente zur Überprüfung der Wirksamkeit von Hausaufgaben (zusf. Hacher/Bischof 2000). Diese sind allerdings größtenteils veraltet und erfassen die elterlichen Bemühungen nicht oder nur unter quantitativem Aspekt. Zum anderen können Interaktionsanalysen häuslichen Lernens (z.B. Krohne/Hock 1994) und Tagebuchstudien (Wild 1999) herangezogen werden, die allerdings meist auf kleinen Stichproben basieren. Schließlich sind noch sozialisationstheoretische Studien (z.B. Fend 1998) zu nennen, die zeigen, dass Schulprobleme häufig Anlass für Konflikte und Streitgespräche zwischen Eltern und Kindern sind, jedoch wenig Aufschluss über (kontra-)produktive Formen der Gestaltung häuslicher Lehr-Lernarrangements geben.

3. Design der Studie

Die Bielefelder Mathematikstudie sieht vor, die Interessen- und Leistungsentwicklung von Schülern über einen Zeitraum von sechs Jahren zu verfolgen. Da sie in der Grundschule startet, handelt es sich um eine prospektiv angelegte quasi-experimentelle Untersuchung, in deren Rahmen die Bewältigung des *Übergangs von der Grundschule in eine weiterführende Schule der Sekundarstufe I* sowie die weitere Leistungs- und Motiventwicklung von Schülern in Abhängigkeit von ihren individuellen Eingangsvoraussetzungen sowie antezedenten und konkurrenten Merkmalen des Schul- und Familienkontextes untersucht werden kann.

Die hier vorgestellten Ergebnisse basieren auf den Angaben einer Zufallsstichprobe von 304 Familien, die erstmalig im Frühjahr 2001 zuhause besucht wurden und weiterhin im Jahresabstand befragt und getestet werden sollen. Vor Beginn der Befragung wurden sämtliche Versuchsleiterinnen geschult, um sicherzustellen, dass die Durchführung des Mathematik-Leistungstests ordnungsgemäß abläuft und den Familien keine Informationen gegeben werden, die die weiteren Ergebnisse verzerren könnten.

Um einerseits eine möglichst große und heterogen zusammengesetzte Gruppe von Familien berücksichtigen zu können und andererseits die aus lern- und motivationspsychologischer Sicht relevanten Prozesse möglichst kleinschrittig nachzeichnen zu können, wurde aus der Eingangsstichprobe eine nach den Leistungen der Schüler im Mathematiktest selektierte Teilstichprobe von 40 Familien ausgewählt. Diese wird intensiv im Halbjahresabstand besucht, sodass neben Selbst- und Fremdauskünften auch Verhaltensbeobachtungen und Tagebuchprotokolle zur Prüfung von Hypothesen herangezogen werden können (vgl. Abbildung 1, S. 280).

G1 Frühjahr 01	I1 Herbst 01	G2 Frühjahr 02	I2 Herbst 02	G3 Frühjahr 03	I3 Herbst 03	G4 Frühjahr 04
		N=80 (-) N=60 (0) N=60 (+) Viertkl. + Eltern		N ≈ 200 Fünftkl. + Eltern		N ≈ 200 Sechstkl. + Eltern
GE- SAMT: N=304 Drittkl. + Eltern	N= 20 (-) N= 10 (0) N= 10 (+) Viertkl. + Eltern	N ≈ 40 Viertkl. + Eltern	N ≈ 40 Fünftkl. + Eltern + Lehrer	N ≈ 40 Fünftkl. + Eltern	N ≈ 40 Sechstkl. + Eltern + Lehrer	N ≈ 40 Sechstkl. + Eltern
Über- gangs- stich- probe		N max. 60 Befragung zu anstehendem Übergang		N ≈ 60 Befragung zu erfolgtem Übergang		N ≈ 60 Befragung zu erfolgtem Übergang
Ergänz. Sampling		max. N=20 Schüler mit Beratungs- bedarf		N ≈ 20		N ≈ 20
Pilot- Stich- probe		Ca. 120 Lehrer zur E-L-Koop. Ca. 40 Familien zur Erprobung		Ca. 10 Lehrer zur Erprobung der E-L-Koop.	max. N=50 Eltern zur Evaluation des B.konzepts	

Legende: grau unterlegt: Intensivstichprobe (Tests, Videos, Tagebücher, Fragebögen),
(+) leistungsschwache, (0) durchschnittliche, (-) leistungsstarke Schüler

Abb.1: Das Design der Bielefelder Mathematikstudie

Zur Gewinnung weiterer Informationen sollen zum einen die Mathematiklehrer der Zielkinder befragt werden, die diese in der Sekundarstufe I unterrichten. Davon unabhängig richtet sich eine Befragung von Lehrkräften an unterschiedlichen Schulformen darauf, ihre derzeitige Praxis der Elternarbeit und Rahmenbedingungen in Erfahrung zu bringen, die die Implementation einer reformierten Form der Eltern-Lehrer-Kooperation behindern könnten.

Ebenfalls begonnen wurde mit der Erprobung eines Gruppentrainings für Eltern. Das Training richtet sich an Familien, in denen es aufgrund von (Teil-)Leistungsschwächen der Schüler zu häufigen Konflikten kommt, die die Eltern-Kind-Beziehung belasten. Das Training soll Häufigkeit und Schärfe dieser Konflikte vermindern und Eltern eine Vorstellung davon vermitteln, wie häusliche Lehr-Lernarrangements in motiv- und lernförderlicher Weise gestaltet werden können. Geplant ist eine in mehreren

Schritten ablaufende, formative und summative Evaluation des Trainings, wobei der Effekt der Intervention durch einen Vergleich der Entwicklungsverläufe der Experimentalgruppen jeweils mit einer Wartekontrollgruppe geprüft werden soll. Langfristig ist geplant, dieses Training mit dem von Gürtler u.a. (in diesem Band) entwickelten Selbstregulationstraining zu verknüpfen.

4. Stichprobe

Das Durchschnittsalter der zum ersten Messzeitpunkt teilnehmenden Kinder (53,8% Mädchen, 46,2% Jungen) beträgt 8,9 Jahre, das der Mütter 39,3 Jahre und das der Väter 42,1 Jahre.

Eine Analyse der soziodemographischen Merkmale der Familien unter Heranziehung der Statistiken des Statistischen Jahrbuchs (2000) ergibt eine leichte Unterrepräsentation von Einzelkindern (15%) sowie eine deutliche Überrepräsentation von strukturell intakten Familien (81%) und Familien der mittleren (34%) und oberen Mittelschicht bzw. Oberschicht (63%). Diese für familienpsychologische Studien typische Mittelschichtslastigkeit gilt es bei den nachfolgenden Befunden ebenso zu berücksichtigen wie die Tatsache, dass 93% der Väter in Vollzeit beschäftigt sind, während 26% der befragten Mütter keiner Erwerbstätigkeit und 60% einer Teilzeitbeschäftigung nachgehen.

Das Notenspektrum der befragten Grundschüler im Fach Mathematik reicht von „sehr gut“ bis „mangelhaft“, wobei die meisten Schüler aber zwischen „gut“ und „befriedigend“ rangieren.

5. Instrumente

Da es bislang an Informationen über die Gestaltung häuslicher Lehr-Lernarrangements in einzelnen Fächern mangelt, wurde den Schülern und Eltern ein Fragebogen vorgelegt, der Einzelitems zu folgenden Fragen enthielt: (1) in welchem Ausmaß die Eltern bei der Bearbeitung der Hausaufgaben involviert sind, (2) ob weitere Personen bei den Hausaufgaben Hilfestellung geben und (3) ob und aus welchem Grund von wem Nachhilfe gegeben wurde. Darüber hinaus wurden den Schülern vier Skalen vorgelegt, um zu erheben (4) inwiefern das Verhalten der Eltern beim häuslichen Lernen als autonomieunterstützend, strukturierend, wertschätzend und nicht direktiv-kontrollierend erlebt wird. Für den Elternfragebogen wurden – soweit sinnvoll – analoge Items formuliert. Tabelle 1 (S. 282) gibt einen Überblick über die Instrumente, nähere Angaben finden sich in Wild u.a. (2001).

Tab. 1: Übersicht über die eingesetzten Instrumente

Konstrukt	Schülerversion		Elternversion	
	Anzahl der Items	Reliabilität	Anzahl der Items	Reliabilität
Quantität der Hilfe bei den Hausaufgaben				
Elternbeteiligung bei den HA	9 Einzelitems	–	9 Einzelitems	–
andere Helfer	6 Einzelitems	–	6 Einzelitems	–
Nachhilfe	8 Einzelitems	–	8 Einzelitems	–
Gründe für Nachhilfe		–	5 Einzelitems	–
Qualität elterlicher Hilfe				
Direktiv-kontrollierende Instruktion	6	$\alpha = .75$	6	$\alpha = .79$
autonomie-unterst. Instruktion	5	$\alpha = .62$	4	$\alpha = .61$
Emotionale Unterstützung	8	$\alpha = .78$	6	$\alpha = .71$
Struktur	4	$\alpha = .58$	4	$\alpha = .67$

6. Ergebnisse

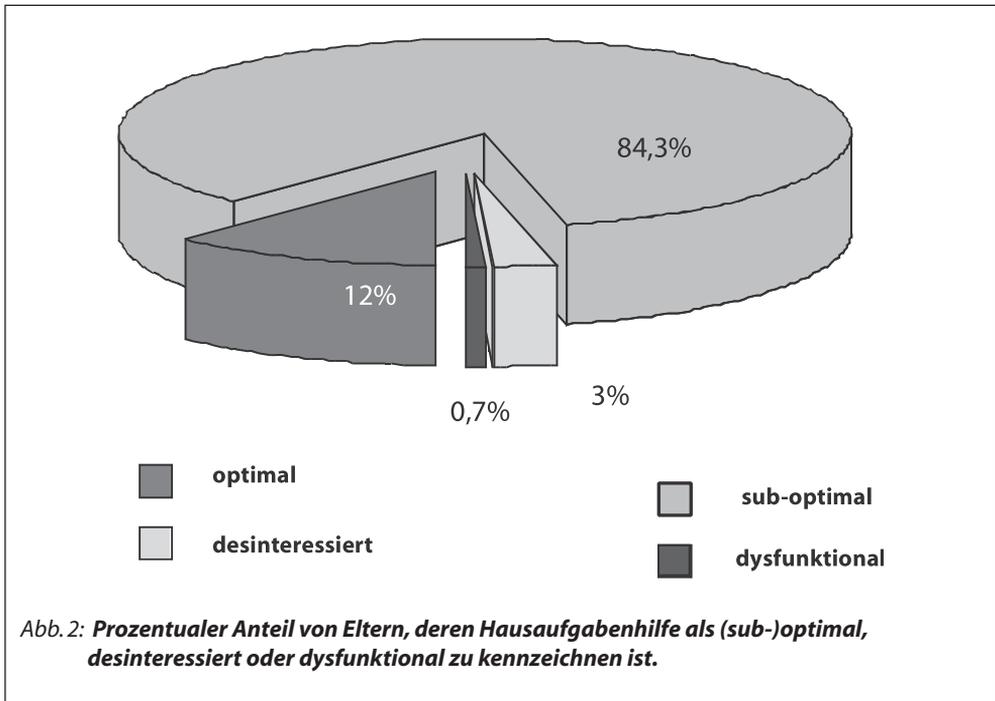
Wie bereits ausgeführt, liegen keine differenzierteren Kenntnisse über Umfang und Anlässe *verschiedener Formen häuslichen Lernens* (darunter Hausaufgaben, Übungen vor Klassenarbeiten und Nachhilfe bei Leistungsproblemen), über die Hilfe leistenden Personen und über Unterschiede in der *Qualität der elterlichen Unterstützung* vor.

Die vorliegende Studie vermag diese Informationen für Familien mit Drittklässlern im Fach Mathematik zu liefern, dem als Hauptfach eine besondere Bedeutung für die Schulkarriere von Kindern zukommt.

6.1 Wer übernimmt die Verantwortung für das häusliche Lernen?

Die Frage, in welchem Umfang Drittklässler Hilfe beim häuslichen Lernen benötigen, um eine positive Interessen- und Leistungsentwicklung zu durchlaufen, kann in Ermangelung aussagekräftiger Studien derzeit nicht beantwortet werden. Wird allerdings in Rechnung gestellt, dass in Schulordnungen die Funktion der Hausaufgaben primär in der Förderung selbstregulierten Lernens gesehen wird und schon Zweitklässler mehrheitlich der Auffassung sind, die Erledigung der Hausaufgaben falle in ihren eigenen Verantwortungsbereich (Warton 1997), dann dürfte weder eine kleinschrittige Anleitung noch eine Ablehnung jeglicher Form der Unterstützung optimal sein.

Betrachtet man vor diesem Hintergrund die Schülerantworten auf die Frage „Wenn ich zuhause lerne, muss ich das *immer alleine* machen“ und „Ich lerne *immer gemeinsam* mit meinen Eltern“, dann erachten 9,2% die erste und 21,4% die zweite Aussage für „genau zutreffend“. Weisen diese Zahlen – in Einklang mit den Befunden von Tietze/Roßbach/Mader (1987) – darauf hin, dass weniger als 10% der Grundschüler ihre Hausaufgaben ohne fremde Hilfe erledigen, aber etwa ein Fünftel ein Zuviel an Anleitung bekommt? Ein näherer Blick in die Daten relativiert diese Aussage. Sie zeigen, dass von den Schülern, die nach eigener Angabe immer alleine lernen, immerhin 51,9% auch die Aussage „Meine Eltern helfen mir beim Lernen, wenn ich sie darum bitte“ für „sehr zutreffend“ halten. Umgekehrt findet sich unter denen, die immer gemeinsam mit den Eltern lernen, immerhin ein Anteil von 52,4%, die gleichzeitig die Aussage „Meine Eltern helfen mir beim Lernen, indem sie mich erst allein lernen lassen und danach die Aufgaben mit mir durchsprechen“ für sehr zutreffend halten. Berücksichtigt man zur Einschätzung des Umfangs der elterlichen Hilfe deshalb gleichzeitig alle Schülerangaben, dann ergibt sich die in Abbildung 2 festgehaltene Verteilung.



Positiv hervorzuheben ist, dass nach Schülerangaben nur ein sehr kleiner Anteil von befragten Eltern entweder ein völliges Desinteresse am außerschulischen Lernen ihres Kindes zeigt oder umgekehrt dem Kind die Verantwortung für die Hausaufgaben vollständig abnimmt und jeden Schritt gemeinsam bearbeitet und kontrolliert. Dennoch kann nur in 12% der Fälle von einer optimalen Hausaufgabenbetreuung gesprochen

werden. Von einer optimalen Hausaufgabenbetreuung sprechen wir, wenn die Eltern sich zwar als Ansprechpartner im Hintergrund halten und ihr Kind auf Bitten oder bei erkennbaren Schwierigkeiten auch anleiten, dieses aber zum selbstständigen Lernen hinführen, indem sie ihm zunächst die Verantwortung für das Vorgehen bei der Aufgabenbearbeitung überlassen.

Es gibt also bereits in der Grundschule Möglichkeiten der Verbesserung, wobei unter dem Gesichtspunkt der sozialen Chancengleichheit interessant ist, dass sich weder in Abhängigkeit von der mütterlichen Berufstätigkeit (also im Vergleich voll berufstätiger und teilzeitbeschäftigter Frauen sowie Hausfrauen), noch von der sozialen Herkunft der Familien (die über die Schul- und Berufsausbildung beider Elternteile operationalisiert wurde) oder der Familienstruktur (strukturell intakt oder nicht) Unterschiede in der Wahrnehmung der elterlichen Hilfe beim Lernen zeigen. Lediglich bei der kleinen (N = 32) Gruppe der Eltern, die nicht in Deutschland aufgewachsen sind, finden sich häufiger Anzeichen für stark kontrollierende oder desinteressierte Formen der Hilfe.

6.2 *Wer ist Ansprechpartner beim häuslichen Lernen?*

Erwartungsgemäß gibt mit 81% die Mehrzahl der Drittklässler an, dass Hilfe bei den Hausaufgaben oder beim Üben von den Eltern geleistet wird, wobei in aller Regel die Mutter hierfür zuständig ist. Auffällig hoch ist mit 60% allerdings auch der Anteil der Kinder, die angeben, Hilfe von Geschwistern zu bekommen. Da weitere 32% der Kinder angeben, dass Mitschüler ihnen beim Lernen helfen, scheinen Gleichaltrige eine nicht unerhebliche Rolle beim außerschulischen Lernen zu spielen. Allerdings deutet die Tatsache, dass dieselben Fragen nach der Unterstützung durch Geschwister und Mitschüler nur von 32% bzw. 2% der Eltern bejaht werden, darauf hin, dass Eltern die Hilfe der Gleichaltrigen stark unterschätzen.

6.3 *Wie gestaltet sich die Nachhilfe?*

Die von Eltern veranlasste und finanzierte *Nachhilfe* stellt insofern einen Sonderfall häuslichen Lernens dar, als sie außerhalb der Familie stattfinden *kann*. Faktisch zeigt sich allerdings, dass von den knapp fünf Prozent der Kinder, die im vergangenen Schuljahr *Nachhilfeunterricht* erhielten, lediglich 29% bezahlten Nachhilfeunterricht bekamen. Da in den von Kramer und Werner (1998) befragten Familien über verschiedene Fächer und Klassenstufen hinweg etwa die Hälfte der Nachhilfe von den Eltern selbst geleistet wurde, deuten unsere Ergebnisse auf eine vergleichsweise hohe Selbstbeteiligung von Eltern hin, die mit der Bedeutung des Fachs und/oder der niedrigen Klassenstufe zu erklären ist. Interessant ist dabei, dass unsere Zahlen weitgehend denen einer groß angelegten Studie (Schmidt 1984) entsprechen, in der die vor 17 Jahren gängige Hausaufgabenpraxis in der Grundschule untersucht wurde. Die seitdem kontinuierlich gestiegenen Bildungsaspirationen der Eltern scheinen insofern nicht mit einer drama-

tisch anwachsenden Nachfrage nach Nachhilfeunterricht einherzugehen. Allerdings zeigt sich anders als noch bei Schmidt (1984), dass 45% der Eltern, deren Kind Nachhilfe bekommt, trotz der frühen Leistungsprobleme an dem Ziel, ihr Kind möge einen Hochschulabschluss erwerben, festhalten. Insgesamt wünschen sich 66,5% der Eltern, ihr Kind möge die allgemeinbildende Schule mit dem Abitur abschließen, und von diesen Eltern halten wiederum 67,9% diesen Abschluss für realistisch. Somit ergibt sich ein Gesamtanteil von 45,1% der Eltern, die ernsthaft die Hochschulreife für ihr Kind anstreben und ihr Engagement möglicherweise schon in Anbetracht der herannahenden Übertrittsempfehlung, spätestens wohl aber in Reaktion auf schulische Misserfolge ihres Kindes steigern werden.

6.4 *Gestalten Eltern das häusliche Lernen in einer Weise, die unter dem Gesichtspunkt der Förderung selbstbestimmter Formen der Lernmotivation als günstig zu erachten ist?*

Wie bereits ausgeführt, sind aus der Perspektive der Selbstbestimmungstheorie jene Lern- und Entwicklungsumgebungen als förderlich zu bezeichnen, die den psychologischen Grundbedürfnissen von Heranwachsenden Rechnung tragen. Angesichts des Mangels einschlägiger Studien lässt sich derzeit aber nicht abschätzen, in wie vielen Familie eine Form der elterlichen Unterstützung gewährt wird, die dem Schüler durchaus Strukturierungshilfen gibt, auf direktiv-kontrollierende Vorgehensweisen aber verzichtet und stattdessen selbstbestimmtes und selbstreguliertes Lernen durch „dosierte Hilfe“ bei der Regulation kognitiver und emotionaler Prozesse anregt.

Werden hierzu die Schülereinschätzungen herangezogen, dann ergibt sich ein insgesamt erfreuliches Bild: Die meisten Schüler glauben, die Werte und Standards ihrer Eltern zu kennen und sich innerhalb dieser vorgegebenen Struktur selbstbestimmt mit den Mathematikaufgaben auseinandersetzen zu können. Die Mehrzahl fühlt sich zudem emotional unterstützt und nicht in einem extremen Maße beim Lernen kontrolliert oder „überfürsorglich“ betreut.

Einschränkend ist jedoch festzuhalten, dass diese Ergebnisse möglicherweise repräsentativ für Drittklässler sind, vermutlich aber nicht auf andere Alters- bzw. Klassenstufen übertragen werden können. Vielmehr ist mit zunehmendem Leistungsdruck in der Schule und unter Umständen schon mit dem herannahenden Wechsel auf eine weiterführende Schule der Sekundarstufe I mit einer Verstärkung kontrollierender und einer Verringerung emotional- und autonomieunterstützender Reaktionen zu rechnen. Diese Vermutung stützt sich auf den bereits jetzt beobachtbaren Zusammenhang zwischen der Leistungsfähigkeit der Schüler und dem Ausmaß der von Schülern ($r = -.37^*$) und von Eltern ($r = -.23^*$) berichteten Kontrolle.

Ein weiterer auffälliger Aspekt betrifft die Frage der Korrespondenz von Schüler- und Elternurteil. Interessanterweise lässt sich bereits in den Familien mit Drittklässlern eine sonst vor allem in jugendpsychologischen Studien (zufs. Wild/Hofer 2001) berichtete intergenerationale Diskrepanz beobachten. Bezüglich aller Einschätzungen mit

Ausnahme der Struktur finden sich signifikante Unterschiede zwischen den Angaben der Kinder- und Elterngeneration. Und auch wenn der in einer Familie vorherrschende Zusammenhang zwischen Selbst- und Fremdeinschätzung ermittelt wird, finden sich relativ durchgängig keine oder allenfalls schwache Korrelationskoeffizienten (vgl. Tabelle 2). Lediglich in der Wahrnehmung der elterlichen Kontrolle scheinen Eltern und Kinder zu einem ähnlichen Urteil zu gelangen.

Tab. 2: Bivariate Korrelationen zwischen den Schülerangaben und den korrespondierenden Selbsteinschätzungen der Eltern (278 < N < 285)

		Eltern			
		(1)	(2)	(3)	(4)
Schüler	(1) autonomieunterstützende Instruktion	.05	.09	-.06	.04
	(2) emotionale Unterstützung	.09	.16**	.02	.11
	(3) direktiv-kontrollierende Instruktion	-.03	-.06	.35**	.16**
	(4) Struktur	.12*	.01	.20**	.12*

* $p < .05$; ** $p < .01$

Gerade weil sich diese Ergebnisse in das Bild familienpsychologischer Studien einfügen, wonach die Qualität der Eltern-Kind-Beziehung im Urteil der Eltern generell positiver ausfällt als in der Einschätzung der Kinder, geben sie zunächst keinen Anlass zur Besorgnis. Wird jedoch in Rechnung gestellt, dass sich Konflikte zwischen Eltern und ihren Kindern ohnehin häufig an schulischen Leistungsprobleme entzünden, dann können Wahrnehmungsdiskrepanzen gerade bei den Hausaufgaben dazu führen, dass es noch schneller zu Missverständnissen kommt und Konflikte eskalieren.

Schließlich sind die grundsätzlich positiven Selbst- und Fremdurteile zur Qualität elterlicher Hausaufgabenhilfe noch zu relativieren, da es schichtspezifische Unterschiede dahingehend gibt, dass in Familien mit niedrigerem sozioökonomischen Status häufiger direktiv-kontrollierende Strategien eingesetzt werden. Da sich eine produktorientierte, bloß formale Kontrolle der Hausaufgaben eher ungünstig auf die Leistungsentwicklung von Kindern (z.B. Trautwein/Köller 2001; Trudewind 1975) und deren intrinsische Motivation (Wild 1999) auszuwirken scheint, könnte hiermit ein Mechanismus angesprochen sein, der die auch in der PISA-Studie wieder dokumentierte ungünstigere Kompetenzentwicklung sozial benachteiligter Schüler zu erklären vermag.

7. Diskussion

In diesem Beitrag wurden erste deskriptive Befunde zu Umfang und Art der elterlichen Hausaufgabenbetreuung im Fach Mathematik vorgestellt. Im Kern bestätigen und erweitern sie das Bild, das sich aus Schüler- und Elternbefragungen zur Akzeptanz von Hausaufgaben ergibt: Häusliches Lernen und Üben wird von beiden Seiten als sinnvoll

erachtet und in der Mehrzahl der Familien auch als eine Angelegenheit geteilter Verantwortung behandelt. Auch unter motivationspsychologischem Aspekt scheinen die Voraussetzungen für verstärkte Einbeziehung der Eltern, wie sie in anderen Ländern längst praktiziert wird, günstig zu sein; zumindest in der dritten Klasse werden die meisten Schülerinnen und Schüler, wenn es um das Fach Mathematik geht, weder völlig allein gelassen mit der Bearbeitung der Aufgaben noch einer kleinschrittigen Kontrolle unterzogen. Dennoch ist ein Optimierungsbedarf zu konstatieren, da im Regelfall günstige und dysfunktionale Strategien nebeneinander verfolgt werden. Hierin spiegelt sich vermutlich weniger ein an didaktischen Prinzipien orientiertes Vorgehen als ein intuitives Reagieren auf situative Hinweisreize, die im Licht der in der eigenen Schulzeit entwickelten, impliziten und wenig reflektierten Lehr-Lerntheorien interpretiert werden.

Mit Blick auf das Verhältnis von Eltern und Lehrern ist interessant, dass die Antwortmuster der Schüler auf eine eigene (Kinder-)Logik verweisen: Aus Sicht von Drittklässlern bedeutet die Aussage, „Ich mache meine Hausaufgaben immer allein“ nicht etwa, dass sich keiner im Hintergrund bereit hält und selbst bei Bedarf keine Unterstützung durch Eltern gewährt wird. Da sich die Zusammenarbeit zwischen Eltern und Lehrern meist auf kurze Gespräche im Rahmen von Elternabenden beschränkt (zusehens Krumm 1996, 1995; Keck 1994), dürften Lehrer ihre Informationen über Qualität und Umfang der elterlichen Hilfe beim häuslichen Lernen vornehmlich von den Schülern beziehen. Weil einzelne Äußerungen aber offenbar leicht ein verzerrtes Bild zeichnen können, gelangen Lehrer möglicherweise zu einer Einschätzung, die der Realität in den Familien nicht entspricht und die Kommunikation mit den Eltern verhindert oder zumindest erschwert.

Weiter ist hervorzuheben, dass zumindest die Schülerangaben auf die besondere Rolle der Gleichaltrigen bei der Hausaufgabenbetreuung verweisen. Es liegt inzwischen vielfältige empirische Evidenz für die – vor allem von Vertretern des strukturgenetischen Ansatzes von Piaget vorgebrachte – These vor, wonach Geschwister und Peers außerhalb der Familie die kognitive und soziomoralische Entwicklung Heranwachsender entscheidend beeinflussen (z.B. Krappmann 1994; Schmid/Keller 1998). Weitgehend im Dunkeln ist jedoch, über welche Mechanismen sich dieser Einfluss im alltäglichen Leben der Schüler vollzieht. Unklar ist, ob lern- und leistungsrelevante Impulse von Geschwistern, Freunden und Klassenkameraden vor allem in den Situationen zum Tragen kommen, in denen Lerninhalte außerhalb der Schule partnerschaftlich bearbeitet werden. Ebenso ist die Frage, wie sich die Hausaufgabenhilfe durch Peers gestaltet – steht hier etwa das Abschreiben der Aufgaben im Bus oder ein reziprokes Lernen im Vordergrund? – noch nicht geklärt.

Ein weiteres, unter anwendungsorientierten Gesichtspunkten bedeutsames Ergebnis betrifft die Diskrepanzen in der Wahrnehmung der elterlichen Hilfe durch Schüler und Eltern und den Umstand, dass vor allem bei leistungsschwachen und sozial weniger privilegierten Schülern bereits in der dritten Klasse Anzeichen für eine Verstärkung dysfunktionaler (weil demotivierender) Strategien erkennbar sind. Sollte sich im Rahmen des Projekts die These einer (von Klassenstufe zu Klassenstufe immer häufiger beobachtbaren) suboptimalen Praxis häuslichen Lehrens und Lernens erhärten lassen, dann

oder allgemein einer *häuslichen Lernkultur als integralem Bestandteil einer Schulkultur* begründen.

Lehrer sind kompetente und vom Grundgesetz zur Kooperation verpflichtete Ansprechpartner der Eltern. Daher liegt es nahe, sie an der an Eltern gerichteten Vermittlung und Einübung von didaktischen Fertigkeiten zu beteiligen. Auch hier ist jedoch von einem Reformbedarf auszugehen, da empirische Befunde die hierzulande praktizierte Form der Eltern-Lehrer-Kooperation als pädagogisch wenig sinnvoll ausweisen. Aus diesem Grund richten sich die mittel- und langfristig anvisierten Ziele des Bielefelder Mathematikprojekts auch auf die Entwicklung und Evaluation eines modular aufgebauten Elterntrainings, in dem die unter lern- und motivationspsychologischen Gesichtspunkten als „optimal“ identifizierten Strategien von Eltern aufgegriffen und umgesetzt werden. Sobald ein bewährtes Training vorliegt, das Eltern hilft, die Lernfreude ihres Kindes, seine Einsicht in die Relevanz mathematischer Kompetenzen und die Fähigkeit zum selbstreflexiven Umgang mit mathematischen Problemen zu fördern, kann es als Ausgangspunkt für die Entwicklung und Erprobung eines vereinfachten Konzepts der Elternarbeit von Lehrern herangezogen werden.

Literatur

- Baumert, J./Bos, W./Lehmann, R. (Hrsg.) (2000): TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Opladen: Leske & Budrich.
- Baumert, J./Lehmann, R./Lehrke, M. u.a. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Deutsches PISA-Konsortium (Hrsg.) (2000): PISA 2000. Opladen: Leske + Budrich.
- Doll, J./Prenzel, M. (2002): Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. In: Zeitschrift für Pädagogik (in diesem Band).
- Eccles, J.S./Midgley, C./Wigfield, A./Buchanan, C.M./Reumann, D./Flanagan, C./MacIver, D. (1993): Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and families. In: *American Psychologist* 48, Heft 2, S. 90–101.
- Fend, H. (1998): Eltern und Freunde. Soziale Entwicklung in der Adoleszenz (Bd. 5). Bern: Huber.
- Gürtler, T./Perels, F./Schmitz, B./Bruder, R. (2002): Training zur Förderung selbstregulativer Kompetenzen in Kombination mit Problemlösen in Mathematik. In: Zeitschrift für Pädagogik, in diesem Band.
- Hascher, T./Bischof, F. (2000): Integrierte und traditionelle Hausaufgaben in der Primarschule – ein Vergleich bezüglich Leistung, Belastung und Einstellungen zur Schule. In: *Psychologie in Erziehung und Unterricht* 47, S. 252–265.
- Keck, R.W. (1994): Hausaufgaben. In: Keck, R. W./Sandfuchs, U. (Hrsg.): Wörterbuch Schulpädagogik. Bad Heilbrunn: Julius Klinkhardt, S. 147–149.
- Kramer, W./Werner, D. (1998): Familiäre Nachhilfe und bezahlter Nachhilfunterricht. Köln: Deutscher Instituts-Verlag.
- Krapp, A. (1998): Entwicklung und Förderung von Interessen im Unterricht. In: *Psychologie, Erziehung und Unterricht* 44, S. 185–201.
- Krapp, A. (1999): Interest, motivation and learning: An educational-psychological perspective. In: *European Journal of Psychology of Education* 14, H.1, S. 23–40.

- Krappmann, L. (1994): Sozialisation und Entwicklung in der Sozialwelt gleichaltriger Kinder. In: Schneewind, K. A. (Hrsg.): *Psychologie der Erziehung und Sozialisation* (Enzyklopädie der Psychologie, Themenbereich D, Serie I). Göttingen: Hogrefe, S. 495–524.
- Krohne, H.W./Hock, M. (1994): *Elterliche Erziehung und Angstentwicklung des Kindes*. Bern: Huber.
- Krumm, V. (1995): Schulleistung – auch eine Leistung der Eltern? Die heimliche und die offene Zusammenarbeit von Eltern und Lehrern und wie sie verbessert werden kann. In: Specht, W./Thonhauser, J. (Hrsg.): *Schulqualität*. Innsbruck: Studien Verlag, S. 256–290.
- Krumm, V. (1996): Über die Vernachlässigung der Eltern durch Lehrer und Erziehungswissenschaft. Plädoyer für eine veränderte Rolle der Lehrer bei der Erziehung der Kinder. In: Leschinsky, A. (Hrsg.): *Die Institutionalisierung von Lehren und Lernen. Beiträge zu einer Theorie der Schule*. Weinheim/Basel: Beltz, S. 119–137.
- Nilshon, I. (1998): Hausaufgaben. In: Rost, D. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim: Psychologie Verlags Union, S. 173–176.
- vom Hofe, R./Pekrun, R./Kleine, M./Götz, T. (2002): Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Konstruktion des Regensburger Mathematikleistungstests für 5 bis 10 Klassen. In: *Zeitschrift für Pädagogik*, in diesem Band.
- Roesner, R.W./Eccles, J.S. (1998): Adolescents' perceptions of middle school: Relation to longitudinal changes in academic and psychological adjustment. In: *Journal of Research on Adolescence* 8, S. 123–158.
- Ryan, R.M./Deci, E.L. (2000): Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. In: *American Psychologist* 55, S. 68–78.
- Schmid, C./Keller, M. (1998): Der Einfluss von Geschwistern auf die kognitive und sozialeschle Entwicklung während der mittleren Kindheit und frühen Adoleszenz. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 30, S. 101–110.
- Schmidt, H.J. (1984): *Hausaufgaben in der Grundschule*. Lüneburg: Klaus Neubauer.
- Statistisches Bundesamt (2000): *Statistisches Jahrbuch 2000*. Stuttgart: Metzler-Poeschel.
- Sumfleth E./Wild, E./Rumann S./Exeler, J. (2002): Wege zur Förderung der naturwissenschaftlichen Grundbildung im Chemieunterricht: Kooperatives Problemlösen im schulischen und familialen Kontext zum Themenbereich Säure-Base. In: *Zeitschrift für Pädagogik*, in diesem Band.
- Tietze, W./Rossbach, H.-G./Mader, J. (1987): Zur Hausaufgaben-situation bei Grundschulern. In: *Empirische Pädagogik* 1, S. 309–329.
- Trautwein, U./Köller, O. (2001): Homework and the development of performance and interest [Internetseite]. http://www.biju.mpg.de/englisch/projects/trautwein_hausaufgaben.htm [26.10.2001].
- Trudewind, C. (1975): *Häusliche Umwelt und Motiventwicklung*. Göttingen: Hogrefe.
- Trudewind, C./Windel, A. (1991): Elterliche Einflussnahme auf die kindliche Kompetenzentwicklung: Schulleistungseffekte und ihre motivationale Vermittlung. In: Pekrun, R./Fend, H. (Hrsg.): *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung*. Stuttgart: Enke, S. 131–148.
- Warton, P.M. (1997): Learning about responsibility: Lessons from homework. In: *British Journal of Educational Psychology* 67, S. 213–221.
- Wild, E. (1999): *Elterliche Erziehung und schulische Lernmotivation*. Unveröffentlichte Habilitation, Sozialwissenschaftliche Fakultät der Universität Mannheim.
- Wild, E. (2001a): Einleitung zum Themenheft „Wider den geteilten Schüler und die Trennung zwischen Schule als „dem“ Lernort und der Familie als „der“ Lebenswelt Heranwachsender“. In: *Zeitschrift für Pädagogik* 47, S. 455–461.
- Wild, E. (2001b): Zum relativen Einfluss von Elternhaus und Schule auf die Lernmotivation von Schülern. In: *Zeitschrift für Pädagogik* 47, S. 455–460.
- Wild, E./Hofer, M. (2001): Innerfamiliale Beziehungserfahrung und Entwicklung in Zeiten sozialen Wandels. In: Walper, S./Pekrun, R. (Hrsg.): *Familie und Entwicklung: Perspektiven der Familienpsychologie*. Göttingen: Hogrefe, S. 131–154.

- Wild, E./Hofer, M./Pekrun, R. (2001): Psychologie des Lernalers. In Krapp, A./Weidenmann, B. (Hrsg.): Pädagogische Psychologie. Weinheim: PVU, S. 207–270.
- Wild, E./Remy, K. (2002): Affektive und motivationale Folgen der Lernhilfen und lernbezogenen Einstellungen der Eltern. In: Unterrichtswissenschaft 30, Heft 1, S. 27–51.
- Wild, E./Remy, K./Gerber, J./Rammert, M./Webler-Pijahn, K./Jonas M. (2001): Dokumentation der Skalen im Schülerfragebogen des Bielefelder Mathematikprojekts. Unveröffentlichtes Manuskript.

Anschrift der Autorinnen:

Prof. Dr. Elke Wild, Universität Bielefeld, Abteilung Psychologie, AE 09 Pädagogische Psychologie, Postfach 100131, 33501 Bielefeld.

Dipl.-Psych. Katharina Remy, Universität Bielefeld, Abteilung Psychologie, AE 09 Pädagogische Psychologie, Postfach 100131, 33501 Bielefeld.

Annette Upmeier zu Belzen/Helmut Vogt/Barbara Wieder/Franka Christen

Schulische und außerschulische Einflüsse auf die Entwicklungen von naturwissenschaftlichen Interessen bei Grundschulkindern¹

1. Einleitung

Interesse entwickelt sich von der frühen Kindheit an. Es besteht Konsens darüber, dass im Vor- und Grundschulalter bereits von Interesse (Schiefele u.a. 1983, S. 6–19) gesprochen werden kann (Weinert 1998, S. 31; Vogt/Wieder 1999, S. 87, 88; Prenzel/Lankes/Minsel 2000, S. 16), wobei sich die Struktur der Interessen von Kindern und Erwachsenen unterscheidet (Prenzel/Lankes/Minsel 2000, S. 16). Unstrittig ist ebenso, dass Entwicklungen im Spektrum von Interessen bis hin zu Nicht-Interessen (Upmeier zu Belzen/Vogt 2001; S. 19) insbesondere von Eltern, Erzieherinnen, Lehrpersonen und Peers von früher Kindheit an maßgeblich und lang anhaltend beeinflusst werden (Vogt/Wieder 1999, S. 87, 88; Prenzel/Lankes/Minsel 2000, S. 16; Wild/Hofer 2000, S. 36).

Vorwiegend aus dem naturwissenschaftlichen Unterricht in der Sekundarstufe ist ein zunehmender Interessenverfall bekannt (Löwe 1987; Krapp 1998, S. 187). Grundschulbezogene Forschung zu den Entstehungs- und Entwicklungsbedingungen von Interessen und Nicht-Interessen – insbesondere über längere Zeiträume – liegen kaum vor. Außerdem fehlen Erkenntnisse über die Wechselwirkung von schulischen und außerschulischen personalen Einflüssen mit der Interessenentwicklung von Kindern.

Entsprechend dem Einstellungsänderungsmodell von Petty/Cacioppo (1986) und der erweiterten Rahmenkonzeption von Vogt (1998, S. 14–16) wird ein unterrichtsrelevanter Einfluss durch Interessen und Nicht-Interessen auf die Einstellungen zu Schule und Sachunterricht erwartet.

Die Längsschnittstudie PEIG² erfasst und beschreibt die Entwicklungen von Interessen bzw. Nicht-Interessen an naturwissenschaftlichen Gegenständen bei Kindern vom Vorschulalter bis zur Sekundarstufe I im Zusammenhang mit der Wirkung personaler Einflüsse im schulischen wie außerschulischen Kontext.

Von den Ergebnissen wird eine Erkenntnis über den Aufbau bzw. Verfall von naturwissenschaftlichen Interessen im Grundschulalter erwartet. Auf dieser Basis würde es möglich, gezielte didaktisch-methodische Unterstützungen im Unterricht zu bieten, um eine fach- und gegenstandsbezogene Interessiertheit aufseiten der Schüler zu fördern.

- 1 Die Studie wurde gefördert durch Mittel der DFG (VO 745/1-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.
- 2 Schulische und außerschulische personale Einflüsse bei Interessenentwicklungen von Grundschulkindern unter besonderer Berücksichtigung sachunterrichtlicher Gegenstandsbereiche (PEIG).

2. Theoretische Grundlagen

Das Spektrum von Interessen bis zu den Nicht-Interessen umfasst das situationale und das individuelle (aktualisierte) Interesse (Krapp 1998, S. 190; Prenzel/Lankes/Minsel 2000, S. 23) sowie das Desinteresse und die Abneigung (Upmeier zu Belzen/Vogt 2001, S. 19–23). Die Einordnung von Person-Gegenstands-Relationen in dieses Spektrum hängt davon ab, ob und wie sich aus Person-Gegenstands-Auseinandersetzungen Person-Gegenstands-Beziehungen (situationsspezifisch) oder Person-Gegenstands-Bezüge (zeit- und situationsübergreifend) entwickeln.

2.1 Interesse und Nicht-Interesse

Interesse

Das Interesse als aktuelle Beziehung zwischen Person und Gegenstand drückt sich in Handlungen aus, die in folgenden Merkmalskategorien positiv ausfallen: kognitive Ausprägung, emotionale Tönung und Wertaspekte (Schiefele u.a. 1983, S. 4, 9). Die Charakteristika der inhaltspezifischen Interessen-Konzeption (vgl. Kap. 3.4.1.) stehen entsprechend Krapp (1998, S. 193; 2001, S. 8–11) in Verbindung mit den drei angeborenen psychologischen Grundbedürfnissen („basic needs“) der Selbstbestimmungstheorie (Bedürfnis nach Kompetenz, Autonomie oder Selbstbestimmung und nach sozialer Eingebundenheit von Deci/Ryan 1993, S. 229).

Die intrinsische Qualität des Erlebens beim Lernen ist hoch, wenn in einer Situation diese Grundbedürfnisse befriedigt werden (Krapp 2001, S. 11). Positive Erfahrungen bei vorausgegangenen Person-Gegenstands-Auseinandersetzungen verstärken die Merkmalsausprägungen von Interesse. Auf diese Weise kann sich entsprechend Krapp (1998, S. 190) nach wiederholtem situationalem Interesse ein individuelles Interesse entwickeln.

Nicht-Interesse

Nicht-Interessen können sich entwickeln, wenn in der Lernsituation die Qualität des intrinsischen Erlebens niedrig ist (Upmeier zu Belzen/Vogt 2001, S. 21–23). Die Nicht-Interessen werden in zwei unterschiedlich starke Ausprägungen differenziert: das Desinteresse und die Abneigung (vgl. Lewalter/Schreyer 2000, S. 54, 55; Upmeier zu Belzen/Vogt 2001, S. 21–23).

Desinteresse beschreibt einen Zustand der Interesselosigkeit bzw. Gleichgültigkeit gegenüber einem Gegenstand, ohne dass eine Person-Gegenstands-Relation zustande kommt. Die Person ist passiv und setzt sich aus eigenem Antrieb nicht mit dem Gegenstand auseinander. Eine *Abneigung* basiert auf einer negativen Person-Gegenstands-Relation – verbunden mit negativen Gefühlen. Die Person meidet bzw. selektiert die weitere Informationsaufnahme. Die handelnde Auseinandersetzung mit dem Gegenstand wird aktiv vermieden.

2.2 *Einstellung zu Schule und Unterricht*

Bei der Einstellung handelt es sich wie bei Interesse um eine Person-Objekt-Relation, bei der jedoch anders als beim Interesse bezüglich der Merkmalskategorien Kognition, Emotion und Wert keine bestimmten Ausprägungen notwendig sind. Einstellung wird als Tendenz verstanden, Objekte, Personen oder Verhalten auf einem evaluativen Kontinuum zu bewerten.

Schulische Einstellungen lassen sich aus den Handlungen und verbalen Äußerungen der Schüler (affektiv und kognitiv) abschätzen (Bachmair 1969, S. 71). Durch die verbalen Aussagen können Einstellungen möglichst objektiv operationalisiert werden.

2.3 *Bedeutung von Interesse/Nicht-Interesse und Einstellung für schulisches Lernen*

Interessen bilden eine der wichtigsten Grundlagen im Kontext schulischen Lernens. Lehr-Lernprozesse sollten die intrinsische Qualität des situationalen Interesses unterstützen und nicht beeinträchtigen. Für das Interesse der Schüler hat die Lehrperson mit ihrer Kompetenz, ihren Interessen bzw. Nicht-Interessen an den Lerninhalten, ihren Einstellungen sowie der Unterrichtsgestaltung mit dem Grad an Unterstützung der „basic needs“ eine entscheidende positive bzw. negative Bedeutung.

Theoretische Rahmenkonzeption zur Klärung der Bedeutung von Interesse sowie Nicht-Interesse und Einstellungen im Unterricht ist das von Vogt (1998, S. 14–16) erweiterte Einstellungsänderungsmodell (Petty/Cacioppo 1986). Es wurde um die inhaltspezifischen Motivationskonstrukte Interesse und Nicht-Interesse sowie um Aspekte der Selbstbestimmungstheorie (Deci/Ryan 1993, S. 224–229) erweitert (Abb. 1, S. 294).

Sowohl das Interesse bzw. Nicht-Interesse an einem spezifischen Interesseobjekt als auch die Einstellung zu einem spezifischen Einstellungsobjekt sind Bestandteile der kognitiven Struktur einer Person. Insofern werden auch die „basic needs“ unterschiedlich erlebt. Somit hängt die Intensität einer Auseinandersetzung mit einem bestimmten Lerninhalt von der persönlichen Relevanz für die Person und der Erlebnisqualität im Unterricht ab. Es kommt zu entsprechenden Wegen der Verarbeitung, welche wiederum eine bestimmte Einstellung hervorbringt, die eine erneute Bereitschaft zu Person-Gegenstands-Auseinandersetzungen unterstützt, blockiert oder wieder herstellt.

2.4 *Fragestellungen des Projektes*

Mit fortschreitender Sozialisation verändern sich die Interessen und Nicht-Interessen strukturell. Veränderungen ergeben sich besonders dann, wenn soziale Bezugsgruppen und bedeutsame Bezugspersonen wechseln (Weinert 1998, S. 32): Eintritt in den Kindergarten, Übergänge in die Grundschule und weiterführende Schule.

Personale Einflüsse wirken im Rahmen von Interaktionen über die Befriedigung der psychologischen Grundbedürfnisse auf die Bereitschaft zum Aufbau einer Person-Gegenstands-Relation (Schiefele 2000, S. 230). Je nach subjektiver Wahrnehmung der Kin-

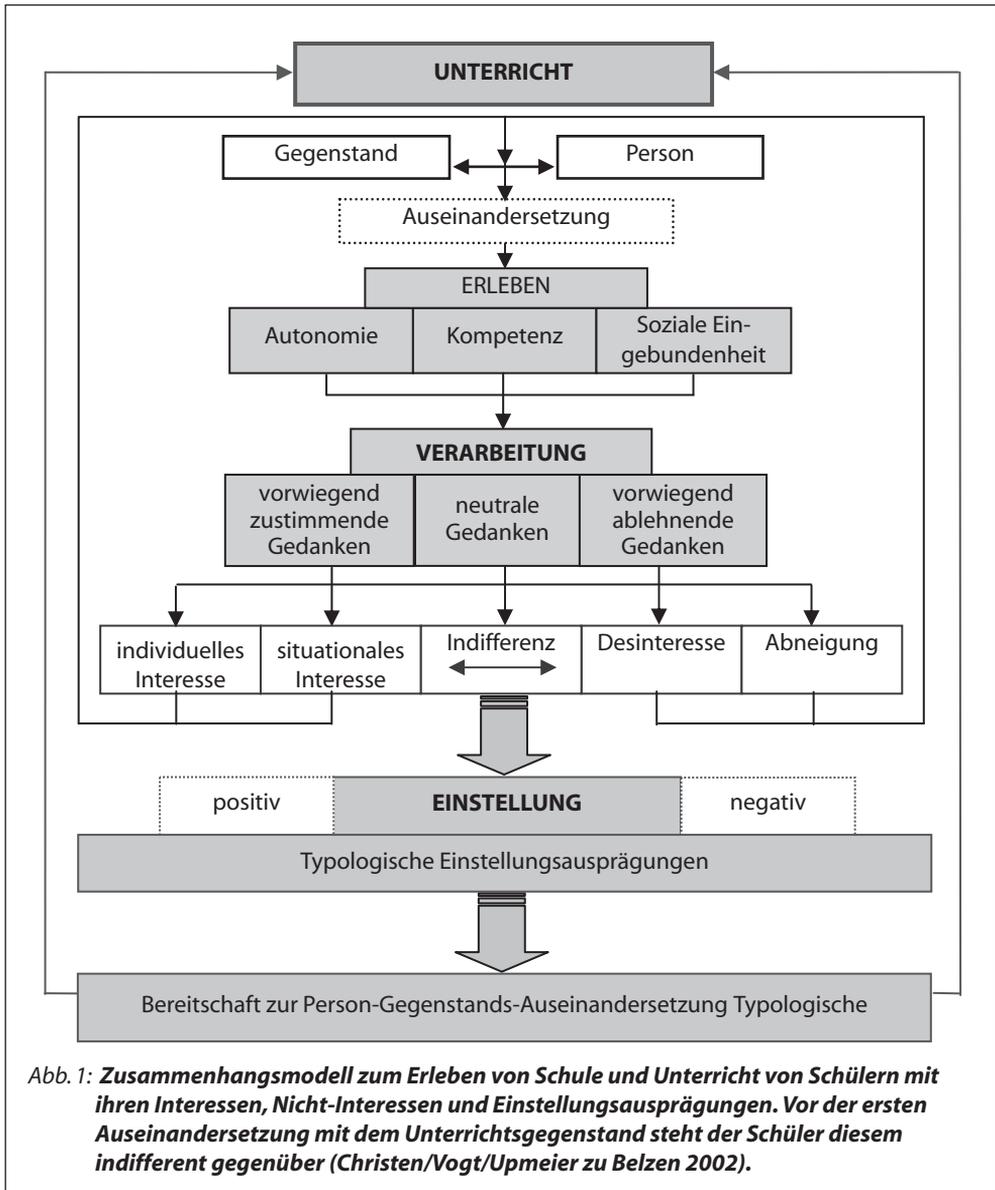


Abb. 1: **Zusammenhangsmodell zum Erleben von Schule und Unterricht von Schülern mit ihren Interessen, Nicht-Interessen und Einstellungsausprägungen. Vor der ersten Auseinandersetzung mit dem Unterrichtsgegenstand steht der Schüler diesem indifferent gegenüber (Christen/Vogt/Upmeier zu Belzen 2002).**

der (Upmeier zu Belzen 1998, S. 203) wird ein Einfluss vom Individuum als unterstützend oder kontrollierend empfunden. Vor diesem Hintergrund kann es je nach Erlebensqualität in allen Altersstufen zur Ausbildung von Interessen oder Nicht-Interessen kommen (Upmeier zu Belzen/Vogt 2001, S. 19). Neben den personalen Einflüssen durch Eltern, Verwandte und Peers gibt es Einflussmöglichkeiten durch Erzieherinnen und Lehrpersonen (Fölling-Albers 1995, S. 25; Hartinger 1997, S. 231; Vogt/Wieder 1999, S. 88; Vogt/Wieder/Schwaab 2000, S. 96–99).

Ausgehend von verschiedenen Strukturen der Interessen und Nicht-Interessen bei Grundschulkindern resultieren unterschiedliche Einstellungen zu Schule und Sachunterricht, welche im Verlauf des weiteren Unterrichtes auf die folgende Entwicklung von Interessen und Nicht-Interessen einwirken. Durch Interessen veränderte Einstellungen beeinflussen wiederum die weitere Entwicklung von Interessen (erweiterte Rahmenkonzeption nach Vogt 1998). Daraus ergeben sich folgende *Forschungsfragen*:

- Welche qualitativen Ausprägungen von Interessen bzw. Nicht-Interessen der Kinder gibt es in Kindergarten und Schule und wie verläuft der Entwicklungsprozess?
- Welchen Einfluss haben Erzieherinnen und Lehrpersonen auf die Entstehung und Entwicklung von Interessen und Nicht-Interessen der Kinder bzw. Schüler?
- Zeigen Schüler mit ähnlichen Strukturen von Interessen bzw. Nicht-Interessen gleiche Einstellungen zu Schule und Sachunterricht und welche typologischen Differenzierungen ergeben sich?

3. Methodik

Die Untersuchung ist als Längsschnitt angelegt, bei dem dieselben Probanden dreimal in sechs Jahren jeweils punktuell zu ihren Interessen und Nicht-Interessen und den Ursachen für Veränderungen umfassend befragt werden. Die Zeitpunkte der Befragungen sind das Ende der Kindergartenzeit, die zweite und vierte Klasse der Grundschule und die sechste Jahrgangstufe. Die zentrale Untersuchungsgruppe im Kindergarten (22 Mädchen und 21 Jungen aus ländlich gelegenen Vororten von Münster Westf.) wurde erstmals im Jahr 2000 befragt, während eine weitere Gruppe im Grundschulalter für Vergleichszwecke im Jahr 2000 begonnen wurde (22 Jungen und 21 Mädchen aus ländlich gelegenen Vororten von Münster Westf.).

Im Rahmen der ersten Befragungen des Längsschnittes im Jahr 2000 wurden die Kinder bzw. Schüler wie auch deren Eltern, Erzieherinnen und Lehrpersonen in halboffenen Interviews befragt. Zuerst wurden jeweils die Kinder Zuhause zu ihren Nicht-Interessen und dann zu ihren Interessen befragt, anschließend die Eltern zu den von ihrem Kind genannten Bereichen. Erzieherinnen und Lehrpersonen wurden in den entsprechenden Institutionen interviewt. Parallel wurden und werden alle Probanden ab dem Grundschulalter mittels eines eigens entwickelten Instrumentes zu ihrer Einstellung zu Schule und naturwissenschaftlichem Unterricht befragt. Die Lehrpersonen werden zu ihren Interessen sowie Nicht-Interessen an naturwissenschaftlichen Inhalten, zu den Einstellungen zu Schule und Sachunterricht sowie zur didaktisch-methodischen Ausgestaltung ihres Unterrichtes interviewt.

Zusätzlich werden für erste längsschnittliche Betrachtungen Daten aus einer ehemaligen Pilotstudie zur Entwicklung von Interessen (Vogt/Wieder 1999, S. 79–91) einbezogen. Dabei wurden 13 der 1994 ausgewählten Kindergartenkinder (6 Mädchen und 7 Jungen einer Vorstadt von Düsseldorf) im Jahr 2000 zum vierten Mal befragt (Sekundarstufe I).

3.1 *Interviewleitfäden für die Kinder und ihre Eltern*

Zur Erfassung der Person-Gegenstands-Relationen der Kinder wurden auf der Grundlage von Vogt/Wieder (1999, S. 82) altersgemäße halbstandardisierte Interviewleitfäden für Probanden des Kindergartens, der Grundschule und der Sekundarstufe I entwickelt. Die Leitfäden bestehen aus Fragen zu den theoretischen Konstrukten Nicht-Interesse (Upmeier zu Belzen/Vogt 2001, S. 21–23, Erfassen des/der Gegenstände – Interesse und Nicht-Interesse –, Ausprägung der Merkmale, Auslöser, Genese und Entwicklung, positive und negative Erfahrungen), Interesse nach Krapp (2000, Erfassen des/der Gegenstände – Interesse und Nicht-Interesse –, Ausprägung der Merkmale, Auslöser, Genese und Entwicklung, positive und negative Erfahrungen), Selbstbestimmungstheorie von Deci/Ryan (1993, S. 224–229, Autonomieerleben, Kompetenzerleben Zuhause und in der Schule, Aspekte der sozialen Eingebundenheit, Unterstützung, Kontrolle, Feedback) und der Flow-Theorie von Csikszentmihalyi/Schiefele (1993, S. 209–211, Flow-Erleben). Mit den Interviewdaten kann die Struktur und Qualität der Interessen und Nicht-Interessen der einzelnen Probanden differenziert beschrieben werden.

Der Interviewleitfaden, nach dem der hauptsächlich betreuende Elternteil befragt wird, greift die vom Kind genannten Gegenstandsbereiche auf und erfasst aus Elternsicht die Umstände der Genese und Entwicklung, den Einfluss der Eltern und Geschwister sowie anderer Personen nach den oben genannten theoretischen Aspekten. Die Interessegegenstände wurden mit BIOLINKS validiert (Schröer/Wessiepe, 1997, S. 39–53).

3.2 *Interviewleitfäden für Erzieherinnen und Lehrpersonen*

Auf der Grundlage des Interviewleitfadens von Upmeier zu Belzen (1998, S. 21) wurden halboffene Interviewleitfäden für Erzieherinnen und Lehrpersonen zur Erfassung der biologie- bzw. sachunterrichtlich orientierten Interessen und Nicht-Interessen der Probanden entwickelt. Im Fokus der Untersuchungen in der Schule liegt zusätzlich die Erfassung wesentlicher Merkmale der didaktisch-methodischen Ausgestaltung des Unterrichts sowie der Grad an Unterstützung bzw. Kontrolle im Sinne der Selbstbestimmungstheorie (Deci/Ryan 1993, S. 224–229) durch den eigenen Unterricht aus Sicht der Lehrperson sowie die Einstellung zu Schule und zum naturwissenschaftlichen Unterricht.

3.3 *Fragebogen zur Einstellung zu Schule und Sachunterricht*

Zur Erfassung der Einstellung der Schüler zu Schule und Sachunterricht wurde mit einer von der Längsschnittstudie unabhängigen Gruppe eine altersgemäße Likert-Skala entwickelt (Christen/Vogt/Upmeier zu Belzen 2001, S. 5–6). An der Vor- und Hauptuntersuchung nahmen dieselben Schüler derselben Grundschulen und Klassen (N = 344)

teil. Ausgewählte Projektkinder (Grundschule, N = 19 [von 43³]) wurden anschließend mittels eines Einstellungsfragebogens befragt und dienen als erste Grundlage für eine Verknüpfung mit den entsprechenden Daten zu deren Interessen und Nicht-Interessen.

3.4 Methoden der Auswertung

Interesse

Alle Interviews wurden auf Tonträgern aufgenommen, transkribiert und mithilfe des Textanalyseprogramms WinMAX (Kuckartz 1999, S. 30–36) inhaltsanalytisch kategorisiert. Gemäß des Kriteriums der interpersonalen Konsensbildung (konsensuelle Validierung) wurden die Interviewdaten (Interviewäußerungen und Interaktionsverlauf) von mehreren Ratern verglichen, analysiert und somit validiert.

Zur Unterscheidung der Qualität der positiven Person-Gegenstands-Bezüge werden die von Vogt/Wieder (1999, S. 83–85) für die Pilotstudie aufgestellten Interessenmerkmale herangezogen. Dabei wird von definierten Interessen gefordert, dass sie die *Mindestanforderungen* von Schiefele u.a. (1983, S. 21) erfüllen. Diese „Mindestanforderungen“ beziehen sich auf die Merkmale Kognition, Emotion und Wert. Darüber hinaus werden entsprechend Schiefele u.a. (1983, S. 31) zur *qualitativen Unterscheidung der Person-Gegenstands-Bezüge* weitere Merkmale heran gezogen. Für die Pilotstudie stellten Vogt/Wieder (1999, S. 83–85) auf Grundlage der Beschreibungsmerkmale von Fink (1992, S. 54–57) und der von Vogt/Wieder (1999, S. 84, Tab. 3) erhobenen Daten aus dem Kindergarten folgende Kriterien zur Unterscheidung der Ausprägungsgrade der einzelnen Person-Gegenstands-Bezüge auf: Dauerhaftigkeit und Häufigkeit; Komplexität (komplexer, themenzentrierter Person-Gegenstands-Bezug); Vorkommen; Anzahl und Art der einzelnen Person-Gegenstands-Auseinandersetzungen.

Neben der Unterscheidung in definierte (D 3 und D 4) und nicht definierte Interessen (D 1 und D 2) wird unter Einbeziehung sämtlicher theoretischer Aspekte eine umfassende qualitative Bewertung aller erhobenen positiven Person-Gegenstands-Bezüge in der Längsschnittstudie ermöglicht.

Nicht-Interesse

Zur qualitativen Unterscheidung der Nicht-Interessen wurden analog zu den Interessenmerkmalen von Vogt/Wieder (1999, S. 84) theoriegeleitet Basisanforderungen und abstufende Merkmale festgelegt (Upmeier zu Belzen/Vogt 2001, S. 21–23). Die Merkmale sind geordnet nach *Basisanforderungen für Abneigung* (vgl. Tab. 1, A) und *Merkmalen zur qualitativen Unterscheidung* (vgl. Tab. 1, B) in Desinteressen und Abneigungen. Sind *alle* Basisanforderungen erfüllt, wird dieses erhobene Nicht-Interesse als Abneigung bezeichnet, ansonsten wird von einem Desinteresse gesprochen.

3 Bei 24 Grundschulern (gesamte Klasse) lehnte die Lehrperson Erhebungen im Bereich der Einstellung zu Schule und Sachunterricht strickt ab.

Tab. 1: Unterscheidung von Nicht-Interessen in Dregees (D) anhand abstufender Merkmale (Desinteressen D-1 und D-2 und Abneigungen D-3 und D-4) analog zu den von Vogt/Wieder (1999, S. 84) aufgestellten Interessenmerkmalen.

Merkmale der negativen PG-Bezüge	Desinteresse		Abneigung	
	D-1	D-2	D-3	D-4
A) Basisanforderungen für Abneigungen				
spezifischer negativer PG-Bezug	±	±	+	+
selektives Wissen	±	±	+	+
negative Gefühle	±	±	+	+
negative Wertschätzung	±	±	+	+
Vermeidung von PGA	±	±	+	+
B) Merkmale zur qualitativen Unterscheidung				
Art der Vermeidung				
Ablehnung subjektiv begründet	±	±	+	+
aktive Vermeidung von PGA		±	+	+
keine PGA trotz Zwang/Kontrolle				+
Art der negativen Wertschätzung				
Wertschätzung besonders negativ		±	+	+
zugenommen negative Einstellung				+
Emotion				
besonders negative Emotion		±	+	+
Kognitiver Aspekt				
Vermeidung von Informationsbeschaffung	±	±	+	+
Ausblenden von Handlungskompetenz				+
Art des spezifischen PG-Bezuges				
Vermeidung zu Hause und in anderer Umgebung	±	±	+	+
themenzentrierter PG-Bezug		±	+	+
Dauer länger als 1 Jahr				+
PGA nur fremdintentional				+
Grundlagen der Merkmalsaufstellung: A) nach Upmeier zu Belzen/Vogt (2001, S. 21-23), B) nach den Beschreibungsmerkmalen für Person-Gegenstands-Bezüge von Fink (1991, S. 55-57) und der Selbstbestimmungstheorie von Deci/Ryan (1993, S. 229) sowie Untersuchungsergebnisse des Längsschnittes. (PG-Bezüge = Person-Gegenstands-Bezüge; PG-Bezug = Person-Gegenstands-Bezug; PGA = Person-Gegenstands-Auseinandersetzung; + = Merkmal beim Probanden erfüllt; ± = Merkmal beim Probanden fakultativ erfüllt)				

Die Einstufung der Desinteressen (D -1 und D -2) sowie der Abneigungen (D -3 und D -4) erfolgt jeweils aufgrund der Anzahl der zutreffenden Unterscheidungsmerkmale (Tab. 1, B).

Folgende Merkmale weisen auf einen „ausgeprägten“ negativen Person-Gegenstands-Bezug (Abneigung) hin: subjektiv begründete Ablehnung eines Gegenstandes, Vermeidung einer Person-Gegenstands-Auseinandersetzung auch unter Zwang, besonders negative Emotionen und Wertschätzung, Vermeidung von Informationsbeschaffung und Ausblenden von Handlungskompetenzen, Dauer des Nicht-Interesses seit mindestens einem Jahr.

Einstellung

Die Analyse der erhobenen Daten bezüglich der Einstellung in der Grundschule (längsschnittunabhängig, N = 344) wurde mit WINMIRA 2001 für Mischverteilungsmodelle (hierbei das Mixed Rasch-Modell) durchgeführt (vgl. Christen 2001, S. 6–7). Das Mixed Rasch-Modell ist ein psychometrisches Modell, mit dem es möglich ist, quantitative und qualitative Persönlichkeitsunterschiede zu erfassen und Personen dementsprechend zu klassifizieren (vgl. Rost 1996, S. 169–177).

4. Darstellung der Ergebnisse

4.1 Interessen und Nicht-Interessen

Interessen

Bei allen befragten Kindern konnten spezifische Person-Gegenstands-Bezüge festgestellt werden, wobei nicht alle den theoretischen Anforderungen für Interessen entsprechen.

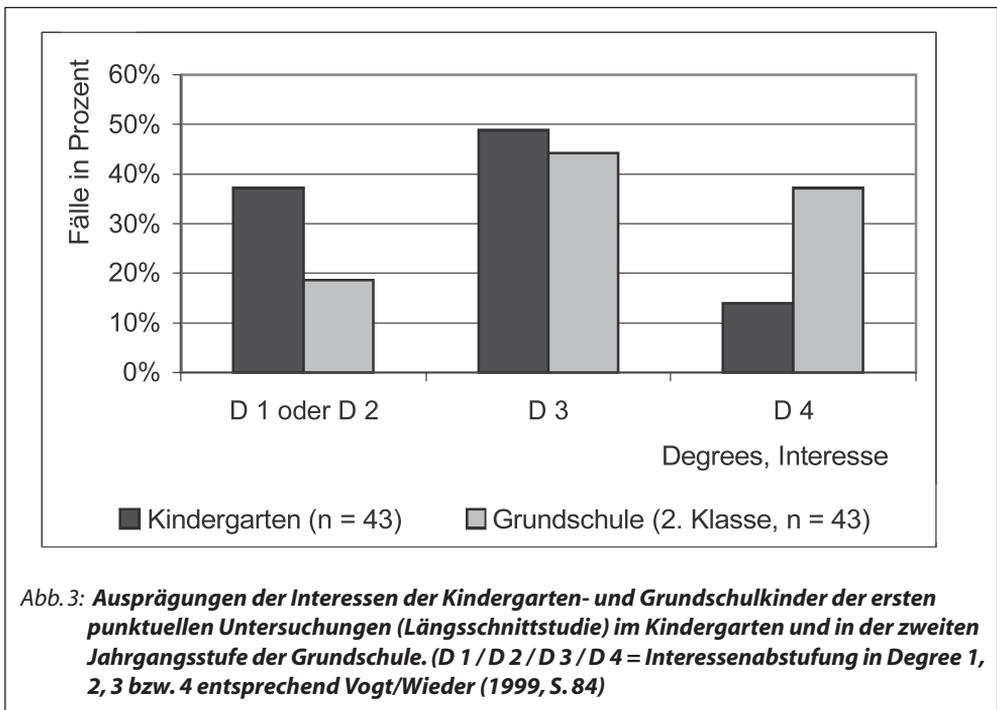
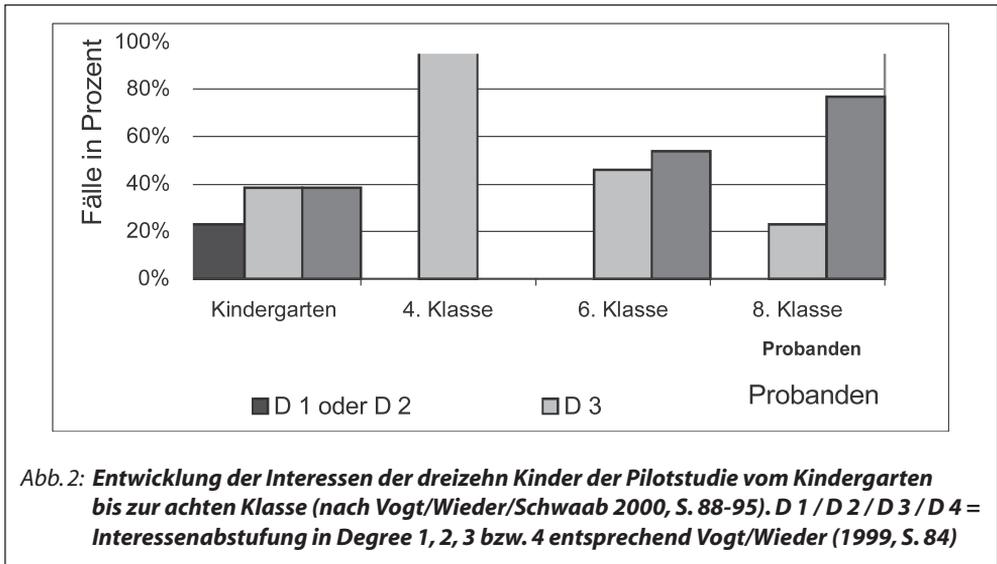
Im Rahmen der Pilotstudie hatten lediglich drei der 13 Kinder im Kindergarten 1994 nicht definierbare Interessen (D 1 oder D 2) und 10 Kinder ein definiertes Interesse. Jeweils 5 dieser Interessen konnten der Gruppe D3 bzw. D4 zugeordnet werden. Im Grundschulalter verfügten alle 13 Kinder über ein definiertes Interesse der Stufe D 3 (Abb. 2, S. 300). In der Sekundarstufe I konnten zunehmend Interessen der Gruppe D 4 zugeordnet werden.

Auch in der Längsschnittstudie (PEIG) verfügten die Grundschul Kinder im Vergleich zu den Kindergartenkindern in der Summe (Abb. 3, S. 300) vermehrt über qualitativ „höherwertige“ Interessen der Gruppen D 3 und D 4.

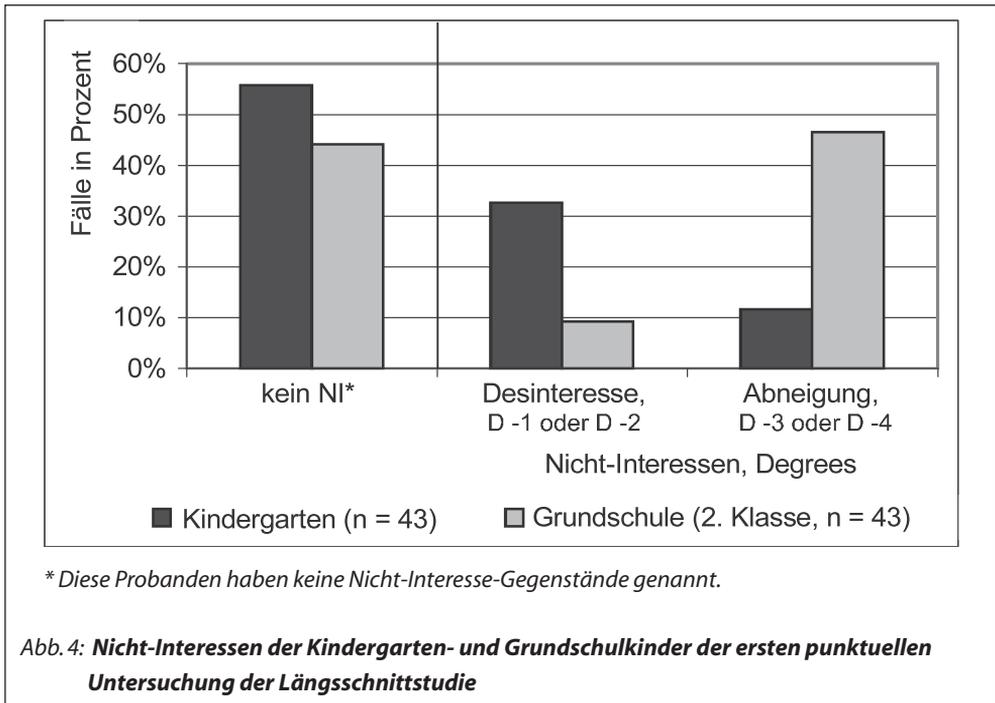
Nicht-Interessen

Vier der 13 Schüler der Pilotstudie benannten in der Sekundarstufe I einen Gegenstand, dem sie desinteressiert gegenüberstanden und weitere vier einen Abneigungs-Gegenstand. Die anderen fünf Schüler nannten weder ein Desinteresse noch eine Abneigung⁴.

4 Die Probanden der Pilotuntersuchung wurden zunächst jeweils lediglich zu ihren Interessen befragt, im Jahr 2000 erstmalig auch zu ihren Nicht-Interessen.



Von den 43 Kindergartenkindern der Längsschnittstudie nannten 33% einen Gegenstand zum Desinteresse und 12% einen Abneigungs-Gegenstand. Bei 56% der Kinder wurde kein Nicht-Interesse festgestellt. Von den 43 Probanden im Grundschulalter wiesen 9% ein Desinteresse und 47% eine Abneigung auf, wobei 44% keinen Nicht-Interesse-Gegenstand nannten (Abb. 4).



Bei den untersuchten Fällen im Vorschulalter ($N = 43$) gab es Ansätze von Nicht-Interessen, aber nur wenige ausgeprägte definierte Nicht-Interessen. Die Kinder im Schulalter ($N = 43$) hatten dagegen vermehrt definierte Nicht-Interessen.

Qualitative Entwicklungen der Interessen und Nicht-Interessen

Die 43 Kindergartenkinder mit definierten Interessen (D 3 und D 4) beschäftigten sich mit ihren Interessegegenständen mindestens einmal wöchentlich Zuhause und im Kindergarten. Informationen zu diesen Gegenständen bekamen sie überwiegend von Personen aus ihrem nahen sozialen Umfeld.

Mit Schuleintritt informierten sich die Kinder zu ihren Interessen vermehrt auch über z.B. Fachbücher oder Filme und erlangten dadurch ein fundierteres fachliches Wissen. Ebenso beschäftigten sich die Kinder mit wachsendem Alter mindestens zwei Jahre mit einem gleichen Interessegegenstand. Im Gegensatz zur Grundschule wandten sich die Kinder in der Sekundarstufe wieder vermehrt Personen zu (insbesondere Freunde), um Informationen hinsichtlich ihrer Interessegegenstände zu erhalten. Vom Kindergarten über die Grundschule bis zur Sekundarstufe I wurden die Interessen (D 3) bei den Kindern stärker und durch eine intensivere, vielfältige Auseinandersetzung mit den Gegenständen komplexer (erhöhte Anzahl der erfüllten Merkmale zur qualitativen Unterscheidung Vogt/Wieder 1999, S. 84, Tab. 3). Aufgrund dieser qualitativ höherwertig einzustufenden Person-Gegenstands-Bezüge wurden die Interessen der Kinder in der Sekundarstufe I vermehrt der Gruppe D 4 zugeordnet. Bis auf eine Ausnahme verfügten alle 13 Pilotkinder der Sekundarstufe I über mindestens zwei definierte Interessengebiete.

te (D 3 oder D 4). Die Ergebnisse der Pilotstudie zeigen, dass in der Sekundarstufe I 11 der 13 Probanden ihre favorisierten Person-Gegenstands-Bezüge aus der Grundschule beibehalten bzw. spezialisiert (z.B.: Musik-Flöte-Klarinette) oder einen favorisierten Person-Gegenstands-Bezug aus dem Vorschulalter erneut aufgegriffen haben. Die Probanden aus Kindergarten und Grundschule fühlten sich besonders stark sozial eingebunden. Mehrfach wurden zwei der „basic needs“ erfüllt. Ab dem Schuleintritt gaben die Schüler vermehrt an, sich bei den Person-Gegenstands-Auseinandersetzungen auch kompetent gefühlt zu haben.

Auch die *Nicht-Interessen* waren mit steigendem Alter von längerer Dauer (mindestens ein Jahr). In allen Altersstufen verfügten Kinder mit Abneigungen über ein selektives Wissen bezüglich des Nicht-Interesse-Gegenstandes. Dementsprechend konnten die Probanden bereits in Einzelfällen im Kindergartenalter und zunehmend im Schulalter ihre Ablehnung subjektiv begründen. Des Weiteren lehnten sie weitere Auseinandersetzungen mit bzw. Informationen über den Nicht-Interesse-Gegenstand – auch unter Zwang – ab. Bei den Nicht-Interessen stand in allen Altersstufen das Bedürfnis nach Autonomieerleben im Vordergrund. Die Kinder und Schüler fühlten sich in ihrer ablehnenden Haltung autonom. Für sie war es unwichtig, ob sich andere Personen mit dem entsprechenden Gegenstand beschäftigten. Wiederholt wurde auch ein fehlendes Kompetenzerleben als Grund für ein Nicht-Interesse angegeben.

4.2 Personale Einflüsse auf die Entwicklung von Interessen und Nicht-Interessen

Personale Einflüsse auf die Entwicklung von Interessen

Die Eltern hatten einen positiven Einfluss auf die Interessenentwicklungen ihrer Kinder durch gemeinsame Erlebnisse oder durch eine Interessiertheit an gleichen Gegenstandsbereichen. Sie unterstützten ihre Kinder z.B. durch gemeinsame Gespräche, Aktivitäten, Anmeldung in Vereinen oder Hilfe bei der Beschaffung von Materialien (z.B. Fachliteratur). Nach den Eltern folgten die Peers, Geschwister und Verwandte mit ihren Einflüssen auf die Entwicklung der Interessen und Nicht-Interessen der Kinder.

Eine positive Beeinflussung durch eine Lehrperson mit ihren eigenen Interessen sowie ihrer Unterrichtsgestaltung wurde lediglich von einem Kind berichtet, bei dem im zweiten Schuljahr Interesse an Vögeln angestoßen wurde. Des Weiteren wurde bei einem Kindergartenkind das Interesse an Pflanzen durch das Pflanzen-Interesse der Erzieherin sowie Angebote zum Thema Pflanzen im Kindergarten positiv beeinflusst.

Personale Einflüsse auf die Entwicklung von Nicht-Interessen

Als Anstoß der Entwicklung von Nicht-Interessen wurden in allen Altersstufen überwiegend vorausgegangene Erfahrungen mit dem Gegenstand wie beispielsweise ein negatives Kompetenzerleben beim Malen im Kindergarten genannt. Eine personale Beeinflussung fand unter anderem dadurch statt, dass Bezugspersonen wie Eltern, Geschwister und Peers gleiche Gegenstandsbereiche ablehnten. Mit zunehmendem Alter der Kinder

wurde jedoch das Favorisieren bestimmter Gegenstandsbereiche von Bezugspersonen (Familienangehörige, Peers, Lehrer) auch als Anstoß und Einflussfaktor für die Entwicklung von Nicht-Interessen gegenüber dem entsprechenden Gegenstand angegeben. Der Versuch der Eltern, Entwicklungen von Nicht-Interessen entgegen zu wirken wurde von einigen Probanden der Sekundarstufe I ebenfalls als negativ empfunden. Der Einfluss der Lehrpersonen bei der Ausbildung von Nicht-Interessen trat deutlich bei drei Schülern der achten Klasse aus der Pilotstudie hervor. Sie gaben den Biologieunterricht und die Fachlehrpersonen als negative Einflussfaktoren an, welche die Ausbildung bzw. Entwicklung biologischer Nicht-Interessen vorangetrieben hat.

4.3 *Typologische Einstellungsausprägungen*

Aus den in der Grundschule erhobenen Daten wurden von Christen/Vogt/Upmeier zu Belzen (2001, S. 8–10) drei typologische Einstellungsausprägungen identifiziert. Der „*Lernfreude-Typ*“ zeichnet sich durch eine grundlegend positive Einstellung zu Schule und Sachunterricht aus. Der „*Gelangweilt-Frustrierte Typ*“ hat demgegenüber eine negativere Einstellung zu Schule und Sachunterricht. Es gibt dabei zwei unterscheidbare Richtungen: Langeweile und Frustration. Der „*Zielorientierte Leistungs-Typ*“ zeigt ebenfalls eine negativere Einstellung zu Schule und Sachunterricht als der Lernfreude-Typ. Die Schüler wissen genau, was sie in der Schule lernen wollen, insbesondere im Hinblick auf ihre individuelle Zukunft.

4.4 *Interessen/Nicht-Interessen und Einstellungen*

Bei den Probanden aus der Grundschule (N = 19) zeigt sich, dass Schüler mit ähnlichen Strukturen ihrer Interessen bzw. Nicht-Interessen vergleichbare Einstellungen zu Schule und Sachunterricht hatten. Acht Schüler gehören zur Gruppe des Lernfreude-Typs, sechs zur Gruppe des Gelangweilt-Frustrierten Typs und fünf zur Gruppe des Zielorientierten Leistungs-Typs.

Alle Grundschüler der Gruppe des Lernfreude-Typs zeigten mindestens ein definiertes Interesse. Dabei hatte wenigstens ein definiertes Interesse die qualitativ höchste Ausprägung (D 4). Desinteressen und Abneigungen wurden vereinzelt festgestellt (N = 3). Alle Grundschulkind der Gruppe des Gelangweilt-Frustrierten Typs zeigten ein definiertes Interesse. Im Bereich der Nicht-Interessen traten sowohl Desinteressen als auch Abneigungen relativ häufiger auf als beim Lernfreude-Typ. In der Gruppe des Zielorientierten Leistungs-Typs verfügten die Grundschüler wenigstens über ein definiertes Interesse mit der Ausprägung D 4. Im Gegensatz zu den Gruppen des Lernfreude- und des Gelangweilt-Frustrierten Typs zeigten die Kinder im Spektrum der Nicht-Interessen ausschließlich Abneigungen und keine Desinteressen. Zusammenfassend zeigt sich, dass die Schüler innerhalb der Gruppe des Lernfreude-Typs und des Zielorientierten Leistungs-Typs alle mindestens ein definiertes Interesse mit der Ausprägung D 4 hatten. In

der Gruppe des Gelaugweilt-Frustrierten Typs zeigten fünf der sechs Schüler je ein definiertes Interesse mit höchster Ausprägung (D 4). Bemerkenswert ist, dass alle Schüler der Gruppe des Zielorientierten Leistungs-Typs neben einem definierten Interesse (D 4) eine Abneigung und kein Desinteresse hatten. Dagegen traten Abneigungen in der Gruppe des Lernfreude-Typs nahezu nicht auf.

5. Diskussion der Ergebnisse

5.1 Entwicklung der Interessen/Nicht-Interessen

Bemerkenswert ist, dass die Kinder, die bereits im Vorschulalter über Interessen verfügen, mit zunehmendem Alter „qualitativ höherwertigere“ Interessen vorwiesen. Entsprechend Vogt/Wieder (1999, S. 84) waren die Ausprägungen der Interessen zunehmend stärker und komplexer (Merkmale zur qualitativen Unterscheidung der Interessen) und konnten der Gruppe D 4 zugeordnet werden. Nahezu alle Schüler der Sekundarstufe I der Pilotstudie (12) und einige Grundschüler des Längsschnittes hatten mindestens zwei definierbare Interessen (vgl. Vogt/Wieder/Schwaab 2000, S. 96). In dieser Weise unterscheidet sich nach bisherigen Erkenntnissen die Struktur der Interessen im Kindergartenalter von der im Grund- bzw. Sekundarstufenalter. Dieses untermauert die Annahme, dass Kinder, die bereits im Vorschulalter ein definiertes Interesse ausgebildet haben, mit zunehmendem Alter ihr Interesse intensivieren oder andere bzw. zusätzliche Interessen ausbilden (Vogt/Wieder/Schwaab 2000, S. 92) sowie ruhende Interessen aktualisieren (Schiefele 2000, S. 230). Auch Weinert (1998, S. 32) geht davon aus, dass es zwar keine alterstypischen Entwicklungsverläufe der Interessen gibt, aber individuelle Präferenzen auf jeden Fall stabil sind solange kein Wechsel bedeutsamer Bezugspersonen oder sozialer Bezugsgruppen stattfindet.

5.2 Personaler Einfluss und Interessen/Nicht-Interessen

Der personale Einfluss durch die Familie ist als Basis zur Ausbildung einer generellen Bereitschaft zu Person-Gegenstands-Auseinandersetzungen bzw. Interessen im Laufe der Persönlichkeitsentwicklung eines Kindes von besonderer Bedeutung (Kasten/Krapp 1986, S. 183). Die Aufgabe von Kindergarten und Schule besteht unter anderem darin, vorhandene Interessen aufzugreifen und Kinder mit Gegenständen vertraut zu machen, die sie in ihren Familien nicht kennen gelernt haben (Hartinger 1997, S. 231). Dagegen deuten die Ergebnisse dieser Untersuchung an, dass das mögliche Interessenpotenzial der Schüler häufig nicht ausgeschöpft wird (vgl. Hansen/Klinger 1997; S. 119). Das oben genannte Beispiel „Interesse an Vögeln“ zeigt, dass Unterricht, in dem die „basic needs“ erfüllt werden bzw. Autonomieunterstützung, Handlungsorientierung und Selbstbestimmung erfahrbar werden, Interessen anregen kann (vgl. Hartinger 1997, S. 224, 225). Stattdessen wurde in anderen Fällen durch die Lehrperson mit ihrem Unter-

richt die Entwicklung von Nicht-Interessen angestoßen (vgl. Prenzel 1997, S. 41). „Lehr-Lernumgebungen könnten allein schon dann eine andere Qualität gewinnen, wenn Demotivierungen unterblieben“ (Prenzel 1997, S. 42).

5.3 *Interessen/Nicht-Interessen und Einstellungen*

Schülergruppen differenziert nach den drei typologischen Einstellungsausprägungen zeigen ähnliche Qualitäten ihrer Interessen bzw. Nicht-Interessen. Aufgrund der Interviewdaten ergeben sich folgende Hinweise: Schüler, die in vielen Bereichen offen für Person-Gegenstands-Auseinandersetzungen bzw. Interessen sind, gehören dem Lernfreude-Typ an. Möglicherweise handelt es sich um „Mitläufer“ und hoch intrinsisch motivierte Schüler. Demgegenüber gibt es Schüler, die sowohl über „hochwertige“ Interessen sowie starke Abneigungen verfügen. Diese Schüler wissen genau, was sie wollen und was sie nicht wollen, sind zukunfts- und leistungsorientiert und gehören dem Zielorientierten Leistungs-Typ an. Bei Schülern, die über nicht so stark ausgebildete Interessen bzw. Nicht-Interessen verfügen, sind die Ursachen oft Langeweile bzw. Frustration. Bei diesen Schülern wurden die „basic needs“ im Unterricht häufig nicht erfüllt und es gab keine situationalen Interessen. Den Interessen wurde vorzugsweise im außerschulischen Bereich nachgegangen.

5.4 *Fazit*

Je früher Kinder definierte Interessen ausbilden, desto besser sind sie in der Lage, erneut Interessen aufzubauen (vgl. Wieder 1999, S. 27). Im Schulbereich wirkt sich diese Fähigkeit bei entsprechender didaktisch-methodischer Ausgestaltung des Unterrichtes (Kleine/Vogt 2001, S. 187–191) vorteilhaft auf die Entwicklung schulischer Interessen aus, wie dieses auch die Ergebnisse des Projektes von Möller/Stern aufzeigen.

Durch eine nachfolgende positive Entwicklung der Interessen wird auch eine Entwicklung positiver Einstellungen zu Schule und naturwissenschaftlichem Unterricht angestoßen.

Literatur

- Bachmair, G. (1969): Einstellungen von Schülern zum Lehrer und zum Unterrichtsfach. Inaugural-Dissertation der Philosophischen Fakultät der Friedrich-Alexander-Universität. Erlangen-Nürnberg.
- Christen, F./Vogt, H./Upmeier zu Belzen, A. (2001): Einstellung von Schülern zu Schule und Sachunterricht – Erfassung und Differenzierung von typologischen Einstellungsausprägungen bei Grundschulern. In: Berichte des Institutes für Didaktik der Biologie Münster, IDB 10, S. 1–16.
- Csikszentmihalyi, M./Schiefele, H. (1993): Die Qualität des Erlebens und der Prozess des Lernens. In: Zeitschrift für Pädagogik 39, H. 2, S. 207–221.

- Deci, E.L./Ryan, R.M. (1993): Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik* 39, H. 2, S. 223–238.
- Fink, B. (1992): Interessenentwicklung im Kindesalter aus der Sicht einer Person-Gegenstands-Konzeption. In: Krapp, A./Prenzel, M.: *Interesse, Lernen, Leistung*. Münster: Aschendorff, S. 53–83.
- Fölling-Albers, M. (1995): Interessen von Grundschulkindern. In: *Grundschule* 27, H. 6, S. 24–26.
- Hansen, K.-H./Klinger, U. (1997): Interesse am naturwissenschaftlichen Lernen im Sachunterricht – Ergebnisse einer Schülerbefragung. In: Marquardt-Mau, B./Köhnlein, Hartinger, A. (1997): *Interessenförderung: Eine Studie zum Sachunterricht*. Forschungen zur Didaktik des Sachunterrichts. Bd. 2. Bad Heilbrunn: Julius Klinkhardt.
- Kasten, H./Krapp, A. (1986): Das Interessen-Genese-Projekt – eine Pilotstudie. In: *Zeitschrift für Pädagogik* 32 H. 2, S. 175–188
- Kleine, A./Vogt, H. (2001): Einfluss der methodisch-didaktischen Ausgestaltung des Unterrichtes auf die Einordnung des Faches Sachunterricht in den Fächerkanon der Grundschule. In: Klee, R./Bayrhuber, H. (Hrsg.): *Biowissenschaften in der Schule und Öffentlichkeit*. Tagungsband zur gleichnamigen Tagung der Sektion Biologiedidaktik in Rendsburg 2001. Kiel: IPN, S. 187–190.
- Krapp, A. (1998): Entwicklung und Förderung von Interesse im Unterricht. *Psychologie, Erziehung, Unterricht*, H. 44, S. 185–201.
- Krapp, A. (2000): Individuelle Interessen als Bedingung lebenslangen Lernens. In: Achtenhagen, E./Lempert, W.: *Entwicklung eines Programmkonzeptes „Lebenslanges Lernen“ für das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie*. Opladen: Leske + Budrich.
- Krapp, A. (2001): An Educational-Psychological Theory of Interest and Its Relations to Self-Determination Theory. In Deci, E.L./Ryan R.M. (Eds.): *The Handbook of Self-Determination Research*. University of Rochester Press.
- Kuckartz, U. (1999): *Computergestützte Analyse qualitativer Daten. Eine Einführung in Methoden und Arbeitstechniken*. Opladen Wiesbaden: Westdeutscher Verlag.
- Lewalter, D./Schreyer, I. (2000): Entwicklung von Interessen und Abneigungen – zwei Seiten einer Medaille? Studie zur Entwicklung berufsbezogener Abneigungen in der Erstausbildung. In: Schiefele, U./Wild, K.-P. (Hrsg.): *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung*. Münster: Waxmann, S. 53–72.
- Löwe, B. (1986): Interessenverfall im Biologieunterricht. *Unterricht Biologie* 124, S. 62–65
- Petty, R.E./Cacioppo, J.T. (1986): *Communication and persuasion. Central and peripheral routes to attitude change*. New York: Springer.
- Prenzel, M. (1997): Sechs Möglichkeiten, Lernende zu demotivieren. In: Gruber, H./Renkl, A. (Hrsg.): *Wege zum Können*. Bern: Hans Huber, S. 33–44.
- Prenzel, M./Lankes, E.-M./Minsel, B. (2000): Interessenentwicklung in Kindergarten und Grundschule: Die ersten Jahre. In: Schiefele, U./Wild, K.-P. (Hrsg.): *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung*. Münster: Waxmann, S. 11–30.
- Rost, J. (1996): *Lehrbuch Testtheorien, Testkonstruktionen*. Bern Göttingen Toronto: Hans Huber.
- Schiefele, H. (2000): Befunde – Fortschritte – neue Fragen. In: Schiefele, U./Wild, K.-P. (Hrsg.): *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung*. Münster: Waxmann, S. 227–241.
- Schiefele, H./Prenzel, M./Krapp, A./Heiland, A./Kasten, H. (1983): *Zur Konzeption einer pädagogischen Theorie des Interesses*. Gelbe Reihe – Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie Nr. 6. München: Institut für Empirische Pädagogik und Pädagogische Psychologie der Universität München.
- Schiefele, U./Schiefele, H. (1997): Motivationale Orientierungen und Prozesse des Wissens. In: Gruber, H./Renkl, A. (Hrsg.): *Wege zum Können*. Bern: Hans Huber, S. 14–31.
- Schröder, T./Wessiepe, K. (1997): Hypertext als Instrument zur Erhebung von Biologieinteressen. – Entwicklung und Einsatz des Autorensystems „BIOLINKS“. In: *Berichte des Institutes für Didaktik der Biologie Münster, IDB* 6, S. 39–53

- Upmeier zu Belzen, A. (1998): Der Zusammenhang zwischen Biologieunterricht und biologieorientiertem Interesse in einer 6. Klasse eines Gymnasiums. Europäische Hochschulschriften, Reihe 11, Bd. 735. Frankfurt a. M.: Peter Lang.
- Upmeier zu Belzen, A./Vogt, H. (2001): Interessen bei Grundschulkindern – Theoretische Basis der Längsschnittstudie PEIG. In: Berichte des Institutes für Didaktik der Biologie Münster, IDB 10, S. 17–31.
- Vogt, H. (1998): Zusammenhang zwischen Biologieunterricht und Genese von biologieorientiertem Interesse. In: Zeitschrift für Didaktik der Naturwissenschaften, ZfDN 4, H. 1, S. 13–27.
- Vogt, H./Wieder, B. (1999): Interessen im Vorschulalter. In: Zeitschrift für Didaktik der Naturwissenschaften, ZfDN 5, H. 2, S. 79–91.
- Vogt, H./Wieder, B./Schwaab, S. (2000): Individuelles Interesse bei Grundschulkindern als „Schutz“ gegen Interessenverfall in der Sekundarstufe I. – Entwicklung von spezifischen Interessen bei ausgewählten Grundschulkindern. In: Berichte des Institutes für Didaktik der Biologie Münster, IDB 9, S. 83–100.
- Wieder, B. (1999): Interessenentwicklung im Vor- und Grundschulalter. Eine grundlegende Studie. In: Berichte des Institutes für Didaktik der Biologie Münster, IDB 8, S. 19–28.
- Weinert, F.E. (1998): Entwicklung im Kindesalter. Weinheim: Beltz, Psychologische VerlagsUnion.
- Wild, E./Hofer, M. (2000): Elterliche Erziehung und die Veränderung motivationaler Orientierungen in der gymnasialen Oberstufe und der Berufsschule. In: Schiefele, U./Wild, K.-P. (Hrsg.): Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung. Münster: Waxmann, S. 31–52.

Anschriften der Autoren:

Annette Upmeier zu Belzen, Westfälische Wilhelms-Universität Münster, Institut für Didaktik der Biologie, Fliednerstr. 21, 48149 Münster.

Prof. Dr. Helmut Vogt, Universität Kassel, FB 19 Biologie/Chemie, Abteilung Didaktik der Biologie, Heinrich-Plett-Straße 40, 43132 Kassel.

Barbara Wieder, Westfälische Wilhelms-Universität Münster, Institut für Didaktik der Biologie, Fliednerstr. 21, 48149 Münster.

Franka Christen, Westfälische Wilhelms-Universität Münster, Institut für Didaktik der Biologie, Fliednerstr. 21, 48149 Münster.

Falko Rheinberg/Mirko Wendland

Veränderung der Lernmotivation in Mathematik: Eine Komponentenanalyse auf der Sekundarstufe I¹

1. Zielsetzung und Einordnung des Projektes

Es scheint keineswegs nur eine pessimistische Feststellung desillusionierter Lehrer zu sein, dass die Lernmotivation von Schülern im Verlauf ihrer Schulzeit häufig abnimmt. Längsschnittstudien zur Motivations- und Interessenentwicklung zeigen nämlich, dass ansatzweise bereits in der Grundschulzeit (z.B. Helmke 1993/1997), aber stärker noch auf der Sekundarstufe, ein Abfall in Indikatoren der Lernmotivation auftritt (zusammenf. Baumert/Köller 1998). Dies gilt natürlich nicht für jeden Einzelfall, sondern für die Mittelwertsbetrachtung auf Klassenstufenebene bei der trendabweichende Einzelfälle eingeebnet werden.

Abgesehen von trendabweichenden Einzelfällen sind aber auch einige systematische Differenzierungen zu beachten. So scheint der Abfall in der Lernmotivation auf der Klassenstufe sieben besonders stark zu sein (Fend 1997). In den gut untersuchten mathematisch-naturwissenschaftlichen Fächern („*science*“) sind davon besonders die „härteren“ Disziplinen, also Mathematik, Physik, Chemie betroffen (Krapp 1998). Dabei gibt es hier meist noch geschlechtsspezifische Effekte: Bei Mädchen zeigen sich diese ungünstigen Entwicklungen stärker und früher als bei Jungen (Hoffmann/Lehrke 1986; zusammenf. Gardener 1997).

Inbesondere vor dem Hintergrund des mäßigen Abschneidens deutscher Sekundarschüler in internationalen Vergleichsstudien (TIMSS, PISA) wäre es gut, Maßnahmen zu finden, die diesem Absinken der Lernmotivation entgegenwirken. Die Frage ist nur, was hierzu im Einzelnen zu tun ist. Je nachdem, was man unter Lernmotivation versteht und wodurch man sie beeinflusst sieht, wird man zu ganz unterschiedlichen Empfehlungen kommen. Das überrascht insofern nicht, als am Zustandekommen von Lernmotivation ganz verschiedene Komponenten beteiligt sind (s. unten). Von daher kann unzureichende Lernmotivation auch auf qualitativ unterschiedliche Faktoren zurückgehen, die entsprechend unterschiedliche Maßnahmen erfordern (vgl. beispielsweise die Entwicklung eines motivbezogenen Elterntrainings von Wild/Remy in diesem Band). So können Lernbemühungen unterbleiben, weil ein Schüler den Unterricht als viel zu schwierig oder, im Gegenteil, als langweilig leicht erlebt. Im ersten Fall erscheinen Lernbemühungen zwecklos, im zweiten überflüssig. Passen dagegen die Anforderungen, so kann es trotzdem zu unzureichender Lernmotivation kommen, weil der mögliche Lern-

1 Die berichtete Studie wurde im Rahmen des DFG-Schwerpunktprogramms »Bildungsqualität von Schule« (Az. RH 14-8/1) gefördert.

zuwachs keinen hinreichenden Anreiz besitzt – weder affektiv (z.B. unmittelbare Kompetenzfreude) noch instrumentell (z.B. erlebte Sinnhaftigkeit oder Karrierenutzen des Lernzuwachses). Zu alledem können noch Selbstregulationsprobleme treten: Man hat sich zwar fest vorgenommen zu lernen, aber schafft es nicht, diesen Vorsatz zu realisieren, weil ständig viel schönere Freizeitaktivitäten konkurrieren (alltagssprachlich: „Willensschwäche“). Je nach vorliegendem Fall wären natürlich ganz verschiedene Maßnahmen zur Motivationsförderung zu ergreifen.

Solche falladaptierten Motivationsförderungen sind das mittelfristige Ziel des jetzigen Forschungsprojektes. Um dieses Ziel zu erreichen, braucht man aber (a) zunächst ein Verfahren, mit dem sich die relevanten Motivationskomponenten erfassen lassen. Dieses Verfahren war also in einer ersten Projektphase zu entwickeln und wurde aus einem bewährten Theoriesystem abgeleitet. (b) Um eine grobe Vorausbeurteilung klassenstufenspezifischer Motivationsprobleme leisten zu können, sind dann mit diesem Verfahren die Veränderungen der relevanten Motivationskomponenten im Längsschnitt zu untersuchen und zwar auf verschiedenen Klassenstufen. Gibt es solche klassenstufenspezifischen Phänomene, so kann man Lehrern auch ohne Kenntnis des Einzelfalls sagen, worauf sie auf einer jeweiligen Klassenstufe besonders achten sollten und welche Fördermaßnahmen sie auf dieser Stufe bereithalten sollten. (c) Schließlich ist es gerade mit Blick auf gewünschte Leistungsverbesserung wichtig zu wissen, welche der verschiedenen Motivationskomponenten den stärksten Einfluss auf nachfolgende Leistungsentwicklungen haben. (d) Da wir Grund zu der Annahme hatten, dass die Motivationsentwicklung nicht lediglich einer Eigendynamik folgt, sondern entscheidend von Kontextfaktoren mitbestimmt wird, sollte schon in der ersten Projektphase der Einfluss zweier (potenziell) motivationsrelevanter Kontextfaktoren grob abgeklärt werden. Bei diesen Faktoren handelt es sich um die Bezugsnormorientierung des Lehrers und um das mathematisch-naturwissenschaftliche Anregungsklima im Elternhaus (s. unten).

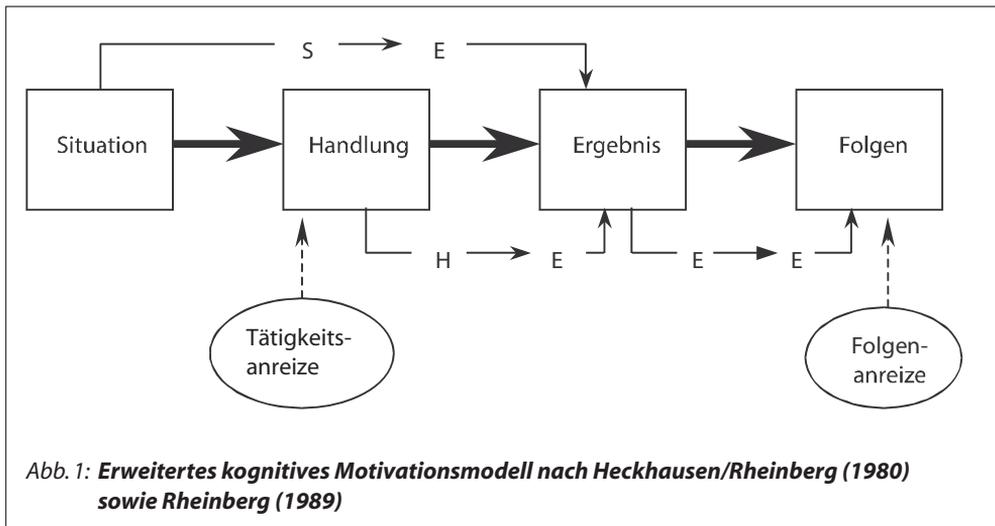
Ordnet man das Projekt in das Rahmenmodell des BIQUA-Programms ein (vgl. Abb. 4 in Doll/Prenzel in diesem Band), so geht es eindeutig um *individuelle Lernvoraussetzungen* motivational-affektiver Art und wie sie sich in *individuellen Verarbeitungsprozessen* niederschlagen (Anstrengung, Emotionen, Aufmerksamkeit etc.). Als Besonderheit dieses Projektes werden diese motivationalen Lernvoraussetzungen theoriegeleitet dekomponiert und komponentenspezifisch erfasst, um sie dann im zeitlichen Längsschnitt zu untersuchen. Dabei wird zusätzlich der mögliche Einfluss von *Lehrermerkmalen* (Bezugsnormorientierung) und Merkmalen des *Systems Eltern/Familie* (Anregungsklima/Interesse) berücksichtigt.

2. Das zugrundegelegte Theoriesystem

Unter *Lernmotivation* verstehen wir die Bereitschaft der Person bestimmte Aktivitäten vornehmlich deshalb auszuführen, weil sie sich davon einen Lernzuwachs verspricht (Rheinberg 1986). Es geht also um *beabsichtigtes Lernen*. Zur theoretischen Rekonstruktion der so verstandenen Lernmotivation stützen wir uns zunächst auf das „Erweiterte

kognitive Motivationsmodell⁴, das von Heckhausen/Rheinberg (1980) auf die Lernmotivation angewandt wurde. Dieses Modell wird weiter spezifiziert und inhaltlich angereichert sowie durch modellexterne Selbstregulationskomponenten ergänzt.

Das Modell rekonstruiert (Lern-)Motivation zunächst strikt zweckrational: Erste Bedingung hinreichender Lernmotivation ist, dass sich ein gewünschtes Lernergebnis nicht schon von alleine aus der Situation ergibt (hohe *Situations-Ergebnis-Erwartung*: „Ich brauche hier nichts zu tun, das kann ich auch so.“). Sehr hohe Situations-Ergebnis-Erwartung wirkt sich motivational also ungünstig aus. Als Zweites muss ich glauben, durch eigene Lernaktivität einen gewünschten Lernzuwachs auch herbeiführen zu können (hohe *Handlungs-Ergebnis-Erwartung*: „Wenn ich mich anstrenge, dann schaffe ich das auch.“) Als Drittes muss dieser Lernzuwachs aber auch attraktive/wichtige *Folgen* haben (Folgenanreize: z.B. Kompetenzfreude, Erkenntnisgewinn, soziale Anerkennung, Freude der Eltern, gute Noten, Versetzung, gewünschter Beruf oder Studienplatz etc.). Schließlich müssen als Viertes diese wichtigen Folgen auch hinreichend eng mit dem Lernzuwachs verknüpft sein (hohe *Ergebnis-Folge-Erwartung*: „Wenn ich das jetzt gelernt habe, bin ich sicher, dass Folge xy auch eintritt“). Bei der Motivationsvorhersage für Einzelfälle werden diese vier Komponenten als notwendige Bedingungen verknüpft (Heckhausen/Rheinberg 1980; Rheinberg 1989). Im jetzigen Kontext der Motivationsdiagnostik benutzen wir diese vier Modellkomponenten zunächst aber lediglich dazu, relevante Subskalen zur Lernmotivation zu erzeugen.



Die empirische Überprüfung dieses Modells zeigte, dass man die rein zweckzentrierte Struktur noch ergänzen musste um Anreize, die im Tätigkeitsvollzug selber liegen (Tätigkeitsanreize; Rheinberg 1989/2002). Gemeint ist, wie attraktiv/aversiv der Vollzug der (Lern-)Aktivität unabhängig vom Anreiz ihrer Ergebnisfolgen ist (z.B. Attraktivität von: „Diskutieren eines Textes mit Freunden“ vs. „Auswendiglernen von Fakten“). Dieser Tätigkeitsanreiz wird mitunter als „intrinsisch“ bezeichnet (z.B. Schiefele/Köller 2001).

Diese Bezeichnung wird aber auch anders verwandt und ist deshalb nicht ganz unproblematisch (Rheinberg 2002).

Dieser zusätzliche Anreiz ist in Abb. 1 als „Tätigkeitsanreiz“ bereits enthalten. In unserem Messinstrument wird dieser Anreiz einmal generell für die Beschäftigung mit mathematischen Gegenständen (*genereller Tätigkeitsanreiz*; z.B. „Mich mit mathematischen Aufgaben zu beschäftigen macht mir großen Spaß.“) und einmal flowspezifisch erfasst (*flowspezifischer Tätigkeitsanreiz*; z.B. „Wenn ich mich mit Mathehausaufgaben beschäftige, vergeht die Zeit wie im Fluge.“). Mit letzterem wird bestimmt, in welchem Ausmaß die Beschäftigung mit mathematischen Problemstellungen den Schüler in den Zustand des freudigen Aufgehens in einer glatt laufenden Tätigkeit bringt, bei dem die Konzentration wie von selber kommt (Csikszentmihalyi 1999; Rheinberg 2002).

Die gesonderte Berücksichtigung von Tätigkeitsanreizen ermöglicht es auch, neuere interessentheoretische Konzepte zu verankern. Interesse wird als besondere Beziehung einer Person zu einem Gegenstandsbereich aufgefasst (Krapp 2001; Prenzel 1988; Schiefele 1996). Diese Beziehung ist zum einen durch positive Vollzuserlebnisse während der Interessenhandlung, zum anderen durch eine hohe Wertschätzung für den Gegenstand und das durch Lernen erreichte tiefere Verständnis charakterisiert („epistemologische Orientierung“ interessierten Lernens; Prenzel 1988). Während Ersteres affektive Tätigkeitsanreize beschreibt, betrifft Letzteres bestimmte Folgen interessengeleiteter Lernaktivitäten, nämlich das tiefere Durchdringen und Verstehen des Interessengegenstandes. Die Wertschätzung für den Lerngegenstand wird in unserem Messinstrument durch eine eigene Skala erfasst (*Sachinteresse*; z.B.: „In meiner Freizeit beschäftige ich mich auch unabhängig vom Unterricht mit Dingen, die mit Mathe zu tun haben.“).

Um neben positiven Anreizen auch negative zu berücksichtigen, wurde eine Kurzskaala aufgenommen, die angstbezogenes Erleben während des Lernens erfasst (*Angst*; z.B.: „Im Matheunterricht habe ich Angst davor, aufgerufen zu werden.“; vgl. Pekrun in diesem Band). Weiterhin wurde der Tatsache Rechnung getragen, dass Lernaktivitäten trotz eigentlich günstiger Motivationsvoraussetzungen aufgrund von Selbstregulationsproblemen unterbleiben können. Hierzu gibt es hoch komplexe Funktionsmodelle, die wir hier nicht darstellen können (für einen Überblick s. Boekaert/Pintrich/Zeidner 2000). Für unsere Zwecke haben wir in Anlehnung an Kuhl und Fuhrmann (1999) für unseren Kontext relevante Skalen entwickelt. Bezogen auf das Mathematiklernen von Sekundarschülern ließ sich das breite Spektrum unterschiedlichster Selbstregulationskompetenzen drei Dimensionen zuordnen. (1) „Betrifft mich der Misserfolg überhaupt und schaffe ich es, mich gerade dann voll auf das Lernen zu konzentrieren?“ (*Betroffenheit und Anstrengungssteigerung nach Misserfolg*; z.B.: „Eine schlechte Mathenote würde mich dazu bringen, mich sofort voll und ganz auf Mathe zu konzentrieren.“) (2) „Wie gut habe ich mich beim Mathematiklernen selbst im Griff und widerstehe verlockenden Alternativen?“ (*Selbststeuerungsprobleme und geringe Ausdauer*; z.B.: „Wenn ich zuhause an Matheaufgaben sitze, schaffe ich es meist nicht lang, daran zu bleiben.“) (3) „Wie gut kann ich meine eigenen Emotionen beim Mathematiklernen beeinflussen?“ (*Emotionskontrolle*; z.B.: „Auch, wenn es mir mal nicht so gut geht, schaffe ich es meist irgendwie, mich für Mathe wieder in eine gute Stimmung zu versetzen.“).

Das unserem Projekt zugrunde liegende theoretische Gesamtsystem enthält ein breites Spektrum solcher Motivations- und Selbstregulationsfaktoren, die nach dem gegenwärtigen Kenntnisstand direkt oder indirekt leistungsrelevant sein müssten. Von daher wäre es gut, diese Faktoren erfassen und in ihrer Veränderung über die Sekundarschulzeit hinweg untersuchen zu können. Man hätte damit eine solide Grundlage für die Entwicklung und den Einsatz fall- und problemspezifischer Interventionen zur Motivationsförderung.

Um schon im ersten Zugriff Hinweise auf Bedingungen unterschiedlicher Motivationsentwicklungen zu erhalten, werden im jetzigen Projekt zwei Kontextfaktoren mit berücksichtigt. Auf der Lehrerseite wird erfasst, ob der Lehrer neben den formal vorgegebenen Leistungsbeurteilungen seine Schüler auch im intraindividuellen Längsschnittvergleich ihrer Leistungsentwicklung wahrnimmt (*individuelle Bezugsnorm-Orientierung*). Letzteres hat sich als motivational förderlich erwiesen (Rheinberg 1980/2001). Dafür liegen inzwischen repräsentative Daten aus der deutschen Sekundarschule vor (Lüdtke/Köller, im Druck).

Beim zweiten Kontextfaktor, nämlich dem mathematisch-naturwissenschaftlichen Anregungsklima im Elternhaus lassen wir uns von der sozialisationstheoretischen Annahme leiten, dass elterliche Wertschätzungen für einen bestimmten Gegenstandsbereich direkt oder indirekt die Wertschätzung der Schüler beeinflussen sollten. So etwas wäre z.B. für den Motivationsfaktor Sachinteresse gut vorstellbar. Wenngleich schon Trudewind (1975) den Einfluss häuslicher Anregungsbedingungen für die Entwicklung der Leistungsmotivation nachgewiesen hat, sind uns analoge Untersuchungen zur Interessenentwicklung in Mathematik nicht bekannt. Von daher hat das Projekt beim Elterneinfluss noch Erkundungscharakter.

Faktorenanalytisch ließen sich bei diesem Kontextfaktor zwei Dimensionen ausmachen. Zum einen unterscheiden sich Elternhäuser in der inhaltsbezogenen Wertschätzung für mathematisch-naturwissenschaftliche Phänomene und Fragen (*Außerschulische Wertschätzung*). Zum anderen unterscheiden sie sich darin, für wie wichtig sie die Fächer Mathematik/Physik für das schulische und berufliche Fortkommen ihres Kindes halten (*Familiäre Nutzenüberzeugung*). Sowohl der lehrerseitige wie auch der elternteilige Kontextfaktor werden aus der Sicht der Schüler wie auch der Agenten (Lehrer, Eltern) erfasst.

3. Das Potsdamer Motivations Inventar (PMI)

Über mehrere Zwischenschritte wurden zu diesen theoretisch verankerten Motivationskomponenten jeweils Fragebogenskalen entwickelt. So wurden in gesonderten Voruntersuchungen mit freier Erhebungsmethode die Folgen erfasst und systematisiert, die deutschen Sekundarschülern im Fach Mathematik wichtig sind (Rheinberg/Wendland 2001).

Die Endfassung des Potsdamer Motivationsinventars (PMI) enthält 14 Skalen (79 Items) zur Erfassung der verschiedenen Komponenten der Lernmotivation. Zwei

weitere Skalen wurden zur Erfassung des Kontextfaktors Anregungsklima im Elternhaus entwickelt. (Aus Raumgründen können wir hier darauf nicht genauer eingehen.) Die Beantwortung des Fragebogens dauert bei klassenweiser Anwendung etwa 10 bis 20 Minuten. Er liegt in einer Version für Mathematik (PMI-M) und Physik (PMI-P) vor. Tabelle 1 informiert im Überblick über die Skalen und relevante Kennwerte des PMI-M. Den Kennwerten liegen N = 747 Sekundarschüler der fünften bis neunten Klassenstufe zugrunde.

Tab. 1: Mittelwerte, Standardabweichungen und Konsistenzen des PMI-M sowie Korrelationen zur späteren Mathematiknote

PMI-Skalen (N = 747 Sekundarschüler der Klassenstufen 5 bis 9)	M _{T1} *	s _{T1}	Cron- bachs α	Stabi- lität r _{tt}	Korrelation der Skalenwerte (T1) mit der rekodierten Halbjahresnote (T2) (beruhend auf pro Klasse z- standardisierten Werten)**					
					5-9	5	6	7	8	9
Handlungs-Ergebnis-Erwartung	3,82	,65	.72	.53	.35	.33	.39	.34	.31	.37
Situations-Ergebnis-Erwartung	2,46	1,03	.86	.67	.41	.37	.41	.23	.44	.53
Ergebnis-Folgen-Erwartung	4,07	,76	.60	.39	.14	-	.18	-	.21	-
Folgenanreiz Gute Noten	4,24	,63	.73	.44	.12	-	-	-	-	.21
Folgenanreiz Sach- und Selbst- bewertungsfolgen	3,55	,90	.84	.56	.30	.20	.21	.36	.37	.37
Folgenanreiz Versetzung	3,60	1,12	.40	.30	-	-	-	-	-	-.24
Folgenanreiz Fremdbewertungs- folgen	3,01	,95	.73	.50	-	.24	-	-	-	-
Generelle Tätigkeitsanreize	3,30	,77	.83	.54	.31	.24	.22	-	.41	.45
Flowspezifische Tätigkeitsanreize	2,80	,88	.82	.54	.22	-	-	-	.34	.35
Angst (rekodiert)	3,73	,98	.79	.53	.36	.40	.46	-	.33	.36
Sachinteresse	2,86	,85	.81	.58	.27	-	-	.25	.37	.48
Betroffenheit, Anstrengungsstei- gerung nach Misserfolg	3,56	,94	.75	.50	-	-	-	-	-	-
Selbststeuerungsprobleme und geringe Ausdauer	2,70	,82	.72	.53	-.33	-.25	-.37	-.22	-.28	-.48
Emotionskontrolle	2,77	,92	.73	.44	.13	-	-	-	.24	.29
Außerschulische Wertschätzung	3,22	,80	.68	.55	-	-	-	-	-	-
Familiäre Nutzenüberzeugung	3,43	,67	.54	.43	-	-	-	-	-	-

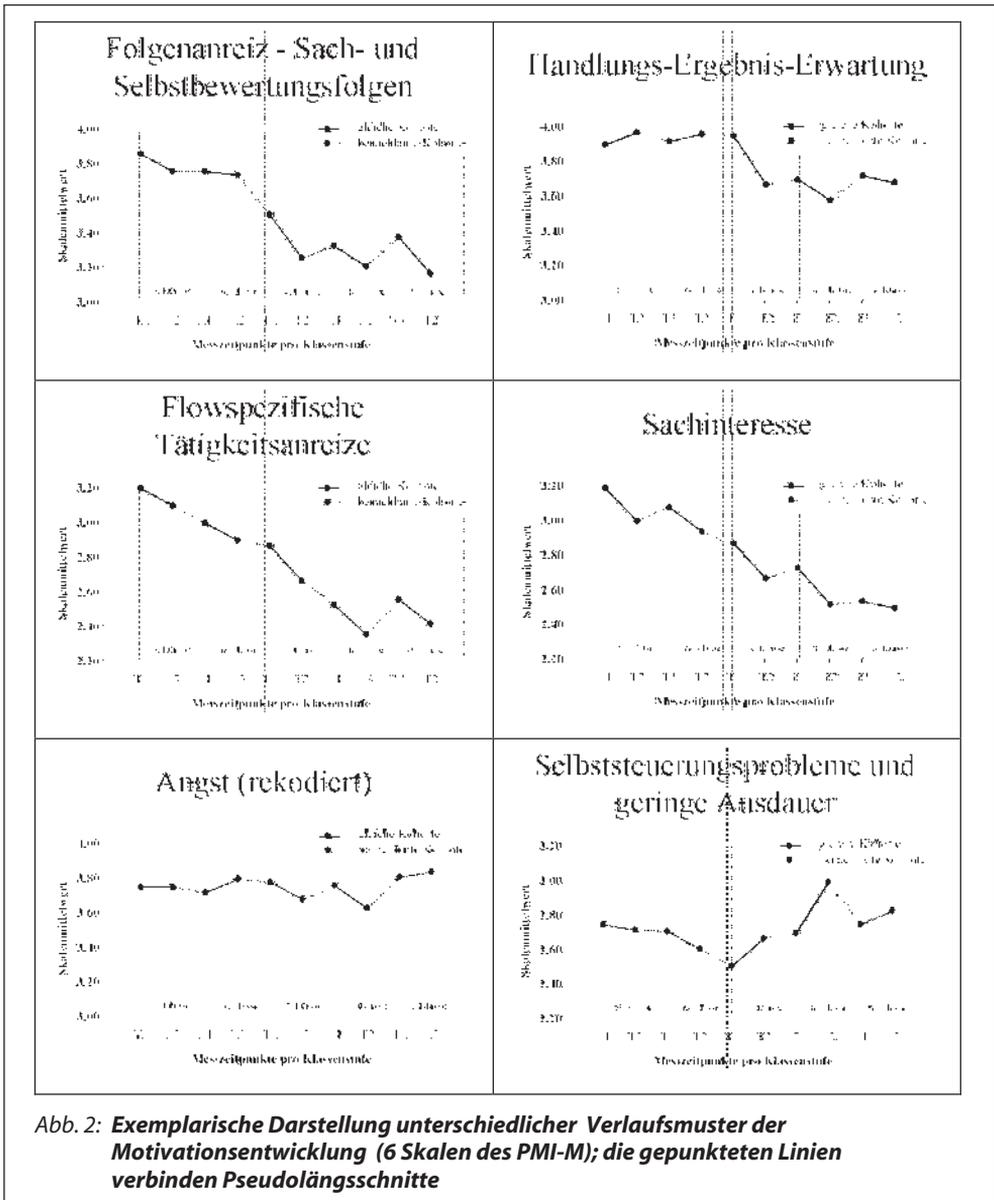
* Jede Skala hat ein Minimum von 1,00 und ein Maximum von 5,00; ** Korrelationen mit $p < .05$

Bis auf die Skala „Folgenanreiz Versetzung“ (nur zwei Items) sind die Konsistenzen akzeptabel bis gut. Interessant ist ein Blick auf die Korrelationen zur späteren Note. Diese Korrelationen drücken aus, in welchem Ausmaß eine jetzt erfasste Motivationskomponente die Mathematik-(Zeugnis-)note sechs Monate später vorhersagt. Auch ohne das vorschnell kausal interpretieren zu wollen (s. unten), sind die Korrelationen teils recht beachtlich, wobei sich deutliche Unterschiede zwischen den Motivationskomponenten zeigen. Über alle Klassenstufen hinweg lässt sich besonders über die jetzige Situations-Ergebnis-Erwartung ($r = .41$), die Angst ($r = .36$), die Handlungs-Ergebnis-Erwartung ($r = .35$), Selbststeuerungsprobleme (und geringe Ausdauer) ($r = -.33$), den Generellen Tätigkeitsanreiz ($r = .31$), die Sach- und Selbstbewertungsfolgen ($r = .30$) sowie das Sachinteresse ($r = .27$) vorhersagen, welche Zeugnisnote der Schüler ein halbes Jahr später erreichen wird. (Genauere Aufschlüsse geben hierzu die Regressionsanalysen in Tabelle 2.) Getrennt nach Klassenstufen analysiert, ist bei einigen PMI-Komponenten der Noteneinfluss auf allen Klassenstufen etwa gleichhoch (z.B. bei Handlungs-Ergebnis-Erwartung), während bei anderen PMI-Komponenten der Noteneinfluss über Klassenstufen hinweg zunimmt (Tätigkeitsanreize, Sachinteresse, Sach- und Selbstbewertungsfolgen). Die Fremdbewertungsfolgen haben nur auf der fünften Klassenstufe eine gewisse Notenrelevanz.

4. Veränderung der Motivationskomponenten

Der PMI wurde zu Schuljahresbeginn und zur Schuljahresmitte bei $N = 615$ Schülerinnen und Schülern der Klassenstufen 5 bis 9 im Großraum Potsdam eingesetzt. (Eine dritte Erhebung läuft zurzeit noch, bei der auch die Kontextfaktoren genauer erfasst werden.) Es handelt sich also um eine kohortengestaffelte Längsschnittstudie, bei der Pseudolängsschnitteffekte (Vergleich zwischen verschiedenen Klassenstufen) mit echten Längsschnitteffekten (Vergleich innerhalb derselben Klassenstufe) kombiniert werden können. Aus Raumgründen können wir hier nur auszugsweise berichten (eine vollständige Darstellung findet sich bei Rheinberg/Wendland 2001). Abbildung 2 zeigt einige typische Verläufe. (Die gepunkteten Linien verbinden Pseudolängsschnittdaten.)

Wie man sieht, ergeben sich je nach Motivationskomponente verschiedene Verlaufsmuster. Eine Art Stufenmodell zeigt sich bei den Sach- und Selbstbewertungsfolgen sowie der Handlungs-Ergebnis-Erwartung. Der entscheidende Abfall tritt zwischen Mitte der sechsten und Mitte der siebten Klassenstufe auf. Das deckt sich mit den eingangs erwähnten Befunden zur Motivationsentwicklung bei Fend (1997). Anders liegen die Dinge bei den Skalen für Flowspezifische Tätigkeitsanreize und Sachinteresse. Statt eines Stufenmodells findet sich hier ein eher kontinuierlicher Abfall. „Angst“ bleibt schließlich etwa unverändert, wohingegen „Selbststeuerungsprobleme und geringe Ausdauer“ auf der achten Klassenstufe stark zunehmen (echte Längsschnittdaten). Schon diese verschiedenen Veränderungsmuster zeigen, dass es erforderlich ist, Lernmotivati-on differenziert in ihren einzelnen Komponenten zu analysieren.



5. Leistungsvorhersagen durch Motivationskomponenten

Tabelle 1 zeigt, dass mehrere PMI-M Faktoren die spätere Mathematiknote vorhersagen. Da diese Faktoren aber nicht gänzlich unabhängig von einander sind, interessiert, wie viel Leistungsvarianz durch vorherige Motivation insgesamt vorhersagbar ist. Zusammengefasst über alle Klassenstufen sind das immerhin 26,3% der Leistungsvarianz, die der PMI-M vorhersagt (Tabelle 2, S. 316)!

Tab. 2: Schrittweise Regressionsanalyse mit den PMI-M-Variablen zum Schuljahresbeginn als Prädiktor zur Vorhersage nachfolgender Schulleistung (Halbjahreszeugnisnote) (N = 615)

PMI-M-Komponente	Beta	r	Varianz- aufklärung	Signifikanz
Situations-Ergebnis-Erwartung	-.312	-.41	16,4%	.000
Handlungs-Ergebnis-Erwartung	-.132	-.35	4,7%	.002
Selbststeuerungsprobleme und geringe Ausdauer	.124	.33	1,8%	.004
Ergebnis-Folgen-Erwartung	-.086	-.14	0,8%	.024
Folgenanreize – Fremdbewertungsfolgen	.103	.08	0,8%	.007
Folgenanreize – Sach- und Selbstbewertungsfolgen	-.158	-.29	0,7%	.001
Emotionskontrolle	.128	-.13	1,1%	.004
			∑ 26,3%	

In den Regressionsgleichungen tauchen jedes Mal diejenigen PMI-Komponenten auf, die auch schon in Tabelle 1 signifikant mit der Leistung korrelierten. Je nach Klassenstufe wechselt mitunter das Beta-Gewicht der Komponenten. Durchgängig gehen die Handlungs- und die Situations-Ergebnis-Erwartungen ein, die Sach- und Selbstbewertungsfolgen sowie Selbststeuerungsprobleme und geringe Ausdauer (s. im Einzelnen Rheinberg/Wendland 2001).

Zeitversetzte Kreuzkorrelationen² („*cross-lagged-panel*“) zeigen, dass vorangehende Leistung nachfolgende Motivation beeinflusst, aber mit etwa gleichem Gewicht vorangehende Motivation die nachfolgende Leistung. Interessant ist deshalb, ob Motivationsfaktoren auch dann noch Leistung vorhersagen, wenn man den Einfluss der vorangegangenen Leistung herausrechnet. Dann würde derjenige Motivationseinfluss erkennbar werden, der im letzten halben Jahr sozusagen hinzugekommen ist. Trivialerweise zeigt sich, dass die künftige Leistung allein zu 40.3% durch die vorangegangene Leistung vorhergesagt wird (Tabelle 3). In letzterem sind neben kognitiven ja auch bereits frühere motivationale Einflüsse mit enthalten.

- 2 Geprüft wurden die beiden Kausalhypothesen, ob vorangehende Leistung nachfolgende Motivation und ob vorangehende Motivation nachfolgende Leistung beeinflusst. Dabei folgten wir der von Bortz (1984) beschriebenen Vorgehensweise: „Zwei Korrelationen eines jeden Merkmals mit sich selbst, gemessen zu zwei Zeitpunkten, zwei Korrelationen zwischen den zwei verschiedenen, zeitversetzt gemessenen Merkmalen und zwei Korrelationen zwischen zwei verschiedenen, gleichzeitig gemessenen Merkmalen. Die vier zuletzt genannten Korrelationen sind für die Entscheidung, welcher der beiden Kausalhypothesen der Vorzug zu geben sei, besonders wichtig.“

Tab. 3: **Schrittweise Regressionsanalyse mit den PMI-M-Variablen zum Schuljahresbeginn sowie der Zeugnisnote als Prädiktor zur Vorhersage nachfolgender Schulleistung (Halbjahreszeugnisnote) (N = 613)**

PMI-M-Komponente	Beta	r	Varianz- aufklärung	Signifikanz
Zeugnisnote (T1)	.527	.64	40,3%	.000
Situations-Ergebnis-Erwartung	-.136	-.41	2,3%	.000
Selbststeuerungsprobleme und geringe Ausdauer	.119	.33	1,5%	.001
Folgenanreize – Sach- und Selbstbewertungsfolgen	-.117	-.29	0,4%	.002
Emotionskontrolle	.090	-.13	0,6%	.016
			∑ 45,1%	

Trotzdem leisten die PMI-M Komponenten auch dann noch einen signifikanten *zusätzlichen* Beitrag von 4,8%. Auf einer Klassenstufe mit besonders niedriger Notenstabilität (siebte Klassenstufe) steigt der durch das PMI vorhergesagte Varianzanteil sogar auf 20,7% und war damit gewichtiger, als der Beitrag der vorausgegangenen Note (9,6%). Ein ähnlicher Effekt wurde kürzlich auch von Schneider (2001) gefunden.

6. Diskussion und Ausblick

Die bisherigen Befunde sprechen dafür, dass es mit dem PMI gelungen ist, ein motivationstheoretisch verankertes Instrumentarium zu entwickeln, das auf Faktoren zielt, die nachfolgende Leistungen beeinflussen. Dabei sind die Analysen hierzu noch etwas grob, weil zwecks Leistungsprognose die Faktoren des Erweiterten kognitiven Motivationsmodells eigentlich konfiguratorisch miteinander verknüpft werden müssten (Heckhausen/Rheinberg 1980; Rheinberg 1989). Mit Blick auf Veränderungen in der Lernmotivation zeigt sich, dass verschiedene Motivationskomponenten ganz offensichtlich unterschiedlichen Verlaufsmustern folgen. Weitere Analysen hatten ergeben, dass es auf jeder Klassenstufe eindeutige „Verschlechterungstypen“ gibt, die aber durch teils unterschiedliche Motivationskomponenten gebildet werden (Rheinberg/Wendland 2001). Wie zu erwarten, präsentiert sich „Lernmotivation“ also keineswegs als strukturhomogene Größe. Entsprechend eindeutige „Verbesserungstypen“ fanden sich nicht.

Als nächste Schritte erscheint es angezeigt, insbesondere für die Klassenstufen fünf bis sieben Interventions- und Trainingsmaßnahmen zu entwickeln, die dem deutlichen Abfall in den Sach- und Selbstbewertungsfolgen sowie in der Handlungs-Ergebnis-Erwartung begegnen. Hier wird man sich an Trainingskonzepten zur Förderung der *individuellen Bezugsnorm-Orientierung* ausrichten können wie sie bei Rheinberg und Krug (1999) beschrieben sind. Es wird aber darauf ankommen, diese Trainingskonzepte mög-

lichst glatt in den laufenden Unterricht einzubringen. Auf anderen Klassenstufen werden womöglich andere Dinge wichtiger, wie z.B. *Selbstregulationskompetenzen* und *Arbeitstechniken*, die ein freudvolles Aufgehen in der Tätigkeit fördern. Hier sind neue Trainingskonzepte zu entwickeln.

Um nicht alles selbst zu entwickeln, wird es hilfreich sein, sich von „besonderen Schulen/Klassen“ anregen zu lassen, in denen auch ohne förmliche Intervention häufig günstige Motivationsentwicklungen auftreten. Hier wird man genauer den laufenden Unterricht, aber auch relevante Kontextfaktoren untersuchen und prüfen, was davon auch für den üblichen Unterricht übernehmbar wäre. Mittelfristiges Ziel ist dann ein Compendium von motivationsbezogenen Fördermaßnahmen, die je nach diagnostiziertem Problemprofil im PMI platziert einsetzbar sind.

Literatur

- Baumert, J./Köller, O. (1998): Interest research in secondary level I: An overview. In: Hoffmann, L./Krapp, A./Renniger, K.A./Baumert, J. (Eds.), *Interest and learning* (pp. 241–256), Kiel: IPN.
- Boekaerts, M./Pintrich, P.R./Zeidner, M. (2000): *Handbook of self-regulation*. San Diego: Academic Press.
- Bortz, J. (1984): *Lehrbuch der empirischen Forschung für Sozialwissenschaftler*. Berlin: Springer.
- Csikszentmihalyi, M. (1999): *Das Flow-Erlebnis. Jenseits von Angst und Langeweile: im Tun aufgehen*. Stuttgart: Klett-Cotta.
- Doll, J./Prenzel, M. (2002): Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. In: *Zeitschrift für Pädagogik* (in diesem Band).
- Fend, H. (1997): *Der Umgang mit Schule in der Adoleszenz*. Bern: Huber.
- Gardner, P.L. (1997): The development of males and females interests in science and technology. In: Hoffmann, L./Krapp, A./Renniger, K.A./Baumert, J. (Eds.), *Interest and learning* (pp. 41–57), Kiel: IPN.
- Heckhausen, H./Rheinberg, F. (1980): Lernmotivation im Unterricht, erneut betrachtet, *Unterrichtswissenschaft*, 8, S. 7–47.
- Helmke, A. (1993): Die Entwicklung der Lernfreude vom Kindergarten bis zur 5. Klassenstufe, *Zeitschrift für Pädagogische Psychologie*, 7, S. 77–86.
- Helmke, A. (1997): Entwicklung lern- und leistungsbezogener Motive und Einstellungen: Ergebnisse aus dem SCHOLASTIK-Projekt. In: Weinert, F.E./Helmke, A. (Hrsg.), *Entwicklung im Grundschulalter*. Weinheim: PVU.
- Hoffmann, L./Lehrke, M. (1986): Eine Untersuchung über Schülerinteressen an Physik und Technik, *Zeitschrift für Pädagogik*, 32, S. 189–204.
- Krapp, A. (1998): Entwicklung und Förderung von Interesse im Unterricht, *Psychologie in Erziehung und Unterricht*, 45, S. 185–201.
- Krapp, A. (2001): Interesse. In Rost, D.H. (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz, PVU, S. 286–293.
- Kuhl, J./Fuhrmann, A. (1999): *Selbststeuerungs-Inventar: SSI-K (Kurzversion)*. Institut für Psychologie, Universität Osnabrück.
- Lüdtko, O./Köller, O. (2002): Individuelle Bezugsnormorientierung und soziale Vergleiche im Mathematikunterricht: Der Einfluss unterschiedlicher Referenzrahmen auf das fachspezifische Selbstkonzept der Begabung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 34, S. 156–166.

- Prenzel, M. (1988): Die Wirkungsweise von Interesse. Opladen: Westdeutscher Verlag.
- Rheinberg, F. (1980): Leistungsbewertung und Lernmotivation. Göttingen: Hogrefe.
- Rheinberg, F. (1986): Lernmotivation. In: Sarges, W./Fricke, R. (Hrsg.), Psychologie für die Erwachsenenbildung. Göttingen: Hogrefe, S. 360–365.
- Rheinberg, F. (1989): Zweck und Tätigkeit. Göttingen: Hogrefe.
- Rheinberg, F. (2001): Bezugsnormen und schulische Leistungsbeurteilung. In: Weinert, F.E. (Hrsg.), Leistungsmessung in Schulen. Weinheim: Beltz, S. 59–71.
- Rheinberg, F. (2002): Motivation. Stuttgart: Kohlhammer.
- Rheinberg, F./Krug, S. (2002): Motivationsförderung im Schulalltag. Göttingen: Hogrefe.
- Rheinberg, F./Wendland, M. (2001): DFG-Bericht zum Projekt „Förderung von Motivationskomponenten“, Universität Potsdam
- Schiefele, U. (1996): Motivation und Lernen mit Texten. Göttingen: Hogrefe.
- Schiefele, U./Köller, O. (2001): Intrinsische und extrinsische Motivation. In: Rost, D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie, S. 304–310.
- Schneider, W. (2001): Kognitive Fähigkeit, Motivation und Schulleistung in unterschiedlichen Altersbereichen. Vortrag auf dem Symposium „Schule als differenzielle Entwicklungsmilieus“, MPI-Berlin am 16./17. November 2001.
- Trudewind, C. (1975): Häusliche Umwelt und Motiventwicklung. Göttingen: Hogrefe.
- Wild, E./Remy, E. (2002). Quantität und Qualität der elterlichen Hausaufgabenbetreuung von Drittklässlern in Mathematik. In: Zeitschrift für Pädagogik (in diesem Band).

Anschrift der Autoren:

Prof. Dr. Falko Rheinberg, Universität Potsdam, Institut für Psychologie, Allgemeine Psychologie II, Postfach 601553, 14415 Potsdam.

Mirko Wendland, Universität Potsdam, Institut für Psychologie, Allgemeine Psychologie II, Postfach 601553, 14415 Potsdam.

Teil V: Einstellungen und Werte als förderliche oder hinderliche Bedingungen schulischer Leistungsfähigkeit

Mathematisch-naturwissenschaftliche Fächer als
Einstellungsobjekte: Einflüsse von Makro- und Mesoebene
auf die Einstellungsbildung

Bettina Hannover

Einleitung

Die Gemeinsamkeit der in diesem Abschnitt zusammengefassten Forschungsvorhaben besteht darin, dass sie sozial geteiltes Wissen, das schulische Lern- und Leistungskulturen konstituiert, in den Mittelpunkt stellen. Im Unterschied zu vielen anderen Projekten des DFG-Schwerpunktprogramms werden in den im Folgenden näher vorgestellten drei Vorhaben individuelle Indikatoren der schulischen Lern- und Leistungsentwicklung – wie z.B. persönliches Leistungsniveau, Selbstkonzept eigener Fähigkeiten, Einstellungen und Motivation – nicht als unabhängige, sondern nur als abhängige Variable in den Blick genommen: Sie können nach Meinung der Autoren nur dann richtig verstanden werden, wenn das Individuum nicht als eine vom sozialen Kontext isolierte psychologische Einheit, sondern als Mitglied oder integraler Bestandteil eines kulturellen Systems betrachtet wird.

Ein kulturelles System umfasst Normen, Ideale und Werte, die vermittelt über gesellschaftliche Sozialisationspraktiken und Institutionen das psychologische System des Individuums formen (z.B. Triandis 2001). Kulturen unterscheiden sich somit nicht nur zwischen Ländern oder Regionen, sondern auch innerhalb einer Gesellschaft können für Gruppen von Personen verschiedene Kulturen existieren (z.B. Geschlechterkulturen, Alterskulturen, „Subkulturen“). Normen, Ideale und Werte reflektieren sich beispielsweise in fundamentalen Schriften einer Kultur, in Erzählungen, Sprichwörtern, Symbolen, oder spezifischer in Gesetzestexten, Werbung, Anstandsregeln oder schulischen Curricula. Sie signalisieren dem Individuum, was es bedeutet, ein akzeptiertes, geschätztes und angemessenes Mitglied der jeweiligen Kultur zu sein und was das Individuum tun soll, um kulturspezifische Normen zu erfüllen oder kulturspezifische Ideale zu erreichen (z.B. Markus/Mullally/Kitayama 1997). Die moderne Sozialpsychologie geht davon aus, dass sich solche impliziten Handlungsimperative in der Art und Weise, wie das Individuum seine Identität konstruiert, niederschlagen, und dass die kulturell geprägte Identität ihrerseits das Denken, Fühlen und Handeln der Person steuert (Markus/Kitayama 1991; für einen Überblick siehe z.B. Oyserman/Coon/Kemmelmeier 2002). Dies gilt unabhängig davon, ob der oder die Einzelne sich über diese Zusammenhänge Rechenschaft ablegt (z.B. Hannover/Kühnen 2002).

Die drei in diesem Abschnitt zusammengestellten Aufsätze untersuchen Aspekte von Kultur, die die individuelle Interessen- und Leistungsentwicklung von Schülerinnen und Schülern prägen. Damit nehmen die Autoren Bezug auf das im Rahmenantrag zum DFG-Schwerpunktprogramm *„Die Bildungsqualität von Schule: Fachliches und fächerübergreifendes Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten“* von Prenzel, Merckens und Noack

(1999) formulierte Ziel, komplexe Bedingungsbeziehungen schulischer Bildung zu untersuchen, in die außerunterrichtliche und außerschulische Kontextfaktoren einbezogen sind. Als Beispiele unterrichtsübergreifender Kontextfaktoren nennen Prenzel u.a. (1999) das Klima oder Profil einer Schule und für schulübergreifend wirksam werdende Kontextfaktoren das Elternhaus oder die Gleichaltrigengruppe der Schülerinnen und Schüler (S. 3).

Die drei in diesem Abschnitt vorgestellten Forschungsvorhaben nehmen unterrichts- und schulübergreifend wirksame Einflussfaktoren auf die individuelle Lern- und Leistungsentwicklung in den Blick, und zwar spezifisch solche, die sich in Form kulturell geteilter Werte, Normen und impliziter Handlungsimperative zeigen. Mit diesen Überlegungen greifen die drei Projekte ein zentrales Anliegen der Internationalen Schulvergleichsstudien auf, für die exemplarisch TIMSS in den Mittelpunkt des DFG-Schwerpunktprogramms gestellt worden ist. In der Schule werden die Weichen für die Bildungsbiographie jedes Einzelnen gestellt. Im Zeitalter der Globalisierung sind Schulen zu zentralen Schaltstellen geworden, die darüber entscheiden, ob alle kulturellen Gruppen einer Gesellschaft – auch diejenigen mit schlechteren Startbedingungen – an den Errungenschaften der Wissens- und Informationsgesellschaft teilhaben können. Entsprechend beabsichtigen die Internationalen Schulvergleichsstudien ausdrücklich nicht, „den Horizont moderner Allgemeinbildung zu vermessen, oder auch nur die Umrisse eines internationalen Kerncurriculums nachzuzeichnen. Es ist gerade (ihre; Einfügung B. Hannover) ... Stärke ..., sich ... mit der Lesekompetenz und mathematischen Modellierungsfähigkeit auf Basiskompetenzen zu konzentrieren, die ... wichtige Voraussetzungen für die ... Generalisierung universeller Prämissen für die Teilhabe an Kommunikation und damit auch für Lernfähigkeit darstellen“ (Baumert/Stanat/Demmrich 2001, S. 21). Die in den Schulvergleichsstudien erfassten muttersprachlichen, mathematischen und naturwissenschaftlichen Kompetenzen können somit als basale Kulturwerkzeuge verstanden werden, in dem Sinne, als sie Voraussetzung zur verantwortungsvollen und verständigen Teilnahme am gesellschaftlichen Leben sind (Baumert 2002).

Vor diesem Hintergrund macht die Kernannahme der drei hier vorgestellten Projekte Sinn, nach der der Erwerb solcher – den „Kernbestand kultureller Literalität“ (Baumert u.a. 2001, S. 20) bildenden – Kompetenzen *selbst* von kulturellen Faktoren geprägt wird, genauer von sozial geteilten Normen, Idealen und Handlungsimperativen, die schulische Lern- und Leistungskulturen ausmachen.

Die drei Projekte fokussieren jeweils einen unterschiedlichen Aspekt von Kultur. In dem Aufsatz von Anna-Katharina Pelkner, Ralph Günther und Klaus Boehnke wird die für das Projekt „*Streber vs. Nurd – Zur Kultur- und Geschlechtsspezifität der (negativen und positiven) Sanktionierung guter mathematischer Schulleistungen bei Jugendlichen in Deutschland und Nordamerika*“ (Projektleiter Klaus Boehnke; Universität Chemnitz, jetzt International University Bremen) zentrale These untersucht, dass die Güte mathematischer Schulleistungen mit der leistungsbezogenen Peer-Kultur zusammenhängt, wobei die Peer-Kultur in Abhängigkeit vom Geschlecht des Schülers und dem Land, in dem er oder sie aufwächst, unterschiedlich geprägt sein kann. Genauer haben die Autoren vermutet, dass starke Anstrengung und gute Leistungen eines Schülers, insbesondere aber

einer Schülerin, bei deutschen Jugendlichen schlecht angesehen sind. Weil leistungsstarke Mathematikschüler und insbesondere -schülerinnen demnach befürchten müssen, sozial ausgegrenzt zu werden, erwarten die Verfasser, dass sie ihr Leistungspotenzial in diesem Fach nicht ausschöpfen.

Während im Projekt von Klaus Boehnke die Peer-Kultur in ihrem Einfluss auf die individuelle schulische Leistungsentwicklung untersucht wird, fokussiert das Projekt von Peter Noack (*Universität Jena*) „*Wert und Nützlichkeit von Bildung: Bewältigung schulischer Anforderungen in Abhängigkeit von Schüler- und Umfeldurteilen*“ die Kultur des Elternhauses. Genauer geht es in dem Aufsatz von Juliane Strecker und Peter Noack um Überzeugungen zur Wichtigkeit und Nützlichkeit von Bildung im Allgemeinen sowie von fachspezifischem Wissen (Deutsch und Mathematik) im Besonderen. Es wird vermutet, dass solche Überzeugungen in globalen Werthaltungen verankert sind und in Abhängigkeit von Merkmalen der Familien und Schulen variieren. Vermittelt darüber, dass Schülerinnen und Schüler die in ihrer Familien- oder Schulkultur kommunizierten Wichtigkeits- und Nützlichkeitsüberzeugungen gegenüber Bildung in ihr eigenes Wertesystem übernehmen, wirken diese sich auf ihr individuelles Lern- und Leistungsverhalten aus.

Die Untersuchung von Peerkultur und Elternhaus-Kultur wird durch das Projekt von Bettina Hannover (*Universität Dortmund, jetzt Freie Universität Berlin*) „*Der Einfluss des Image von Mathematik und Naturwissenschaften auf die schulische Interessen- und Leistungsentwicklung*“ ergänzt, in dem die Kulturen spezifischer Schulfächer im Vordergrund stehen. Genauer werden im Aufsatz von Bettina Hannover und Ursula Kessels kulturell geteilte Annahmen über mathematisch-naturwissenschaftliche Schulfächer im Unterschied zu geistes- und sozialwissenschaftlichen Disziplinen untersucht; nämlich stereotype Vorstellungen über den Gegenstandsbereich, über Unterrichtsscripts, über typische Lehrpersonen und Schüler/innen, die die jeweiligen Fächer mögen oder ablehnen. Es wird vermutet, dass Lernende solche Schulfach-Kulturen zu ihrem Selbstbild in Beziehung setzen und das Mögen bzw. Ablehnen der jeweiligen Fachdisziplin von dem Ausmaß abhängig machen, in dem die sozial geteilten Annahmen dem Bild entsprechen, das sie von der eigenen Person haben.

Im Folgenden präsentieren sich die drei Projekte mit ersten Ergebnissen, die die These veranschaulichen, dass Aspekte von Kultur das Denken, Fühlen und Handeln der an Schule beteiligten Individuen bestimmen.

Literatur

- Baumert, J. (2002): Deutschland im internationalen Bildungsvergleich. *Universitas*, 57, S. 114–135.
- Baumert, J./Stanat, P./Demmrich, A. (2001): PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In Deutsches PISA-Konsortium (Hrsg.): *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*, S. 15–68. Opladen: Leske + Budrich.
- Hannover, B./Kühnen, U. (2002): Der Einfluss unabhängiger und interdependenter Selbstkonstruktionen auf die Informationsverarbeitung im sozialen Kontext. *Psychologische Rundschau*, 53, S. 61–76.

- Markus, H./Kitayama, S. (1991): Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, S. 224–253.
- Markus, H./Mullally, P./Kitayama, S. (1997): Selfways: Diversity in modes of cultural participation. In U. Neisser/D.A. Jopling (Hrsg.), *The conceptual self in context: Culture, experience, self-understanding* (S. 13–61). Cambridge: Cambridge University Press.
- Oyserman, D./Coon, H.M./Kimmelmeier, M. (2002): Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, 128, S. 3–71.
- Prenzel, M./Merkens, H./Noack, P. (1999): Die Bildungsqualität von Schule: Fachliches und fächerübergreifendes Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten (Antrag an den Senat der DFG auf Einrichtung eines Schwerpunktprogramms). Kiel: Institut für die Pädagogik der Naturwissenschaften.
- Triandis, H.C. (2001): Individualism-collectivism and personality. *Journal of Personality*, 69, S. 907–924.

Anschrift der Autorin:

Prof. Dr. Bettina Hannover, Freie Universität Berlin Fb Erziehungswissenschaft und Psychologie, Habelschwerdter Allee 45, 14195 Berlin.

Anna-Katharina Pelkner/Ralph Günther/Klaus Boehnke

Die Angst vor sozialer Ausgrenzung als leistungshemmender Faktor¹

Zum Stellenwert guter mathematischer Schulleistungen unter Gleichaltrigen

1. Einleitung

Gegenstand der Untersuchung „Streber‘ versus ‚Nerd‘ – Zur Kultur- und Geschlechtsspezifität der (negativen und positiven) Sanktionierung guter mathematisch-naturwissenschaftlicher Schulleistungen bei Jugendlichen in Deutschland und Nordamerika“ ist der Stellenwert guter mathematischer Leistungen unter Gleichaltrigen. Die Studie ist kulturvergleichend angelegt, doch werden zunächst nur Ergebnisse der deutschen Teilstudie vorgestellt.

Während für SchülerInnen in Deutschland gute Schulleistungen oft mit dem Vorwurf einhergehen, ein Streber zu sein, scheinen gute Leistungen in Nordamerika unter Peers hoch im Kurs zu stehen.

Die jüngste Veröffentlichung der PISA-Studie dokumentiert – wie bereits die TIMS-Studie – die im internationalen Vergleich schlechten Leistungen von SchülerInnen deutscher Schulen. Ihre Leistungen liegen deutlich unter dem Durchschnitt der OECD-Länder (Baumert u.a. 2001).

Mit der Einrichtung des Schwerpunktprogramms „Bildungsqualität von Schule“ (BIQUA) hat die Deutsche Forschungsgemeinschaft auf die Diskussion um Bildungsqualität in deutschen Schulen, speziell im mathematisch-naturwissenschaftlichen Unterricht reagiert. Während viele im Schwerpunkt geförderte Projekte Möglichkeiten der Optimierung von Unterricht erforschen, steht im hier vorgestellten Projekt der Zusammenhang zwischen Leistungsbewertung durch Peers und mathematischer Schulleistung im Vordergrund.

Zentrale These ist, dass gute Schulleistungen in Deutschland einem erheblichen Peerdruck unterliegen, sodass gute deutsche SchülerInnen befürchten müssen, bei ihren gleichaltrigen Mitschülern und Freunden schlecht angesehen zu sein und als *Streber* diffamiert zu werden. Dieses Erleben, so weiter die letztendlich im Projekt zu prüfende These, senkt die Bereitschaft insbesondere begabter SchülerInnen der Sekundarstufe I, ihre eigentlich vorhandene Leistungsfähigkeit voll auszuschöpfen.

Im Folgenden wird die Projektidee erläutert und ein kurzer Literaturüberblick gegeben; zudem werden einige theoretische Vorüberlegungen dargelegt. Anschließend wird empirisches Datenmaterial aus der am Standort des Projekts in Chemnitz durchgeführten Erhebung präsentiert.

1 Die berichtete Studie wurde im Rahmen des DFG-Schwerpunktprogramms »Bildungsqualität von Schule« (Az. BO 929/16-1) gefördert.

2. Peerdruck als Leistungshemmnis: Die Projektidee

Informelle Gespräche mit Jugendlichen in Deutschland machen deutlich, dass *Streber* – insbesondere im Altersbereich der Sekundarstufe I – eine der gefürchtetsten Etikettierungen im Peergruppenkontakt ist. Wer in deutschen Schulen oft durch gute Leistungen auffällt und erkennen lässt, dass diese Schulleistungen Ergebnis intensiver Lernbemühungen und nicht „natürlicher Begabung“ sind, wird schnell damit konfrontiert, von MitschülerInnen – in der Regel in der Sekundarstufe I die relevanteste Peergruppe – als Streber benannt und tendenziell ausgegrenzt zu werden.

Wissen vorwiegend anekdotischer Natur aus Nordamerika legt nahe, dass es sich hierbei um ein Kulturspezifikum handelt. Einiges spricht für die These, dass negative Sanktionierung guter Schulleistungen für Deutschland typisch ist, während in Nordamerika gute Schulleistungen unter Peers positiver bewertet werden. Hierfür spricht u.a. die Tatsache, dass eine wörtliche Übersetzung von *Streber* ins Englische schwer fällt. Es scheint lediglich der „nurd“ bzw. „nerd“ in Frage zu kommen, doch hat dieser deutlich andere – zum Teil liebevolle – Konnotationen, die eher dem deutschen Gebrauch des Anglizismus „Freak“ entsprechen (Mathe-Freak). Die Bezeichnung „nerd“ scheint nahe zu legen, dass außergewöhnliche Leistungen zwar den Beigeschmack des Absonderlichen haben, in Nordamerika unter Peers aber tendenziell positiv bewertet werden.

Bildungspolitisch ist die Möglichkeit einer kulturspezifisch in Deutschland besonders negativen Bewertung guter Schulleistungen hochgradig relevant. Wenn nämlich deutsche SchülerInnen der Sekundarstufe I eine massive (negative) Sanktionierung positiver Schulleistungen durch Peers befürchten müssen, so ist zu erwarten, dass hervorragende Leistungen im Sinne einer Art „Selbstbeschneidung“ insbesondere von begabten SchülerInnen, und dabei vor allem von Mädchen vermieden werden.

Köller u.a. (1999) zeigen, dass Mädchen sich stärker vom Urteil anderer abhängig machen, dass zudem die Erwartungen von Eltern und Peers in Bezug auf gute Schulleistungen bei Mädchen signifikant höher als bei Jungen sind und dass sie mehr Angst haben zu versagen als Jungen. Aus derartigen Überlegungen könnte man folgern, dass die Güte der Schulleistungen weniger das Resultat von curricularen, didaktischen, schulstrukturellen oder personalen Bedingungen, sondern vielmehr Ergebnis des leistungsbezogenen Peerklimas ist. Zugespitzt formuliert: Deutsche SchülerInnen der Sekundarstufe I vermeiden besonders gute Schulleistungen, um sich die Anerkennung der Peergruppe zu erhalten, nordamerikanische SchülerInnen suchen besonders gute Schulleistungen, um die Anerkennung der Peers zu gewinnen.

Zu der Frage, ob es einen Zusammenhang zwischen mathematischen Schulleistungen und dem Peerklima in einer Schulklasse gibt, liegen nur wenige kulturvergleichende Studien vor. Eine Ausnahme bilden einige Arbeiten aus der Forschungsgruppe um Herbert Walberg. Walberg/Singh/Tsai (1984) zeigen für indische SchülerInnen, dass Peergruppenstatus und Schulleistung nach Auspartialisierung anderer Variablen in einem positiven Zusammenhang stehen. Uguroglu/Walberg (1986) zeigen für den nordamerikanischen Kontext, dass ein (positiver) Zusammenhang zwischen einem unterstützenden Peerkontext und Mathematikleistungen besteht. Beide Studien verfolgen allerdings

eine andere Kausalitätsannahme als das hier vorgestellte Projekt. Sie gehen davon aus, dass ein hoher Peergruppenstatus gute (Mathematik-)Leistungen *zur Folge* hat. Hier wird hingegen der These nachgegangen, dass (sehr) gute Leistungen in Mathematik – in Deutschland – zu negativen Peersanktionen *führen*, was im weiteren dann zu einer Reduktion der schulischen Performanz führt, während dies in Nordamerika nicht der Fall ist. Juvonen/Murdock (1995) arbeiten eine mögliche Altersspezifik negativer Zusammenhänge zwischen guten Schulleistungen und Peerstatus heraus, wenn sie zeigen, dass insbesondere Achtklässler (im Vergleich zu Viert- und Sechstklässlern) sehr zurückhaltend sind, ihren Peers mitzuteilen, dass sie viel Arbeit in gute Schulleistungen investieren.

Gelegentlich wird in Diskussionen um das Thema der hier vorgestellten Studie davon gesprochen, die Studentenbewegung mit ihrer Ablehnung des Leistungsgedankens sei „schuld“ an der Negativbewertung von Leistung unter deutschen Jugendlichen. Der „Streber“ hat jedoch in Deutschland Geschichte und ist keineswegs erst seit ‘68 *en vogue*². Er hat zudem in Westdeutschland ebenso wie in Ostdeutschland, wo die Studentenbewegung nicht oder nur in sehr viel geringerem Maß wirksam war, die gleiche negative Bedeutung in jugendlichen Peergruppen.

Der niedrige Stellenwert von hervorragenden Leistungen bei deutschen SchülerInnen scheint nicht zuletzt durch die jüngsten Ergebnisse der PISA-Studie belegt, die zumindest für ein nordamerikanisches Land, nämlich Kanada hervorragende Ergebnisse erbrachte.

Versucht man nun die vermuteten Unterschiede zwischen Deutschland und Kanada in einer Art und Weise zu erfassen, die die Moderatorfunktion von Kultur empirisch erfassbar macht, so bietet sich der wertetheoretische Ansatz von Shalom Schwartz an. Das Schwartzsche Konzept kennt auf der Individualebene zehn motivationale Typen von Werthaltungen, nämlich Hedonismus, Stimulation, Selbstbestimmung, Universalismus, Humanismus, Tradition, Konformität, Sicherheit, Macht und Leistung. Besonders in der Bewertung von Leistung als Wert unterscheiden sich nach dem von Schwartz vorgelegten Datenmaterial deutsche und nordamerikanische – US-amerikanische wie kanadische – Individuen (Schwartz 1992). In Nordamerika wird Leistung durchgehend positiver bewertet als in Deutschland. Schwartz zeigt, dass Nordamerika stärker durch Erfolgs- und Machtorientierung, Deutschland stärker durch ein Klima des Egalitarismus und der intellektuellen Autonomie geprägt ist, was sich durchaus auf den Widerspruch von Leistung und Gleichheit zuspitzen lässt (Schwartz 1994). Dieser Widerspruch, so die zweite zentrale These des Projekts, ist letztlich die Ursache für die unterschiedliche Sanktionierung von Leistung in Deutschland und Nordamerika, speziell für die SchülerInnen der Sekundarstufe I, für die Peergruppenintegration eine zentrale Entwicklungsaufgabe ist (vgl. Noack 1990).

2 Im deutschen Sprachgebrauch taucht Streber zuerst im 16. Jahrhundert auf, wobei der Begriff zunächst die Bedeutung von *Widersacher* hatte. Im 18. Jahrhundert diente er zur Kennzeichnung karrieresüchtiger Beamter. Später wurde der Begriff des *Strebers* dann in der Studentensprache und schließlich in die Sprache der Schule übernommen.

Von besonderem Interesse in diesem Kontext ist, wie sich die Peergruppenintegration für Mädchen bzw. Jungen gestaltet. Empirisch zu prüfen wäre die Frage, ob sich Mädchen und Jungen in ihrer Angst vor negativen Peersanktionen guter Schulleistungen unterscheiden und ob auch Geschlechtsunterschiede sich auf unterschiedliche Werthaltungen zurückführen lassen.

Neben der Kultur- die Geschlechtspezifität in den Mittelpunkt eines Projekts zu außerunterrichtlichen Einflüssen auf mathematische Schulleistungen zu rücken, wird u.a. durch die Ergebnisse der TIMS-Studie nahe gelegt, da sich im Bereich der mathematisch-naturwissenschaftlichen Fächer „substantielle Leistungsunterschiede zwischen Schülerinnen und Schülern nachweisen [lassen]“ (Baumert/Bos/Watermann 1998).

Ziel unserer Studie ist es, diese – wie wir meinen nur vordergründigen – Leistungsunterschiede zwischen Mädchen und Jungen genauer zu untersuchen und aus dem Blickwinkel der Leistungsakzeptanz unter Peers zu fragen, warum Mädchen – scheinbar – schlechtere Schulnoten in Mathematik haben als Jungen.

Um dieser Frage nachzugehen ist es sinnvoll, Interaktionen in der Schule in Hinblick auf Geschlechter- und Machtbeziehungen zu analysieren. Von zentraler Bedeutung ist das „kulturelle System der Zweigeschlechtlichkeit“ (Hagemann-White 1988). Menschen werden jeweils dem männlichen bzw. dem weiblichen Geschlecht zugeordnet. Für beide Geschlechter gelten verschiedene Normen, Werte und Standards, die sich in Stereotypen ausdrücken. Die beiden Welten stehen in einer hierarchischen Beziehung zueinander, wobei die männliche den Anspruch erhebt, höherwertig zu sein.

Die Schule als eine der prägendsten Sozialisationsinstanzen vermittelt Geschlechterstereotypen und zementiert so die hierarchisch strukturierten Beziehungen zwischen Männern und Frauen. Das kulturelle System der Zweigeschlechtlichkeit ist eingeschrieben in Texten und Subtexten, sowohl im offiziellen Lehrplan, als auch im sogenannten „heimlichen Lehrplan“ (Zinnecker 1975). Prengel (1986) definiert die „versteckten Lernziele der Geschlechtererziehung“ als die Botschaften „zwischen den Zeilen“. Mädchen und Jungen lernen, dass Frauen und Männer nicht gleich und ihre Beziehung hierarchisch durch Überlegenheit und Unterlegenheit strukturiert sind. Das kulturelle System der Zweigeschlechtlichkeit bestimmt sowohl das tatsächliche Verhalten als auch die Interpretation von Handlungen. Selbst wenn sich Mädchen und Jungen gleich verhalten, lernen sie, dass dieses Verhalten – je nach Geschlecht – eine andere Bedeutung erhält. So wird zum Beispiel aggressives Verhalten eines Jungen in der Regel als normal wahrgenommen, während Mädchen ein solches Verhalten eher nicht zugebilligt wird (vgl. Nyssen 1990, S. 35).

Für ein Mädchen bedeutet es etwas anderes als für einen Jungen, gut in Mathematik zu sein: Wie Hannover in einem weiteren BIQUA-Projekt herausarbeitet, sind Jungen ggf. in einem Fach gut, das von SchülerInnen als „Jungenfach“ apostrophiert wird (Hannover 1999), während Mädchen, die in einem derartigen Fach womöglich gute Schülerinnen sind, eine Art Hautgout des Unweiblichen anhängt. In einem retrospektiv angelegten und international vergleichenden Projekt zu weiblichen Bildungsbiografien von Frauen aus über 10 Ländern arbeitet Pelkner (2002) heraus, dass derartige Ungleichheitserlebnisse die schulische Interaktion weltweit zu bestimmen scheinen.

Seit Einführung der Koedukation werden Schulen in dem Sinne als geschlechtsneutral wahrgenommen, dass sie „keine Unterschiede aufgrund von Geschlecht, Rasse, Klasse, sexueller Orientierung [und] körperlicher Unversehrtheit“ machen (McLaren/Gaskell 1995, S. 136). Eine genauere Analyse zeigt jedoch, dass Schule und Interaktion im Klassenraum weit davon entfernt sind, geschlechtsneutral zu sein. Jungen bekommen in der Schule mehr – negative wie positive – Aufmerksamkeit als Mädchen. Diese Tendenz scheint verstärkt in Fächern wie Mathematik und Naturwissenschaften aufzutreten, da diese nach wie vor eher als männliche Domäne gelten. Lehrpersonen tendieren dazu, Mathematik als männliche Domäne zu stereotypisieren (Li 1999). Im Unterricht wird – bewusst oder unbewusst – der Schein männlicher Überlegenheit aufrecht erhalten und von Jungen oftmals nicht gefordert, sich über tatsächliche Leistung zu qualifizieren (vgl. Kreienbaum 1988). Gleichzeitig wird das Lernengagement der Mädchen als selbstverständlich vorausgesetzt und bleibt ohne positive Rückmeldung. Auf kritisches Lehrerfeedback reagieren Mädchen zudem sensibler als Jungen (vgl. Dweck u.a. 1978).

Nach einer kanadischen Studie (McLaren/Gaskell 1995) hängt das Selbstvertrauen von Mädchen in Bezug auf Mathematik und Naturwissenschaften stärker vom Urteil ihrer Peers als von der Unterstützung der Lehrpersonen ab. Leistungsstarke Mädchen, die ihre Kompetenz in der Klasse demonstrieren, werden von Jungen unter Druck gesetzt (Enders-Drägässer/Fuchs 1988). Für die Bundesrepublik hat Horstkemper (1987) auf die Abhängigkeit des (geringeren) weiblichen Selbstwerts von schulischen Lernbedingungen hingewiesen; sie macht deutlich dass Lehrkräfte beiderlei Geschlecht in bestimmten Fächern geringere Leistungen von Mädchen erwarten und diese dann im Sinne einer sich selbst erfüllenden Prophezeiung auch erhalten. Ähnliche Entwicklungen sind für die DDR zu konstatieren (vgl. Krause 1990). Insbesondere im Hinblick auf die mathematisch-naturwissenschaftlichen Fächer stellten seit den frühen 70er-Jahren SoziologInnen und SchulpraktikerInnen überall dieselben Tendenzen zur Ungleichheit fest: Um die Pubertät herum verlieren Mädchen ihr Interesse an Mathematik. Spätestens in der Sekundarstufe I ist eine deutliche Distanz der Mädchen gegenüber dem mathematisch-naturwissenschaftlichen Unterricht zu beobachten (Eccles 1984; Hoffmann u.a. 1997). Bis heute jedoch fehlt es an breiter angelegten empirischen Studien, die die verschiedenen Hypothesen zu geschlechtsspezifischen Effekten von Interaktionen in der Klasse untersuchen. Die existierenden Studien konzentrieren sich in der Regel auf die Erwachsenenperspektive der Lehrkraft oder anderer erziehender Personen (Glumpler 1995) und sind meist nicht quantitativ angelegt.

3. Die empirische Untersuchung

Dass befürchtete negative Peersanktionen etwas mit den erbrachten Schulleistungen zu tun haben, konnten wir bereits in den Vorstudien zu dem hier vorgestellten Projekt darlegen (Pelkner/Boehnke 2002). Danach „überlagert“ die bei Mädchen größere Angst vor negativen Peersanktionen und ihr geringeres mathematisches Selbstvertrauen in gewisser Weise ihre eigentliche Leistungsfähigkeit: Einerseits haben die Vorstudien erbracht,

dass sich die Mathematiknoten von Jungen und Mädchen nicht unterscheiden. Andererseits konnte statistisch gezeigt werden, dass Mädchen bei gleicher Angst vor negativer Peersanktionierung und bei gleichem mathematischen Selbstwert wie Jungen (sie haben in der Regel mehr Angst vor negativen Peersanktionen und einen geringeren mathematischen Selbstwert als die Jungen) signifikant bessere Mathematiknoten hätten als ihre männlichen Altersgenossen.

Dies legt die Vermutung nahe, dass Mädchen aus Angst vor negativen Sanktionen von Gleichaltrigen dazu neigen, ihre Leistungen nicht in Gänze zu entfalten. Je besser die Leistungen eines Mädchens in Mathematik, desto mehr fürchtet sie als Streber bezeichnet zu werden. Die Vorerhebungen haben die Plausibilität der Erwartung belegt, dass Angst vor dem Strebervorwurf sich bei Mädchen stärker mit der Mathematiknote kovariiert als bei Jungen (Pelkner/Boehnke 2002).

Der hier vorgelegte Beitrag will zunächst die Ergebnisse der Vorerhebungen replizieren. Es soll gezeigt werden, dass der Strebervorwurf bzw. die Angst vor diesem bei deutschen Schülerinnen *und* Schülern mit der Mathematiknote kovariiert. Die Frage, ob es bestimmte Prädiktoren gibt, mit der die Schulnote in Mathematik, insbesondere von Mädchen, vorausgesagt werden kann, soll im Folgenden anhand der Daten der Chemnitzer Teilstudie des Projekts exploriert werden. Leitidee ist dabei letztlich, wie bereits ausgeführt, dass befürchtete negative Peersanktionen als *Folge* besonders guter Leistungen in Mathematik bei guten Schülern und insbesondere bei leistungsstarken Mädchen *die längerfristig* die Ausschöpfung des vorhandenen Leistungspotenzials begrenzt.

In einem weiteren Auswertungsschritt soll gezeigt werden, dass Wertorientierungen den Zusammenhang zwischen Streberangst und Schulleistungen moderieren. Dieser Teil der Auswertungen wird geleitet von dem Gedanken, dass nur Individuen, bei denen Leistung als Wert hoch im Kurs steht (wie wir es in den noch ausstehenden kulturvergleichenden Analysen für Nordamerika typischerweise annehmen), in ihrem Wunsch, gute Leistungen zu erbringen so gefestigt sind, dass sie der Vorwurf „Streber“ zu sein, eher kalt lässt, während für Individuen, die Leistungswerte eher ablehnen, die Zuschreibung guter, durch „Streberei“ erreichter Schulleistungen stärker angstbesetzt sein sollte.

In die Untersuchung einbezogen wird auch die Frage der geschlechtsspezifischen Ausprägung des mathematischen Selbstwerts. Diese ist, dass dieser bei Mädchen geringer ist als bei Jungen und die eigentlich vorhandene Leistungsfähigkeit der Mädchen massiv behindert. Versucht man die geäußerten Vermutungen in Hypothesenform zu präzisieren, so geht es um folgende Annahmen:

- 1) Die Angst als Streber benannt zu werden, kovariiert negativ mit der Schulnote in Mathematik³.
- 3) Hierzu ist anzumerken, dass die Hypothese sich auf den erwarteten Momentanbefund bezieht, nicht auf die eigentlich unterliegende Prozessannahme: Es wird erwartet, dass die besten SchülerInnen die meiste Angst davor haben, Streber genannt zu werden. Numerisch muss dabei die erwartete Korrelation negativ sein, da Schulnoten revers gepolt sind, niedrige Werte stehen für gute Leistungen. Im aktuellen Design der Studie nicht testbar ist die Prozessannahme, ob sich *frühere* hohe Streberangst in *späterer* niedrigerer Leistungsfähigkeit manifestiert.

- 2) Die Kovariation ist bei Mädchen stärker als bei Jungen.
- 3) Die Höhe der Kovariation wird durch die Akzeptanz von Leistungswerten moderiert. Bei SchülerInnen, die Leistungswerten positiv gegenüber stehen, ist die Kovariation von Streber-Angst und Mathematiknote geringer als bei SchülerInnen, die Leistungswerte ablehnen.
- 4) Der in Hypothese (3) formulierte Zusammenhang ist bei Mädchen stärker ausgeprägt als bei Jungen.
- 5) Nach einer statistischen Ausparialisierung von Unterschieden in der „Streber-Angst“ und im mathematischen Selbstwert haben Mädchen bessere mathematische Schulleistungen.
- 6) Streber-Angst und mathematischer Selbstwert sind bessere Prädiktoren der Schulnote in Mathematik als die mathematische Leistungsfähigkeit, wie sie in TIMSS-Aufgaben erfasst wird.
- 7) Erneut gilt dies in ausgeprägterer Form für Mädchen als für Jungen.

3.1 Stichprobe

Zur empirischen Überprüfung der Hypothesen werden die Daten der Chemnitzer Hauptuntersuchung herangezogen. Am Standort Chemnitz wurden – altersanalog zur TIMS-Studie – SchülerInnen achter Schulklassen von Gymnasien und Mittelschulen in die Studie einbezogen, insgesamt 661, die zu 55,7% der Mittelschule und zu 44,3% dem Gymnasium entstammen; 52,5% sind Mädchen, 47,5% Jungen.

3.2 Erhebungsinstrumente

Der hier berichtete Teil der „Streber“-Studie wurde als Fragebogenstudie durchgeführt. Dabei wurden ausschließlich vorgetestete Erhebungsinstrumente eingesetzt. Für die hier vorgelegte Auswertung relevant sind

- a) ein selbstentwickeltes Instrument zur Bedeutung von „Streber“-Angst
- b) die Porträtversion des Wertefragebogens von Schwartz (vgl. Schwartz u.a. 2001)
- c) die letzte Mathematiknote
- d) drei TIMSS-Aufgaben
- e) die Selbsteinschätzung der eigenen mathematischen Leistungsfähigkeit (vgl. Westbury/Travers 1990).

Zu (a) wurden die Befragten gebeten, zu sieben Stereotypen (u.a. dem „Streber“) anzugeben, in welchem Ausmaß sie Angst davor haben, so genannt zu werden. Für die hier vorgelegten Auswertungen wurde nur das Item berücksichtigt, das unmittelbar danach fragt, ob die Befragten (0 = nie bis 3 = häufig) Angst haben, Streber genannt zu werden.

Die zu (b) eingesetzte Porträt-Version des Schwartzschen Wertinventars fragt: „Im Folgenden findest du einige kurze Personenbeschreibungen. Bitte lies die einzelnen Beschreibungen aufmerksam durch. Überlege dann, wie ähnlich oder unähnlich dir die jeweilige Person ist und trage dann rechts eine der folgenden Ziffern ein: 1 = sehr unähnlich, 2 = unähnlich, 3 = eher unähnlich, 4 = eher ähnlich, 5 = ähnlich, 6 = sehr ähnlich.“ Es folgen 22 Personenbeschreibungen. Eine hier relevante Personenbeschreibung zu Leistungswerten lautete z.B. „Es ist ihr/ihm sehr wichtig, erfolgreich zu sein. Sie/er mag es, hervorragend abzuschneiden und andere Leute zu beeindrucken“. Die 4-Item-Skala zur Präferenz von Leistungswerten hatte eine interne Konsistenz von $\alpha = 0,72$.

Die Frage zur letzten Mathematiknote (c) lautete „Welche Note hattest du auf dem letzten Zeugnis in Mathematik?“

Aus der TIMS-Studie (d) wurden drei Aufgaben (Stichwort: „GLANZI“, „Wildpark“ und „Geschenkschachtel“) eingesetzt. Die GLANZI-Aufgabe lautete „GLANZI Waschpulver wird in würfelförmigen Kartons verkauft. Ein Karton hat eine Kantenlänge von 10 cm. Die Herstellerfirma beschließt, die Länge jeder Kante des Kartons um 10 Prozent zu vergrößern. Um wie viel nimmt das Volumen zu? A 10 cm³, B 21 cm³, C 100 cm³, D 331 cm³“.

Die der Second International Mathematics Study (SIMS, Westbury/Travers 1990) entstammende Skala zum mathematischen Selbstwert (e) umfasste 9 Items u.a. z.B. „Egal wie sehr ich mich anstrenge, ich bin nicht gut in Mathe“. Die Skala hatte eine interne Konsistenz von $\alpha = 0,92$.

3.3 Analysen

Zur Überprüfung der oben formulierten Hypothesen wurden Korrelations- und Regressionsanalysen eingesetzt.

Hypothese (1) nimmt an, dass „Streber-Angst“ mit der Mathematiknote kovariiert. Zur Prüfung dieser Hypothese wurden Produkt-Moment-Korrelationskoeffizienten errechnet. Dabei wurde Schultyp als Dummyvariable (Mittelschule = 1, Gymnasium = 2) auspartialisiert, um der Möglichkeit Rechnung zu tragen, dass in Mittelschule und Gymnasium unterschiedlich benotet wird⁴. Die Partialkorrelation von Streberangst und letzter Mathematiknote lag bei $r = -0.10$, $p = .01$. Das Ergebnis bestätigt Hypothese (1): Streber-Angst ist ausgeprägter bei Achtklässlern mit guten Mathematiknoten.

Hypothese (2) besagt, dass der in Hypothese (1) postulierte Zusammenhang bei Mädchen höher sein sollte als bei Jungen. Für Mädchen lag die Partialkorrelation von Streber-Angst und Mathematiknote bei $r = -0.16$, $p < .001$, für Jungen lag diese bei

4 Auch Schulnoten werden als intervallskalierte Variablen betrachtet. Im Kontext der hier vorgelegten Studie legt insbesondere auch die Notwendigkeit der Berechnung von Partialkorrelationen und Regressionsanalysen dies nahe, die als nonparametrische Analysen einen sehr hohen Aufwand erfordern würden und ebenfalls an einige zusätzliche Voraussetzungen gebunden wären (Bortz/Lienert/Boehnke 2000).

$r = 0,03$ (n.s.). Die Differenz zwischen beiden Korrelationskoeffizienten ist – geprüft über Fischers-Z-Test – ebenfalls signifikant ($p < .01$). Auch Hypothese (2) findet Bestätigung. Die Ergebnisse machen allerdings deutlich, dass der Befund zu Hypothese (1) in dem Sinne neu bewertet werden muss, dass es einen Zusammenhang zwischen Mathematiknote und Streber-Angst ausschließlich bei Mädchen gibt und dass die Signifikanz in der Gesamtstichprobe einzig auf einen stärkeren Zusammenhang bei Mädchen zurück zu führen ist.

Hypothese (3) postuliert, dass die Kovariation von Streberangst und Mathematiknote durch die Zustimmung zu Leistungswerten moderiert ist, eine besonders ausgeprägte Kovariation wird für solche SchülerInnen vermutet, die Leistungswerte ablehnen, während für SchülerInnen mit einer hohen Akzeptanz von Leistungswerten eine geringere Kovariation erwartet wird. Diese These wurde geprüft, indem die Gesamtstichprobe hinsichtlich ihrer Zustimmung zu Leistungswerten auf der Basis des Skalenwerts der vier Schwartz-Leistungsitens medianhalbiert wurde und dann die bereits berichteten Korrelationsanalysen getrennt für die Gruppe der Leistungswertbefürworter (LWB) und der Leistungswertablehner (LWA) durchgeführt wurden. Die Partialkorrelation von Streber-Angst und Mathematiknote liegt bei den LWA bei $r = -0.12$ (n.s.)⁵, bei den LWB hingegen bei $r = -0.09$ (n.s.). Die Differenz zwischen beiden Korrelationskoeffizienten entspricht zwar prima face der erwarteten Richtung verfehlt statistische Signifikanz aber deutlich. Hypothese (3) muss deshalb zurückgewiesen werden.

Hypothese (4) nimmt auch für den Moderatoreffekt der Wertorientierung Geschlechtsunterschiede an. Um diese Hypothese zu prüfen, wurden die zu (3) berichteten Analysen noch einmal getrennt nach Geschlecht durchgeführt. Die Partialkorrelationen zwischen Streber-Angst und Mathematiknote lagen für LWA-Mädchen bei $r = -0.18$, $p < .05$, für LWB-Mädchen bei $r = -0.14$, $p < .10$. Bei Jungen lagen die entsprechenden Koeffizienten bei $r = 0.01$ (n.s.) (LWA) bzw. bei $r = -0.05$ (n.s.) (LWB). Die Analysen machen erneut die Bedeutsamkeit des Geschlechtseffekts deutlich, können aber die Bedeutsamkeit des Moderators „Leistungswerte“ erneut nur prima face aufzeigen, ohne dass ein Beleg seiner statistischen Bedeutsamkeit gelingt.

Hypothese (5) postuliert, dass Mädchen bessere Mathematiknoten als Jungen hätten, wenn sie nicht eine stärkere „Streber-Angst“ als die Jungen und ein geringeres mathematisches Selbstwertgefühl als diese hätten. Für die Prüfung von Hypothese (5) wurden zwei Regressionsanalysen durchgeführt. Abhängige Variable (AV) der Analysen war die Mathematiknote. In der ersten Analyse wurde die AV ausschließlich durch das Geschlecht vorhergesagt. In der zweiten Analyse wurden zusätzlich die Dummyvariable Schultyp, die Streberangst und der mathematische Selbstwert in die Regressionsgleichung aufgenommen. Die Selbstwertskala, die ausschließlich Items enthielt, in denen geringer mathematischer Selbstwert zum Ausdruck kam, wurde für diese Analyse umgepolt, sodass hohe Werte für hohen mathematischen Selbstwert stehen. Tabelle 1 berichtet die ermittelten β -Regressionskoeffizienten.

5 Die exakte Wahrscheinlichkeit liegt bei $p = .051$ und verfehlt somit Signifikanz nur sehr knapp.

Tab. 1: Prädiktion der Mathematiknote^a			
Prädiktor		Analyse 1	Analyse 2
		β	β
Geschlecht		-0.004	0.076*
Schultyp		-	-0,088**
Streberangst		-	-0.121***
Selbstwert		-	-0.555***
R ²		0.000	0.319 ^c
(ggf. partialisierter) Notenmittelwert		Mädchen/Jungen 2.88/2.88	Mädchen/Jungen 2.82/2.96 ^d
<p>a Bei der Interpretation ist wiederum zu beachten, dass bei Noten gute Mathematikleistungen durch niedrige Zahlenwerte wiedergegeben werden.</p> <p>b + p <.10, * p <.05, ** p <.01, *** p <.001</p> <p>c Hier nicht berichtete sukzessive Analysen, in denen die Prädiktoren Schultyp, Streberangst und mathematischer Selbstwert nacheinander in die Gleichung aufgenommen werden, zeigen, dass R²_{change} in jedem Fall signifikant (p <.01) ist.</p> <p>d Diese Berechnungen entstammen analogen Kovarianzanalysen.</p>			

Die Tabelle macht deutlich, dass sich die Rohwerte der Mathematiknoten von Jungen und Mädchen nicht unterscheiden. Führt man jedoch zusätzlich den Schultyp, die Streber-Angst und den mathematischen Selbstwert als Prädiktoren ein, so zeigt sich, dass Mädchen nach Auspartialisierung dieser Variablen signifikant bessere Mathematiknoten haben als Jungen. Gleichzeitig wird deutlich, dass Selbstwert und Streber-Angst zwei voneinander unabhängige Variablen sind, von denen allerdings der mathematische Selbstwert in einem wesentlich höheren Zusammenhang mit der Schulnote steht als die Streberangst.

Hypothese (6) postuliert, dass Streberangst und mathematischer Selbstwert bessere Prädiktoren der Mathematiknote sind als die mit TIMSS-Aufgaben gemessene mathematische Leistungsfähigkeit. Um die Hypothese zu testen wurde eine weitere Regressionsanalyse mit dem Kriterium Mathematiknote durchgeführt, in der nun – neben dem Schultyp – die Streber-Angst, der mathematische Selbstwert und die Anzahl der gelösten TIMSS-Aufgaben als Prädiktoren fungiert. Tabelle 2 (S. 336) berichtet die β -Koeffizienten dieser Analyse.

Hypothese (7) postuliert, dass der in Hypothese (6) postulierte Effekt bei Mädchen ausgeprägter ist als bei Jungen. Hierzu wurden die gerade beschriebenen Regressionsanalysen noch einmal getrennt für Mädchen und Jungen durchgeführt. Die β -Koeffizienten dieser Analyse werden ebenfalls in Tabelle 2 dokumentiert.

Tab. 2: Prädiktion der Mathematiknote auf der Basis von Schultyp, Streber-Angst, mathematischem Selbstwert und gelösten TIMSS-Aufgaben				
Prädiktor		Gesamtstichprobe	Mädchen	Jungen
		β	β	β
Schultyp		-0.051	-0.055	-0.029
Streberangst		-0.133***	-0.181***	-0.042
Selbstwert		-0.524***	-0.578***	-0.479
TIMSS-Leistung		-0.131***	-0.164***	0.108 ⁺
R ²		0.329	0.404	0.272

Die Tabelle zeigt, dass Hypothese (6) prima face angenommen werden kann. Die Bedeutung von Streber-Angst und mathematischem Selbstwert als Prädiktoren der Mathematiknote ist höher als die der mit TIMSS-Aufgaben gemessenen mathematischen Leistungsfähigkeit. Der Unterschied in der Bedeutsamkeit der Streber-Angst und der TIMSS-Leistung ist dabei allerdings nicht statistisch signifikant. Gleichzeitig wird auch im Lichte der in Tabelle 2 dokumentierten Analysen erneut deutlich, dass für die Möglichkeit einer Annahme der Hypothese erneut ausschließlich die Mädchen „verantwortlich“ sind. Bei den Jungen spielt die Streber-Angst in der Vorhersage der Mathematiknote überhaupt keine Rolle; es überrascht allerdings, dass auch die Anzahl der gelösten TIMSS-Aufgaben keine statistisch signifikante Vorhersage der Mathematiknote erlaubt.

4. Diskussion

Die „Streber“-Studie postuliert, dass die Leistungsbewertung durch Peers in einem Zusammenhang mit den tatsächlichen Schulleistungen in Mathematik steht und dass dieser Zusammenhang unmittelbar mit der Akzeptanz von Leistungswerten in Beziehung steht. Gegenstand des Projekts sind dabei letztlich Kulturunterschiede zwischen Deutschland und Kanada. Es wird angenommen, dass in Deutschland der postulierte Zusammenhang wesentlich stärker ist als in Kanada. Gleichzeitig wird postuliert, dass die Abhängigkeit der mathematischen Schulleistungen vom Peerurteil bei Mädchen wesentlich ausgeprägter ist als bei Jungen.

Der postulierte Kulturunterschied kann derzeit noch nicht angemessen überprüft werden. Deshalb wurde versucht, die zentralen Hypothesen des Projekts zunächst anhand des Chemnitzer Datensatzes zu prüfen. Um auch die Bedeutung der Präferenz von Leistungswerten, von denen angenommen wird, dass sie den Kulturunterschied zwischen Deutschland und Kanada „ausmachen“, ansatzweise prüfen zu können, wurde die Chemnitzer Stichprobe auf der Basis von Ratings zu vier Leistungsitems aus dem Schwartzschen Wertinventar medianhalbiert. Analysen zur Hypothese einer vermuteten Bedeutsamkeit von Kulturunterschieden wurden getrennt für die Teilgruppen der Leistungswertbefürworter und der Leistungswertablehner vorgenommen.

Die in Abschnitt 3.3 dokumentierten Analysen erbringen insgesamt ein „durchwachsendes“ Bild. Die postulierten Zusammenhänge zwischen der mathematischen Schulleistung und dem leistungsbezogenen Peerklima (Stichwort: Streber-Angst) ließen sich für die Gesamtstichprobe belegen, doch sind die Effektgrößen gering. Nur etwa 1% der Varianz in den Schulnoten lässt sich auf Basis der Kenntnis des Ausmaßes von Streber-Ängsten deuten.

Nach Geschlecht getrennte Analysen machen zudem deutlich, dass ein Konnex zwischen Streber-Angst und Mathematiknote ausschließlich bei Mädchen besteht. Nur Mädchen leiden substantiell unter der Angst von anderen wegen guter Mathematikleistungen ausgegrenzt zu werden; nur für Mädchen lässt sich deshalb die dem Projekt zugrunde liegende Prozesshypothese aufrechterhalten, dass Streber-Ängste dazu führen könnten, dass mathematische Schulleistungen nach unten „justiert“ werden, wodurch sich auf längere Sicht dann auch die mathematische Leistungsfähigkeit reduzieren dürfte.

Die Analysen zur möglichen Bedeutung von Werthaltungen erbringen keine rundum überzeugenden Ergebnisse. Zwar zeigt ein Blick auf die ermittelten Koeffizienten in allen Fällen ein mit den aufgestellten Hypothesen kompatibles Ergebnis, doch lässt sich statistische Bedeutsamkeit in keinem Fall belegen: Bei SchülerInnen, die Leistungswerte ablehnen, und dabei insbesondere bei Mädchen, ist die Bedeutung der Streber-Angst als Prädiktor der Schulnote jeweils größer als bei SchülerInnen, die Leistungswerten zustimmen, doch sind die gefundenen Effekte nur marginal.

Summa summarum lässt sich festhalten, dass negative Peersanktionen guter Schulleistungen für Mädchen die Gefahr negativer Auswirkungen auf ihre Mathematiknote in sich bergen. Dieser Befund ist trotz geringer Effektstärken bedeutsam, da gezeigt werden kann, dass Ängste vor Peersanktionen im Verbund mit einem bei Mädchen gering ausgeprägten mathematischen Selbstwert durchaus dafür „verantwortlich“ gemacht werden können, dass Mädchen nicht die – eigentlich möglichen – besseren Mathematiknoten als Jungen erreichen. Wie wichtig dieser Befund ist, wird deutlich, wenn man in einer weiteren Regressionsanalyse den relativen Anteil von Peersanktionen, mathematischem Selbstwert und objektiver Leistung an der Mathematiknote – unter Ausparialisierung des Schultyps – exploriert. Bei Jungen lauten die entsprechenden β -Koeffizienten -0.07 für die Streber-Angst, -0.47 für den mathematischen Selbstwert und -0.08 für die TIMSS-Leistung. Bei Mädchen ist das Ergebnis -0.16 , -0.58 und -0.17 . Diese Koeffizienten machen zum einen die geringe Bedeutung der objektiven Leistungen für die vergebenen Noten deutlich, verweisen aber zugleich darauf, dass Streber-Ängste für die Prädiktion der Schulnote eine annähernd gleich hohe Bedeutung haben wie die objektive Leistungsfähigkeit. Weiterhin machen sie noch einmal deutlich (vgl. Boehnke 1996) von welcher herausragenden Bedeutung der Selbstwert für die in der Schule erworbenen Noten ist.

Die zu den Hypothesen (5) und (6) vorgelegten Analysen zeigen, dass Geschlechtsunterschiede in objektiven Mathematikleistungen zugunsten der Jungen sich im Wesentlichen als Unterschiede in der Folge eines unterschiedlichen mathematischen Selbstwerts deuten lassen. In den Noten als weiterem Leistungsindikator finden sich keine Geschlechtsunterschiede. Auch hier allerdings haben Unterschiede im mathematischen

Selbstwert und darüber hinaus hier auch die Streber-Angst eine große Bedeutung. Gäbe es in letzteren keine massiven Unterschiede zuungunsten der Mädchen, hätten diese deutlich bessere Mathematiknoten als die Jungen.

Dieser Befund wird nicht von Wertorientierungen moderiert. Es spielt keine Rolle, ob ein Schüler bzw. eine Schülerin Leistung hoch bewertet oder nicht. Die Geschlechtsunterschiede in den Auswirkungen von negativen Peersanktionen und geringem Selbstwert auf die mathematische Leistungsfähigkeiten hängen nicht von der Zustimmung zu Leistungswerten ab. Nicht Leistungswerte von Individuen sondern objektiv unterschiedliche Rollen von Mädchen und Jungen könnten hier die Ursache sein.

Insgesamt hat die Chemnitzer Teiluntersuchung der „Streber“-Studie gezeigt, dass die zentralen Annahmen des Projekts sich für die Deutung Geschlechts- und – hier nur simulierten – Kulturunterschieden in der schulischen Leistungsbewertung durch Noten als valide erwiesen haben, für die Deutung von Unterschieden in der objektiven mathematischen Leistungsfähigkeit aber nicht.

Forschungsdieser Nr. 1 wäre, die hier nur simulierten Kulturunterschiede im deutsch-kanadischen Vergleich auch tatsächlich zu prüfen. Diese Aufgabe wird noch im Rahmen dieses Projekts zu erledigen sein.

Anzumerken wäre darüber hinaus aber, dass sämtliche hier dargelegten Befunde unter der Kautele stehen, dass es sich um Korrelationsdaten handelt. Solche Daten schließen die insinuierte Kausalität forschungslogisch aus. Postuliert wird in der Streberstudie, ohne dass dies explizit so formuliert wurde, dass Mädchen, *weil* sie Angst vor negativen Peersanktionen haben, ihre mathematischen Leistungen antizipativ nach unten korrigieren. Die Überprüfung eines derart starken Kausalitätspostulats aber ist in einer korrelativen Studie nicht möglich. Forschungsdieser 2 ist deshalb, die hier erneut als plausibel belegte Annahme eines Zusammenhangs zwischen negativen Peersanktionen und Mathematikleistungen in einer Panelstudie und/oder in einer quasi-experimentellen Studie zu prüfen.

Literatur

- Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiß, M. (Hrsg.) (2001): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Baumert J./Bos, W./Watermann, R. (1998): Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse. http://www.mpib-berlin.mpg.de/TIMSS_III/Zusammenfassung.htm.
- Boehnke, K. (1996): Is intelligence negligible? The relationship of family climate and school behavior in a cross-cultural perspective. Münster: Waxmann.
- Bortz, J./Lienert, G.-A./Boehnke, K. (2000): Verteilungsfreie Methoden in der Biostatistik, 2. Auflage. Berlin: Springer
- DIE ZEIT, 6. Dezember 2001: Wieso, weshalb, warum? Über die Ursachen der Bildungsmisere und wie man Schule besser machen kann. Jürgen Baumert und Hermann Lange im ZEIT-Gespräch., S. 46–47.

- Dweck, C.S./Davidson, W./Nelson, S./Enna, B. (1978): Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. In: *Developmental Psychology*, Vol. 14, S. 268–276.
- Eccles, J. (1984): Sex differences in achievement patterns. In: *Nebraska Symposium on Motivation*, Vol. 32, S. 97–132.
- Enders-Drägässer, U./Fuchs, C. (1988): Interaktionen und Beziehungsstrukturen in der Schule. Eine Untersuchung an hessischen Schulen im Auftrag des hessischen Instituts für Bildungsplanung und Schulentwicklung. Frankfurt: Hessisches Institut für Bildungsplanung und Schulentwicklung.
- Glumpler, E. (1995): Feministische Schulforschung. In: Roloff, H.-G. (Hg.): *Zukunftsfelder von Schulforschung*. Weinheim: Beltz, S. 133–155.
- Hagemann-White, C. (1988): *Sozialisation: weiblich-männlich*, Opladen: Leske + Budrich.
- Hannover, B. (1999): Schulischer Kontext, geschlechtsbezogenes Selbstwissen und Fachpräferenzen. In: Hannover, B./Kittler, U./Metz-Göckel, H. (Hg.), *Sozialkognitive Aspekte der pädagogischen Psychologie*. Band I. Dokumentation des 3. Dortmunder Symposions für Pädagogische Psychologie 1998. Essen: Die Blaue Eule, S. 125–139.
- Hoffmann, L./Häussler, P./Lehrke, M. (1997). *Die IPN-Interessenstudie Physik*. Kiel: Institut für die Pädagogik der Naturwissenschaften an der Univerität Kiel.
- Horstkemper, M. (1987): *Schule, Geschlecht und Selbstvertrauen: Eine Längsschnittstudie über Mädchensozialisation in der Schule*. Weinheim: Juventa.
- <http://www.mpib-berlin.mpg.de/TIMSS-Germany/home-d.htm>; Testaufgaben TIMSS/III, S. 35+40; Stichwort: Bausausflug in einen Wildpark bzw. Glanzi Waschpulver.
- Juvonen, J./Murdoch, T.B. (1995): Grade-level differences in the social value of effort: Implications for self-presentation tactics of early adolescents. In: *Child Development*, Vol. 66(6), S. 1694–1705.
- Kahl, R. (2001): Depressive Zirkel gibt es genug. PISA zur Mutter der Erneuerung machen. In: *Erziehung und Wissenschaft*. Allgemeine Deutsche Lehrerzeitung, 53. Jg., S. 2.
- Kölller, O./Baumert, J./Clausen, M./Hosenfeld, I. (1999): Predicting Mathematics Achievement of Eighth Grade Students in Germany: An Application of Parts of the Model of Educational Productivity to the TIMSS Data. In: *Educational Research and Evaluation* 5, S. 180–194.
- Krause, C. (1990): Mädchen und Jungen in der Schule der DDR. In: *Die Deutsche Schule*, Beiheft 1, S. 126–138.
- Kreienbaum, M.A. (1988): Frauen bilden Macht. Fachkongress Frauen und Schule. In: *Frauen und Schule*, 7. Jg., H. 23, S. 24–25.
- Lee, V. E./Bryk, A. S. (1986): Effects of single-sex secondary schools on student achievement and attitudes. In: *Journal of Educational Psychology*, Vol. 78, S. 381–395.
- Li, Q. (1999): Teachers' beliefs and gender differences in mathematics: a review. In: *Educational Research*, Vol. 41 (1), S. 63–76.
- McLaren, A./Gaskell, J. (1995): Now you see it, now you don't: Gender as an issue in school science. In: Gaskell, J./Willinsky, J. (Eds.): *Gender informs Curriculum: From Enrichment to Transformation*. Toronto, pp. 136–156.
- Ministry of Education (1995): *Royal Commission on Learning provides a blueprint for changing Ontario schools*. <http://www.edu.gov.on.ca/eng/general/abcs/rcom/news.html>
- Noack, P. (1990): *Jugendentwicklung im Kontext. Zum aktiven Umgang mit sozialen Entwicklungsaufgaben in der Freizeit*. München: PsychologieVerlagsUnion.
- Nyssen, E. (1990): Aufwachsen im System der Zweigeschlechtlichkeit. In: Metz-Göckel, S./Nyssen, E.: *Frauen leben Widersprüche*. Weinheim: Beltz, S. 25–48.
- Pechtl, S. (1994): *Chancengleichheit zwischen Knaben und Mädchen: ein Überblick über die feministische Schulforschung mit Bezugnahme auf das gegenwärtige österreichische Bildungssystem*. Frankfurt: Lang.
- Pelkner, A.-K. (2002): *Brave Mädchen und freche Gören: Erfahrungen mit Schule aus Internationalen und Interkulturellen Perspektiven*. In: *Project Body* (Hg.): *Korporealitäten*. In(ter)ventionen in ein omnipräsentes Thema, Königsstein/Taunus, Ulrike Helmer Verlag, (Arbeitstitel).

- Pelkner A.-K./Boehnke K. (2002): Streber als Leistungsverweigerer? Projektidee und erstes Datenmaterial einer Studie zu mathematischen Schulleistungen”, vorgesehen zur Veröffentlichung in der Zeitschrift für Erziehungswissenschaft.
- PISA 2001: Kurzfassung der PISA-Ergebnisse. www.gew-berlin.de
- Prenzel, A. (1986): Gleichberechtigung – ein utopisches Ziel?. In: *Frauen & Schule*, Nr. 14, Berlin.
- Royal Commission on Learnig (1994): *For the Love of Learning: Report of the Royal Commission on Learnig*. Toronto: Queen’s Printer for Ontario.
- Schwartz, S.H. (1992): Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In: *Advances in Experimental Social Psychology*, Vol. 25, S. 1–65.
- Schwartz, S.H. (1994): Beyond Individualism/Collectivism: New cultural dimensions of values. In: KIM, U./Triandis, H.C./Kagitcibasi, C./Choi, S.-C./Yoon, G. (Eds.): *Individualism and collectivism: Theory, method and applications*. Newbury Park, CA, S. 85–119.
- Schwartz, S.H./Melech, G./Lehmann, A./Burgess, S./Harris, M., Owens, V. (2001): Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. In: *Journal of Cross-Cultural Psychology* 32, S. 519–542.
- Uguroglu, M.E./Walberg, H.J. (1986): Predicting achievement and motivation. In: *Journal of Research and Development in Education*, Vol. 19(3), S. 1–12.
- Walberg, H.J./Singh, R.P./Tsai, S.-L. (1984): The achievement productivity of psychological variables: A replication and extension in Rajasthan, India. In: *Alberta Journal of Educational Research*, Vol. 30(3), S. 171–178.
- Westbury, I./Travers, K. (1990): *Second International Mathematics Study*. Urbana: University of Illinois.
- Zinnecker, J. (1975): *Der heimliche Lehrplan. Untersuchungen zum Schulunterricht*. Weinheim: Beltz.

Anschrift der Autoren:

Dipl.-Soz. Anna-Katharina Pelkner, Technische Universität Chemnitz, Institut für Soziologie, 09107 Chemnitz.

Dipl.-Soz. Ralph Günther, Technische Universität Chemnitz, Institut für Soziologie, 09107 Chemnitz.

Prof. Dr. Klaus Boehnke, International University Bremen, School of Humanities and Social Science, Campus Ring 1, 28759 Bremen.

Bettina Hannover/Ursula Kessels

Challenge the science-stereotype!¹

Der Einfluss von Technik-Freizeitkursen auf das Naturwissenschaften-Stereotyp von Schülerinnen und Schülern

1. Ausgangslage

In der „Third International Mathematics and Science Study“ (Baumert/Bos/Lehmann 2000) lagen deutsche Schülerinnen und Schüler der Sekundarstufen I und II beim Vergleich von insgesamt 40 Ländern bezüglich ihrer Leistungen in der Mathematik lediglich auf Platz 25 und bezüglich ihrer Leistungen im naturwissenschaftlichen Bereich auf Platz 20. Die geringen Leistungen korrespondieren einem schwachen Interesse an diesen Fächern (Krapp 2000). Weiter machte TIMSS substanzielle Geschlechtsunterschiede im Interesse und in den Leistungen zuungunsten der Schülerinnen sichtbar: Während sich Mädchen und Jungen in der Primarstufe in ihren schulischen Interessen und Leistungen noch gleichen, setzt mit Beginn der Sekundarstufe I eine Polarisierung ein, die in stark geschlechtstypisierten Fachinteressen und -leistungen endet (vgl. z.B. Hoffmann/Häußler/Lehrke 1998; Seidel u.a. 2002).

Aktuelle Statistiken über Leistungs- und Studienfachwahlen an deutschen Schulen und Universitäten zeigen einen anhaltenden Trend weg von den Naturwissenschaften. Beispielsweise sank bei den Leistungskurswahlen in Baden-Württemberg der Anteil der naturwissenschaftlichen Fächer (Physik, Chemie und Biologie) im Zeitraum von 1979 bis 1999 von ehemals 31% auf nur noch 16%, wohingegen Deutsch und musische Fächer ihren gemeinsamen Anteil von 13% auf 18% ausweiten konnten (Zwick/Renn 2000)². Korrespondierend zeigt die Veränderung der Studierendenzahlen in den verschiedenen Fachgruppen – wiederum exemplarisch für Baden-Württemberg – ein ähnliches Bild: Einem allein demographischen Entwicklungen geschuldeten generellen Rückgang der Studierendenzahlen um minus 10,5% (Zeitraum: 1993/94 bis 1997/1998) stehen unter- und überproportionale Abnahmen in einzelnen Fachgruppen und Fächern gegenüber: Die Fachgruppe der Rechts-, Wirtschafts- und Sozialwissenschaften wies bloß eine Abnahme um minus 1,3% der Studierenden auf und in den Sprach- und Kulturwissenschaften gab es eine unterdurchschnittliche Abnahme von minus 6,5%. Dagegen im Bereich der Ingenieurwissenschaften sind die Zahlen der Studierenden im gleichen Zeitraum um insgesamt 19,2% zurückgegangen und im Bereich der mathematisch-naturwissenschaftlichen Studiengänge um 20,6% gesunken; Spitzenreiter darunter

- 1 Die berichtete Studie wurde im Rahmen des DFG-Schwerpunktprogramms »Bildungsqualität von Schule« unter Az. HA 2381/5-1 gefördert.
- 2 LK-Wahlen 1999, verbleibende Prozentzahlen: Fremdsprachen 30%, Mathematik: 18%, Gesellschaftswissenschaften, Theologie und Sport: 18%.

die Physik/Astronomie mit einem Rückgang um 39,5% und Chemie mit 40,1% (Zwick/Renn 2000). In Folge der geringen Beliebtheit mathematisch-, natur- und ingenieurwissenschaftlicher Fächer im Vergleich zu den geistes-, kultur-, wirtschafts- und sozialwissenschaftlichen Fächern besteht in Deutschland ein erheblicher Mangel an entsprechenden Fachkräften. Eine Konsequenz dieser beunruhigenden Entwicklungen war die Einrichtung des DFG-Schwerpunktprogrammes „Bildungsqualität von Schule“ (BIQUA).

2. Image und Prototypen

Im Fokus unseres im Schwerpunktprogramm BIQUA angesiedelten Projektes steht das Image naturwissenschaftlicher Fächer und das Image von Personen, die sich für diese Fächer in besonderem Maße begeistern (Projekttitle: „Der Einfluss des Image von Mathematik und Naturwissenschaften auf die schulische Interessen- und Leistungsentwicklung“). Wir gehen davon aus, dass das individuelle Interesse an Mathematik und Naturwissenschaften nicht nur von idiosynkratischen Bedingungskonstellationen – wie z.B. persönlichem Leistungsniveau, Selbstkonzept eigener Fähigkeiten, Einstellungen und Motivation – abhängig ist, sondern auch von sozial geteiltem Wissen (Image), genauer von sog. Stereotypen oder Prototypen, die in der deutschen Schule mit diesen Disziplinen assoziiert werden. Im „BIQUA-Rahmenmodell schulischer und außerschulischer Lernumwelten“ (vgl. Abb. 4 in Doll/Prenzel in diesem Heft) könnte diese Herangehensweise am besten durch ein normatives/kulturelles Makrosystem gekennzeichnet werden, das die dort genannten Lernumwelten wie „Gleichaltrigengruppe“, „Medienumwelt“, „Lehrermerkmale“ und „Unterrichtsprozesse“ beeinflusst (für ähnliche Ansätze siehe auch Boehnke 1996; Noack u.a. 2001; Wild/Hofer 1999).

Unter einem Stereotyp versteht man sozial geteilte Annahmen über Persönlichkeitsmerkmale und Verhaltensweisen der Mitglieder einer sozialen Gruppe (z.B. Stroebe/Hewstone/Stephenson 1996). Der Begriff des Prototypen meint ganz ähnlich den typischen oder durchschnittlichen Repräsentanten einer kognitiven Kategorie, der aus verschiedenen Beispielen extrahiert worden ist (Rosch 1978). Das Stereotyp oder der Prototyp³ über ein Schulfach ist entsprechend eine sozial geteilte Vorstellung über seinen Gegenstandsbereich, über typische Unterrichtsskripts, typische Lehrpersonen oder typische Schüler, die das Fach mögen oder ablehnen.

In einem ersten Schritt haben wir verschiedene explorative Studien durchgeführt, die darauf abzielten, das Stereotyp über Mathematik und Naturwissenschaften zu beschreiben (Kessels/Hannover in Vorbereitung). In einer Untersuchung haben wir Schüler und Schülerinnen der Oberstufe gebeten, in freier Beschreibung die Fächer Mathematik, Deutsch, Physik und Gemeinschaftskunde zu charakterisieren. In einer weiteren

3 Der Begriff des Stereotyps wird in der Regel auf Personengruppen angewendet, während ein Prototyp auch andere Objekte beschreiben kann. Konzeptuelle Unterschiede zwischen den Begrifflichkeiten sind im hiesigen Zusammenhang irrelevant; deshalb verwenden wir die Begriffe Stereotyp und Prototyp synonym.

Studie haben wir Lehramtsstudierende die Fächer Mathematik, Physik, Deutsch und Französisch anhand eines Semantischen Differenzials einschätzen lassen. Die Ergebnisse zeigten, dass folgende Aspekte für die differenzielle Beschreibung mathematisch-naturwissenschaftlicher und sprachlich-geisteswissenschaftlicher Fächer wesentlich sind:

- a) Die Fächer Mathematik und Physik werden als schwieriger wahrgenommen als die geisteswissenschaftlichen Fächer Deutsch und Gemeinschaftskunde (*Schwierigkeit*).
- b) In Physik und vor allem in Mathematik gezeigte Leistungen bzw. Testergebnisse werden als stärker diagnostisch für die zugrunde liegende Intelligenz der Schülerinnen und Schüler angesehen als dies in den geisteswissenschaftlichen Fächern der Fall ist (*Fähigkeitsdiagnostizität*).
- c) Physik- und Mathematikunterricht bieten weniger Möglichkeiten zu freier Gestaltung – und damit zum Ausdruck der eigenen Persönlichkeit oder Identität der Lernenden (*Selbstaffirmation*) – und gehen weniger wahrscheinlich mit sinnlich positiven Erlebnissen (kinästhetisch, haptisch, auditiv, visuell, olfaktorisch, geschmacklich) einher (*Sinnlichkeit*).

In der Studie, in der Lehramtsstudierende die Schulfächer auf einem Semantischen Differenzial eingeschätzt haben, zeigte sich darüber hinausgehend, dass Physik und Mathematik als stärker maskulin konnotiert wahrgenommen werden als die Fächer Deutsch und Französisch (*Maskulinität*). In den freien Beschreibungen der Oberstufenschüler/-innen hingegen wurde die geschlechtsspezifische Konnotation der Schulfächer spontan unerwartet selten thematisiert.

Über die Deskription hinausgehend haben wir in einer Studie die zentrale Annahme geprüft, dass sich ein negatives Stereotyp abträglich auf die individuelle Interessen- und Leistungsentwicklung in der entsprechenden Fachdisziplin auswirkt (Hannover/Kessels eingereicht). Die Wirkung von Stereotypen auf das Individuum verstehen wir dabei im Rahmen unseres generellen Ansatzes, in dem wir die schulische Entwicklung im Zusammenhang mit der Identitätsentwicklung von Kindern und Jugendlichen interpretieren (Hannover 1998, 2002, im Druck). Eine zentrale Entwicklungsaufgabe, die das Individuum über die Lebensspanne hinweg, in besonderem Maße jedoch während der Adoleszenz, zu bewältigen hat, besteht darin herauszufinden, wer es selbst ist und diese Definition der eigenen Person der Außenwelt zu kommunizieren. Jugendliche nutzen Angebote aus ihrer Umwelt zu diesem Zweck; nämlich um Antwort auf die Fragen zu finden: „Wer bin ich selbst?“, „Wer möchte ich gerne sein/werden?“, oder: „Wie sollen andere mich sehen?“ (vgl. auch Fuhrer/Josephs 1999; Habermas 1999; Noack 1990).

Angewendet auf unsere Frage nach den Ursachen der Interessen- und Leistungsdefizite von Schülerinnen und Schülern betrachten wir die Schule als ein Angebot der Umwelt, das das Kind oder der Jugendliche nutzen kann, um sich selbst zu definieren und diese Definition gegenüber anderen zu demonstrieren. In der Schule kann der Lernende die eigene Person darüber definieren, für welche Fächer er Interesse und Selbstvertrauen entwickelt und darüber, welche Fähigkeiten er entwickelt. Gleichzeitig kann er seine Selbstdefinition auch dadurch etablieren und aufrechterhalten, dass er an bestimmten Fächern Desinteresse bekundet und dies durch mangelnde Anstrengung, die in mangel-

haften Leistungen resultiert, noch unterstreicht. Aus diesen Überlegungen ergibt sich unsere Annahme, dass Kinder und Jugendliche Interesse oder Leistungsbereitschaft nur in solchen Unterrichtsfächern entwickeln, die sie als relevant für die Definition der Person erachten, die sie gerne sein wollen und nach außen kommunizieren wollen.

Wie aber gehen Schüler/innen nun vor, wenn sie herausfinden wollen, welche Schulfächer sich dazu eignen, angestrebte Selbstdefinitionen zu erreichen? Wir nehmen an, dass sie zu diesem Zweck das Stereotyp über verschiedene Schulfächer oder aber prototypische Schüler/innen, die die verschiedenen Fächer präferieren, zu ihrem Selbst in Beziehung setzen. Immer dann, wenn der Lernende in der Schule eine Entscheidung trifft – also z.B. welches Fach er als persönliches Lieblingsfach betrachtet, als Arbeitsschwerpunkt oder Leistungskurs wählt – nutzt er den sog. „Selbst-Prototypen-Abgleich“ (vgl. Niedenthal/Cantor/Kihlstrom 1985; Setterlund/Niedenthal 1993) als eine Art Entscheidungsheuristik. Dazu stellt er sich für jede der möglichen Entscheidungen die prototypische Person vor, die diese Option wählen würde; also z.B. die prototypischen Schüler/innen, die Physik, Französisch, Mathe usw. als Leistungskurs wählen würden. Nun überprüft die Person das Ausmaß der Übereinstimmung jedes einzelnen Prototypen mit dem Bild, das sie von sich selbst hat. Je größer die wahrgenommene Ähnlichkeit zwischen dem Prototypen und ihrem Selbst ist, desto wahrscheinlicher wird sich die Person den Interessensgebieten des prototypischen Repräsentanten dieser Gebiete zuwenden.

Voraussetzung dafür, dass ein Lernender diese Entscheidungsheuristik verwenden kann, ist allerdings, dass (a) es klare Prototypen über die betreffenden Schulfächer gibt und (b) die Person ein klares Bild davon hat, wer sie selbst ist. Campbell (1990; Campbell u.a. 1996) hat Selbstklarheit als das Ausmaß definiert, in dem ein Individuum persönliche Eigenschaften, Präferenzen, Fähigkeiten usw. sicher, konsistent und zeitstabil definieren kann. In unserer eigenen Untersuchung haben wir nur für Schüler/innen, die eine hinreichend hohe Selbstklarheit haben, erwartet, dass sie in Schulfächern, über die ein Prototyp besteht, das Selbst-Prototypen-Abgleichen als Entscheidungs-Heuristik verwenden.

Zusammengefasst gehen wir in unserem Forschungsvorhaben davon aus, dass schulisches Verhalten aus der Sicht des Lernenden funktional für die Entwicklung seiner Identität ist. Dies kann beispielsweise bedeuten, dass sich ein Lernender für ein bestimmtes Schulfach interessiert oder auch gerade nicht interessiert, weil er sich dadurch selbst als eine Person definieren kann, die die entsprechenden Vorlieben bzw. Abneigungen hat. Entsprechend entfaltet ein negatives Stereotyp oder ein negativer Prototyp über Mathematik und Naturwissenschaften seine negative Wirkung vermittelt über das Selbstbild oder die Identität des Lernenden: nämlich dadurch, dass der Lernende das Stereotyp oder den Prototypen zu seinem Selbstbild in Beziehung setzt. Erste Belege für diese Annahmen stammen aus einer Studie von Hannover/Kessels (eingereicht). In einer Stichprobe von Achtklässlern fanden wir hohe Übereinstimmungen in den Vorstellungen, wie „prototypische Schüler/innen“ sind, die Physik als Lieblingsfach haben. Der Physik-Prototyp zeigte von allen Schulfachprototypen im Mittel die geringste Übereinstimmung mit dem Selbstbild der befragten Schülerinnen und Schüler.

3. Veränderung des Stereotyps oder des Prototyps über die naturwissenschaftlichen Schulfächer

Das Schwerpunktprogramm BIQUA zielt darauf ab, langfristig Interventionen für Schule und Unterricht zu entwickeln, durch die die Interessen- und Leistungsentwicklung in der Mathematik und den Naturwissenschaften verbessert werden kann. Bezogen auf das Thema unseres Forschungsprojekts bedeutet dies, dass zu prüfen ist, ob und wenn ja wie Stereotype und Prototypen von Schulfächern geändert werden können.

Zahlreiche Studien haben belegt (z.B. Brewer 1996; Hamilton/Trolier 1986; Tajfel/Wilkes 1963), dass Stereotype und Prototypen auf einem kognitiven Kategorisierungsprozess basieren, als dessen Ergebnis Unterschiede zwischen konkreten Vertretern verschiedener Kategorien (also z.B. zwischen dem Schüler Tom, der Physik als Lieblingsfach hat, und dem Schüler Max, der Deutsch als Lieblingsfach hat) überschätzt und die Verschiedenheit von Vertretern innerhalb der betreffenden Kategorien (also z.B. zwischen Physik-Liebhaber Tom und Physik-Liebhaber Jürgen) unterschätzt werden. Im Ergebnis werden Repräsentanten einer stereotypisierten Gruppe oder eines Prototyps nicht mehr als einzigartige Individuen wahrgenommen, sondern als anonyme Repräsentanten der betreffenden sozialen Kategorie. So werden beispielsweise idiosynkratische Merkmale des Physik-Liebhabers Tom (z.B. seine Leidenschaft für Bassgitarre) in dem Maße ignoriert, wie ihm Merkmale des Physik-Prototypen zugeschrieben werden (z.B. unsinnlich zu sein); ein Phänomen, das als Depersonalisierung bezeichnet wird (z.B. Tajfel/Turner 1986).

Vor dem Hintergrund dieser Erkenntnisse über die Funktionsweise von Stereotypen ist angenommen worden, dass Stereotype durch individualisierten Kontakt mit Mitgliedern stereotypisierter Gruppen abgebaut werden können (z.B. Aronson/Bridgeman 1979). Dabei ist allerdings entscheidend, dass das Mitglied der stereotypisierten Gruppe als einzigartiges Individuum – und nicht als anonymen Vertreter der stereotypisierten Gruppe – wahrgenommen wird. Durch eine solche Personalisierung wird die mit der Stereotypisierung einhergehende Depersonalisierung aufgehoben (Brewer/Miller 1988, 1996; Hewstone 1996). Entsprechend haben wir angenommen, dass der individualisierte Kontakt mit Personen, die zu Gruppen gehören, auf die das negative Stereotyp oder der negative Prototyp über die Naturwissenschaften angewendet wird, zu einer Abschwächung des Stereotyps oder Prototyps führen würde.

4. Ziele der vorliegenden Studien

Im Folgenden werden zwei Studien geschildert, die darauf abzielen,

- a) die uneindeutigen Ergebnisse, die unsere explorativen Untersuchungen zum Stereotyp/Prototyp (v.a. zur Geschlechtskonnotation) verschiedener Schulfächer ergeben haben, aufzuhellen und (Studie 1, Abschnitt 5.1)
- b) erste Evidenz zu der Frage beizubringen, ob Stereotype/Prototyp über Schulfächer verändert werden können (Studie 2, Abschnitt 5.2).

In unseren explorativen Vorstudien hatten sich uneinheitliche Ergebnisse bezüglich der Frage ergeben, ob Geschlechtskonnotation eine relevante Dimension des Stereotyps über verschiedene Schulfächer darstellt. Während die von uns befragten Lehramtsstudierenden naturwissenschaftliche Fächer als stärker maskulin konnotiert eingeschätzt hatten als sprachliche Fächer, tauchten in den freien Beschreibungen der Schulfächer durch Oberstufenschülerinnen und -schüler nur selten Verweise auf Geschlechtskonnotationen auf. Diese Befunde können auf unterschiedliche Weise interpretiert werden. Sie können darauf verweisen, dass Physik und Mathematik von Schülerinnen und Schülern mehrheitlich nicht als männliche Domäne angesehen werden. Sie können aber auch bedeuten, dass die Geschlechtstypisierung der Fächer als so selbstverständlich angesehen wird, dass sie in den freien Beschreibungen noch nicht einmal erwähnt wurde oder schließlich, dass die Geschlechtstypisierung der Fächer von den Schülerinnen und Schülern nicht bewusst wahrgenommen und entsprechend auch nicht artikuliert wird.

Um zwischen diesen verschiedenen Interpretationsmöglichkeiten entscheiden zu können, haben wir eine (Studie 1, Abschnitt 5.1) durchgeführt, in der eine andere Antwortmodalität als in der Explorationsstudie verwendet wurde: Die befragten Schülerinnen und Schüler wurden direkt aufgefordert, vorgegebene Schulfächer danach zu klassifizieren, ob sie Mädchenfächer, Jungenfächer oder aber weder Mädchen- noch Jungenfächer (also nicht geschlechtskonnotiert) sind.

In der zweiten (Studie 2, Abschnitt 5.2) soll erste Evidenz für die Frage gewonnen werden, ob Stereotype/Prototypen über Schulfächer durch individualisierten Kontakt mit Personen, die Mitglieder von Gruppen sind, auf die das Stereotyp/der Prototyp angewendet wird, verändert werden können. Dazu haben wir Jugendliche untersucht, die im Herbst 2001 an den so genannten „Schülerinnen&Schüler-Technik-Tagen der TU Berlin“ teilgenommen haben. Hier wurden in den Schulferien über vier Tage hinweg für Schülerinnen und Schüler ab der 9. Klasse praktische Projekte in den Laboren, Instituten und Versuchshallen der Technischen Universität Berlin angeboten. Die Jugendlichen konnten aus insgesamt 58 Projekten wählen, deren Gemeinsamkeit darin bestand, dass a) Naturwissenschaften und Technik durch praktische Tätigkeiten direkt erfahren wurden und b) Studienmöglichkeiten und Berufsfelder im Bereich Naturwissenschaften und Technik vorgestellt wurden. Die Organisatoren erhofften sich durch dieses Programm, die Leistungskurswahlen der Schülerinnen und Schüler in Richtung auf Mathematik, Informatik, Physik, Chemie oder Biologie zu lenken und damit die Zahl potenziell für die Aufnahme eines Hochschulstudiums im Bereich Naturwissenschaften und Technik geeigneter Schulabgänger zu erhöhen; wobei Mädchen besonders stark angesprochen und gefördert werden sollten.

Sämtliche Projekte wurden von (einer oder mehreren) im Bereich Naturwissenschaften und Technik tätigen Personen geleitet. Alle Projektleiter/innen haben sich den teilnehmenden Jugendlichen zu Beginn des Kurses vorgestellt und dabei auch ihre berufliche Tätigkeit erwähnt. Entsprechend kann angenommen werden, dass der Kursleiter/die Kursleiterin als individualisierter Repräsentant der Gruppe der „im Bereich von Naturwissenschaften und Technik Berufstätigen“ – die gemäß unseren Voruntersuchungsergebnissen negativ stereotypisiert sein sollte – wahrgenommen wurde.

Die Teilnehmerinnen und Teilnehmer wurden zu ihren Prototypen über das Schulfach Physik befragt. Weiter wurde der Altersgruppe (Oberstufenschülerinnen und -schüler) entsprechend ein berufsbezogener Prototyp erfasst, nämlich der Prototyp über Ingenieure/Ingenieurinnen, der als besonders repräsentativ für akademische Berufe im Bereich von Naturwissenschaften und Technik gelten kann. Um den Grad der Übereinstimmung zwischen den Prototypen und dem Selbstbild der Jugendlichen bestimmen zu können, wurde das Selbstbild jedes Jugendlichen erfasst. Mit der Evaluation wurden drei Ziele verfolgt:

- a) Es sollte beschrieben werden, wie Prototypen über Physik und den Ingenieursberuf in einer Stichprobe von Schülerinnen und Schülern aussehen, die eine positive Voreinstellung zu Technik und Naturwissenschaften mitbringen (freiwillige Teilnahme an den Techniktagen).
- b) Es sollte geprüft werden, inwieweit sich die schulische Interessenentwicklung aus dem Grad der Übereinstimmung zwischen Selbst und relevanten Prototypen vorhersagen lässt. Wir haben erwartet, dass Schüler/innen mit hoher Selbstklarheit in dem Maße stärkere Präferenzen gegenüber der Physik zeigen, in dem die Übereinstimmung zwischen ihrem Selbstbild und dem Prototypen über Physik und dem Prototypen über den Ingenieursberuf zunimmt. Für Schüler/innen mit geringer Selbstklarheit hingegen sollte sich kein Zusammenhang zwischen schulischen Präferenzen gegenüber der Physik auf der einen Seite und dem Ausmaß der Übereinstimmung zwischen Selbst und relevanten Prototypen auf der anderen Seite zeigen.
- c) Es sollte geprüft werden, ob sich durch den persönlichen Kontakt mit einer im Bereich Naturwissenschaften tätigen Person (Kursleiter/in) die Prototypen der Jugendlichen über Physik und den Ingenieursberuf dahingehend verändern, dass die Distanz zwischen ihrem Selbstbild und den Prototypen geringer ausfällt.

5. Methode

5.1 Studie 1

Stichprobe

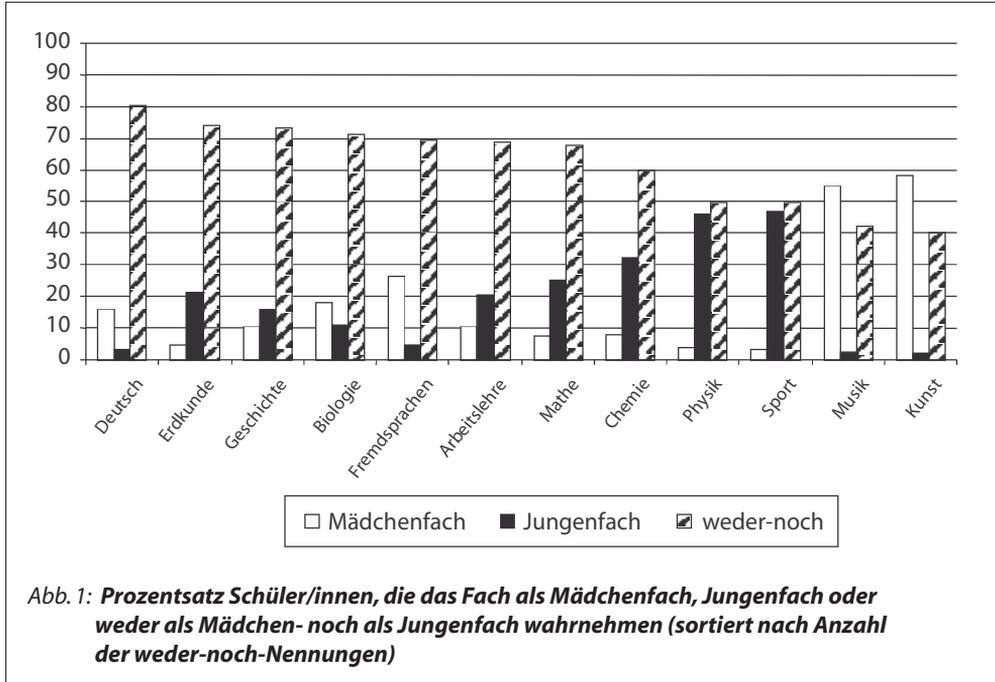
An der Befragung beteiligten sich die Schülerinnen und Schüler aus den neunten Jahrgängen von fünf Berliner Gesamtschulen. Insgesamt wurden $N = 631$ Jugendliche befragt (48% Mädchen, 52% Jungen). Das mittlere Alter betrug 15,3 Jahre ($s = 0,57$).

Messinstrumente

Den Schülerinnen und Schülern wurden während einer Schulstunde umfangreiche Fragebögen zum Physik- und Chemieunterricht vorgelegt. Die für unsere Studie relevante Frage bezog sich auf die Geschlechtstypisierung der an der Gesamtschule unterrichteten Fächer Deutsch, Biologie, Physik, Fremdsprachen, Chemie, Mathematik, Kunst, Geschichte/Sozialkunde, Musik, Erdkunde, Sport und Arbeitslehre. Sie lautete: „Welche der

folgenden Schulfächer sind deiner Meinung nach eher ‚Jungenfächer‘ und welche eher ‚Mädchenfächer‘?‘ (dreistufiges Antwortformat: ‚Jungenfach‘, ‚Mädchenfach‘, ‚weder-noch‘).

Ergebnisse



In Abbildung 1 sind die Ergebnisse für sämtliche erfassten Schulfächer dargestellt. Dabei fällt auf, dass die hier im Fokus stehenden naturwissenschaftlichen Fächer in besonderem Maße als „Jungenfächer“ galten. 25% der Befragten äußerten, dass Mathematik ein Jungenfach sei, 32% bezeichneten Chemie als Jungenfach und sogar 46% fanden dies für Physik zutreffend. Eine vergleichbare Geschlechtstypisierung zeigte sich nur noch bei den Fächern Musik und Kunst (55% bzw. 58% nannten diese „Mädchenfächer“) und Sport (47%: „Jungenfach“). Fächer wie Deutsch und Fremdsprachen wurden in geringerem Maße als geschlechtstypisiert angesehen als die naturwissenschaftlichen Fächer (Deutsch: 81% „weder-noch“; Fremdsprachen: 69% „weder-noch“).

Getrennte Analysen für männliche und weibliche Befragte zeigten weiterhin, dass Mädchen das Fach Physik anders als Jungen beurteilten, $\chi^2 = 9,44$; $df = 2$; $p < .01$; sie bezeichneten Physik häufiger als „Jungenfach“ (52%) als Jungen (41%). Auch bei der Beurteilung des Faches Chemie unterschieden sich Jungen und Mädchen, $\chi^2 = 8,91$; $df = 2$; $p < .05$; 31% der Mädchen und 34% der Jungen hielten es für ein „Jungenfach“ und 11% der Mädchen und 5% der Jungen für ein „Mädchenfach“. Biologie wurde ebenfalls unterschiedlich beurteilt, $\chi^2 = 10,56$; $df = 2$; $p < .01$, Mädchen bezeichneten es öfter als

„Mädchenfach“ (22%) als Jungen (15%) und Jungen nannten Biologie häufiger ein „Jungenfach“ (14%) als Mädchen (7%). Auch Deutsch wurde unterschiedlich beurteilt, $\chi^2 = 8,18$; $df = 2$; $p < .05$; Mädchen bezeichneten es häufiger als „Mädchenfach“ (20%) als Jungen (12%). Sport wurde dagegen häufiger von Jungen als „Jungenfach“ bezeichnet (58%) als von Mädchen (34%), für ein nicht geschlechtskonnotiertes Fach hielten es 60% der Mädchen, aber nur 40% der Jungen, $\chi^2 = 41,45$; $df = 2$; $p < .001$. Für alle anderen Fächern ergaben sich keine bedeutsamen Unterschiede in der Kategorisierung in Abhängigkeit vom Geschlecht der Befragten.

5.2 Studie 2

Stichprobe

Vor Beginn der „Schülerinnen&Schüler-Technik-Tage der TU Berlin“ (STT) wurden Fragebögen per Post an die angemeldeten Teilnehmer/innen verschickt. Die ausgefüllten Bögen wurden von den Jugendlichen zu Beginn der STT mitgebracht. An der Vorher-Befragung beteiligten sich insgesamt 183 Jugendliche (94 Mädchen, 87 Jungen). Zwecks Nachher-Befragung wurden Fragebögen jeweils am Ende eines Projektes, an dem die Jugendlichen teilgenommen hatten, ausgegeben und nach dem Ausfüllen von den Jugendlichen bei einer Sammelstelle abgegeben. Von der Nachher-Befragung liegen Bögen von 104 Jugendlichen (60 Mädchen, 44 Jungen) vor. 85 Jugendliche nahmen an beiden Befragungen teil (49 Mädchen, 36 Jungen). Im Mittel waren die Teilnehmer/innen 16.4 Jahre alt ($s = 1.7$). 62% wohnten in Berlin, 38% in Brandenburg. Die meisten besuchten ein Gymnasium (86%), 10% eine Gesamtschule. 43% waren entweder in der 9. oder 10. Klasse und 55% in der Oberstufe (Klasse 11 bis 13). Drei Personen wurden aus allen Auswertungen ausgeschlossen, da sie offensichtlich die Teilnahme durch unvollständige Antworten oder „Durchkreuzen“ boykottiert hatten.

Messinstrumente

Zum ersten und zweiten Messzeitpunkt wurde der *Prototyp über den Schüler* (für männliche Teilnehmer) bzw. *über die Schülerin* (für weibliche Teilnehmer), *der/die Physik als Lieblingsfach hat* und der *Prototyp des Ingenieurs* (für männliche Teilnehmer) bzw. *der Ingenieurin* (für weibliche Teilnehmer) erfasst⁴. Dazu wurde den Jugendlichen jeweils eine Liste von 40 Adjektiven in alphabetischer Reihenfolge vorgegeben. 25 dieser Adjektive entstammten einem bereits erprobten Erhebungsinstrument für Prototypen (Hannover/Kessels eingereicht). Die Adjektive sind Abbildung 1 zu entnehmen. Sie bilden die in der zitierten Untersuchung faktorenanalytisch identifizierten fünf Dimensionen „Soziale und physische Attraktivität“ (schön, attraktiv, begehrt, beliebt, angesehen) „Soziale

4 Die Befragten sollten jeweils nur Personen ihres eigenen Geschlechts beschreiben, weil sonst die Beurteilung der Ähnlichkeit zwischen Selbst und Prototyp vermutlich stärker durch die (andere) Geschlechtszugehörigkeit als durch die Berufsbezeichnung der prototypischen Person bestimmt worden wäre.

Integration und soziale Kompetenz“ (kontaktfreudig, verklemmt, langweilig, einsam, aufgeschlossen), „Selbstbezogenheit und Arroganz“ (besserwisserisch, wichtigtuerisch, arrogant, streberhaft, selbstbezogen), „Intelligenz und Motivation“ (klug, intelligent, gebildet, ehrgeizig, fleißig), „Kreativität und Emotionalität“ (kreativ, fantasievoll, einfühlsam, redegewandt, romantisch) ab. Sie wurden ergänzt durch fünf maskuline (fühlt sich überlegen, stark, kraftvoll, hartnäckig, unerschrocken) und fünf feminine Adjektive (gefühlbetont, sanft, ordentlich, weichherzig, feinfühlig) aus einem Instrument von Kessels (2002) sowie durch fünf weitere, die besonders repräsentativ bzw. atypisch für das Bild über Ingenieur/inn/en sind (kann gut organisieren, kulturell interessiert, logisch denkend, vernarrt in Computer, hat gute Menschenkenntnis).

Die Jugendlichen sollten zunächst beurteilen, wie stark jedes der 40 Adjektive auf „das typische Mädchen mit dem Lieblingsfach Physik“ (weibliche Teilnehmer) bzw. auf „den typischen Jungen mit dem Lieblingsfach Physik“ (männliche Teilnehmer) zutrifft. Im Anschluss beurteilten die weiblichen Teilnehmer, wie stark die Adjektive „die typische Ingenieurin“ und die männlichen Teilnehmer, wie stark sie den „typischen Ingenieur“ beschreiben. Abschließend wurden die Jugendlichen zu beiden Messzeitpunkten gebeten, sich anhand der Adjektivliste selbst zu beschreiben (*Selbstbild*) (siebenstufige Antwortskalen von 1 = „trifft überhaupt nicht zu“ bis 7 = „trifft völlig zu“).

Die *Selbstklarheit* wurde beim ersten Messzeitpunkt mit dem Fragebogen von Campbell u.a. (1996, eigene Übersetzung) erhoben, der aus 12 Items besteht (wie z.B. „Meine Meinung über mich selbst ändert sich von einem Tag zum anderen“ und „Im Großen und Ganzen habe ich eine klare Vorstellung davon, wer und was ich bin (-)“, Antwortskala 1 = „trifft völlig zu“ 5 = „trifft überhaupt nicht zu“ Chronbach's alpha = .76; ein hoher Wert bedeutet hohe Selbstklarheit).

Um prüfen zu können, ob das Ausmaß der Übereinstimmung zwischen Selbstbild und relevanten Prototypen prädiktiv für die schulische Interessenentwicklung ist, wurden *Präferenzen gegenüber der Physik* erfasst. Dazu sollten die Jugendlichen a) eine Liste von 11 Schulfächern danach sortieren, wie gern sie sie mögen (je höher der *Rangplatz von Physik* [d.h. je kleiner der Wert, z.B. 1., 2. oder 3. Platz], desto lieber wurde es gemocht) und b) eine aus 3 Items bestehende *Einstellungsskala zur Physik* beantworten („Ich werde Physik als Leistungskurs wählen“; „Ich würde Physik, wenn es ginge, abwählen (-)“; „In einem Physik-Leistungskurs könnte ich gute Noten erzielen“; fünfstufiges Antwortformat 1 = „keinesfalls“ bis 5 = „ganz sicher“; Chronbach's alpha = .88).

Ergebnisse

Selbstbeschreibung und Beschreibung der Prototypen für Physik und den Ingenieursberuf zum ersten Messzeitpunkt

In Abbildung 2 sind die Selbstbeschreibungen sowie die Beschreibungen der Prototypen für Physik und Ingenieur/innen auf den einzelnen Adjektiven dargestellt. Diese Darstellung hat rein deskriptiven Charakter; von einer statischen Testung auf Einzelitem-Ebene wurde abgesehen. Zunächst zeigt sich, dass die verschiedenen Adjektive relativ unab-

hängig von der zu beschreibenden Zielperson unterschiedlich stark bejaht bzw. zurückgewiesen wurden: Die Kurven weisen einen ähnlichen Verlauf auf. Allerdings wurde der Physik-Prototyp im Vergleich zu dem Bild, das die Jugendlichen von sich selbst haben, als weniger schön, attraktiv, begehrt, angesehen, kontaktfreudig, kreativ, stark, kraftvoll und mit weniger Menschenkenntnis ausgestattet angesehen. Gleichzeitig galt er als stärker verklemmt, langweilig und einsam. Dem Physikprototypen wie auch dem Ingenieurprototypen wurden außerdem besonders wenige feminine Eigenschaften (z.B. gefühlsbetont, sanft, weichherzig) bescheinigt. Eigenschaften, die auf Intelligenz, Leistungsorientierung und Professionalität verweisen, wurden den Prototypen hingegen teilweise sogar stärker zugeschrieben als der eigenen Person; z.B. klug, intelligent, ehrgeizig, kann gut organisieren.

Präferenzen gegenüber der Physik

		schulische Wahl Physik ¹	Rangplatz Physik ²
Mädchen	Mittelwert	2,84	6,53
	s	1,25	3,51
Jungen	Mittelwert	3,93	3,45
	s	0,96	2,64
gesamt	Mittelwert	3,36	5,08
	s	1,25	3,50
t-Test Mädchen/ Jungen	t	-6,53	6,43
	p	.000	.000

¹ Skala von 1 (keinesfalls) bis 5 (ganz sicher)
² mittlerer Rangplatz, als 11 Schulfächer in eine Rangfolge gebracht wurden

Die Analyse der individuellen Schulfach-Beliebtheits-Ranglisten ergab, dass Physik bei den männlichen Teilnehmern deutlich beliebter war als bei den weiblichen: Wie Tabelle 1 zeigt, platzierten die Jungen das Fach Physik auf der persönlichen Schulfach-Rangliste deutlich höher als die Mädchen, $t_{(175)} = 6,57$, $p < .001$. Weiter zeigt Tabelle 1 eine positivere Einstellung gegenüber der Physik für Jungen als für Mädchen, $t_{(177)} = -6,53$, $p < .001$. Die insgesamt hohen Werte illustrieren, dass die hier untersuchte Stichprobe dem Fach Physik gegenüber deutlich positiv eingestellt ist.

Selbst-Prototypen-Abgleich

Wir haben angenommen, dass unter der Bedingung, dass ein Jugendlicher ein klares Bild seines Selbst hat, er das Fach Physik in dem Maße präferiert, wie sein Selbstbild seinen Prototypen über Physik und über den Ingenieursberuf ähnlich ist. Um diese Annahme zu prüfen, wurde die Stichprobe per Mediansplit in Jugendliche mit niedriger

($N = 92$; Selbstklarheit kleiner als 3,17) und hoher Selbstklarheit ($N = 79$; Selbstklarheit gleich oder größer als 3,17) eingeteilt. Des Weiteren wurden für jede Person die individuelle Distanz zwischen Selbstbild und Physikprototyp bzw. Ingenieurprototyp berechnet. Genauer wurde für jedes Adjektiv die absolute Differenz zwischen der Beschreibung des jeweiligen Prototypen und der Selbstbeschreibung berechnet, aufsummiert und durch 40 dividiert (*Selbst-Prototypen-Distanzscore*). Je größer der Score, desto unähnlicher waren sich Selbstbeschreibung und Physik- bzw. Ingenieurprototyp. Dieser Score kann theoretisch von 0 (Selbstbeschreibung und Beschreibung Prototyp identisch) bis 6 (Selbstbeschreibung und Beschreibung Prototyp maximal verschieden) reichen (die empirischen Mittelwerte werden weiter unten berichtet). Nun wurden separat Rangkorrelationen zwischen den beiden Physik-Präferenzmaßen auf der einen Seite und den beiden Selbst-Prototypen-Distanzscores auf der anderen Seite berechnet. Folgendes Muster wurde erwartet. Je kleiner der Distanzscore für die Physik bzw. für den Ingenieursberuf ist, a) um so höher sollte der Rangplatz der Physik in der individuellen Präferenzliste sein und b) um so positiver sollte die Einstellung gegenüber der Physik sein. Die resultierenden Korrelationskoeffizienten sind in Tabelle 2 dargestellt.

Für die Gruppe der Jugendlichen mit hoher Selbstklarheit zeigten sich folgende Zusammenhänge: Je ähnlicher sie dem Prototypen des Schülers/der Schülerin mit dem Lieblingsfach Physik waren, desto höher stand das Fach Physik auf ihrer Schulfach-Rangliste, $r = .56$ ($p < .001$), und desto positiver war ihre Einstellung gegenüber der Physik, $r = -.59$ ($p < .001$). Zwischen der Nähe zum Ingenieur/innen-Prototyp und dem Selbstbild zeigten sich ähnliche, wenn auch weniger starke Zusammenhänge. Um so ähnlicher ein Jugendlicher sich selbst dem/der typischen Ingenieur/in einschätzte, desto höher hatte er das Fach Physik auf der Rangliste platziert, $r = .32$ ($p < .01$) und desto positiver war seine Einstellung gegenüber der Physik, $r = -.40$ ($p < .001$).

Vergleichbare Zusammenhänge ließen sich bei der Gruppe der Jugendlichen mit gering ausgeprägter Selbstklarheit nicht finden: Zwar korrelierte auch für sie die Distanz zwischen Selbst und Physik-Prototyp signifikant mit dem Rangplatz des Faches Physik, $r = .32$ ($p < .01$) und der Einstellung gegenüber der Physik, $r = -.22$ ($p < .05$), allerdings deutlich niedriger als bei der Gruppe mit hoher Selbstklarheit. Der Distanzscore Selbst-Ingenieurprototyp war sogar völlig unabhängig von dem Rangplatz des Faches Physik ($r = .05$, *ns*) und der Einstellung gegenüber der Physik, $r = -.06$ (*ns*) (siehe Tabelle 1). Die Korrelationen in der Gruppe mit niedriger Selbstklarheit unterschieden sich signifikant von den Korrelationen in der Gruppe mit hoher Selbstklarheit (Physik-Prototyp/Rangplatz Physik: $z = 1.87$; $p < .05$; Physik-Prototyp/schulische Wahl Physik: $z = -2.82$; $p < .01$; Ingenieurprototyp/Rangplatz Physik: $z = 1.75$; $p < .05$; Ingenieurprototyp/schulische Wahl Physik: $z = -2.26$; $p < .05$).

Tab.2: **Rangkorrelationen zwischen dem Selbst-Prototypen-Distanzscore und Beliebtheit von Physik bei hoher und niedriger Selbstklarheit**

		Rangplatz Physik		schulische Wahl Physik	
Distanzscore Selbst-Physikprototyp	hohe Selbstklarheit	.56	***	-.59	***
	niedrige Selbstklarheit	.32	**	-.22	*
Distanzscore Selbst-Ingenieurprototyp	hohe Selbstklarheit	.32	**	-.40	***
	niedrige Selbstklarheit	.05		-.06	

* = $p < .05$; ** = $p < .01$; *** = $p < .001$

Veränderung der Distanzscores

Schließlich sollte geprüft werden, ob durch die Teilnahme an den STT die Prototypen beeinflusst wurden. Eine deskriptive Analyse der Prototypen zum zweiten Messzeitpunkt ergab ein ähnliches Verlaufmuster der Kurven wie zum ersten Messzeitpunkt: Gegenüber dem in Abbildung 2 dargestellten Muster ergab sich keine Verschiebung in der relativen Position zwischen den drei zu beurteilenden Zielpersonen (Physik-Prototyp, Ingenieurs-Prototyp, Selbst).

Um zu prüfen, ob sich die absolute Entfernung zwischen Prototypen und Selbstbild über die Messzeitpunkte hinweg verringert hat, wurden die Distanzscores zum Prototypen Physik und Ingenieur/in zwei separaten 2 (Geschlecht: männlich/weiblich) x 2 (Messung vor/nach den STT) -faktoriellen Varianzanalysen mit Messwiederholung unterworfen.

Für die physikbezogenen Distanzscores zeigte sich erwartungsgemäß, dass die wahrgenommene Unterschiedlichkeit von Selbstbild und dem/der typischen Schüler/in mit Lieblingsfach Physik nach den STT kleiner war ($M = 1,28$; $s = 0,54$) als vor den STT ($M = 1,47$, $s = 0,60$), Haupteffekt des Faktors Messwiederholung, $F_{(1,78)} = 26,34$, $p < .001$. Es zeigte sich kein Haupteffekt des Geschlechts und keine Interaktion.

Auch für den ingenieurbezogenen Distanzscore ergab sich, dass sich die wahrgenommene Unterschiedlichkeit von Selbstbild und dem Bild des/der typischen Ingenieur/in über die Messzeitpunkte hinweg verringerte, Haupteffekt des Faktors Messwiederholung, $F_{(1,74)} = 15,74$, $p < .001$ (vor den STT: $M = 1,39$, $s = 0,50$; nach den STT: $M = 1,20$; $s = 0,52$); für das Geschlecht der Teilnehmer ergaben sich weder ein Haupteffekt noch Interaktionen.

6. Diskussion

Die Ergebnisse unserer ersten Studie zeigen deutlich, dass die Naturwissenschaften von Schülerinnen und Schülern der Sekundarstufe I geschlechtskonnotiert wahrgenommen werden. Insbesondere die Physik wurde von fast der Hälfte der Jugendlichen als Jungenfach eingeschätzt; wobei der Anteil der Schülerinnen hier noch größer war als der der

Schüler. Dies deutet darauf hin, dass Mädchen in besonderem Maße annehmen, für dieses Fach wenig geeignet zu sein. Dieselben Zusammenhänge gelten auch für die Chemie und für die Mathematik, wenn auch in geringerem Ausmaß. Eine vergleichbar starke Kategorisierung als Mädchenfach hatte sich nur für die Fächer Musik und Kunst gezeigt.

Diese Ergebnisse klären die uneindeutigen Befunde der qualitativen Befragungen von Schülerinnen und Schülern auf, in denen die Geschlechtskonnotation von Schulfächern nur sehr selten erwähnt worden war (Kessels/Hannover in Vorbereitung). Es scheint, dass die Zuordnung verschiedener Fächer zur Gruppe der Mädchen oder zur Gruppe der Jungen den Jugendlichen so selbstverständlich erscheint, dass sie die sen Aspekt der Wahrnehmung der Fächer in freier Beschreibung gar nicht erwähnen.

Die Befunde der zweiten im vorliegenden Aufsatz vorgestellten Studie sprechen dafür, dass die Maskulinität bzw. Femininität eines Schulfachs eine bedeutsame Dimension des Prototypen über das betreffende Schulfach darstellt. Hier hatte sich gezeigt, dass sich die erfassten Prototypen über Physik und den Ingenieursberuf insbesondere auf geschlechtskonnotierten, femininen Eigenschaftsdimensionen vom Selbstbild der Jugendlichen unterschieden: Beide Prototypen wurden als weniger feminin eingeschätzt, als die Jugendlichen sich selbst gesehen haben.

Die Beschreibung der Prototypen hatte auch auf den anderen erfassten Eigenschaftsdimensionen Abweichungen vom Selbstbild der Jugendlichen ergeben. Im Vergleich zum Bild, das die Schülerinnen und Schüler von sich selbst haben, wurde insbesondere der Prototyp über Physik auf den Dimensionen „Soziale und physische Attraktivität“ und „Soziale Integration und soziale Kompetenz“ negativer wahrgenommen. Diese Befunde sind um so bedeutsamer, als es sich bei unserer Stichprobe um Schülerinnen und Schülern handelt, die freiwillig während der Schulferien ein Kursangebot zum Thema Naturwissenschaften und Technik in Anspruch nahmen. Dies bedeutet, selbst in einer Gruppe von Jugendlichen, die eine positive Voreinstellung zu Naturwissenschaften und Technik mitbringt, bestehen negative Prototypen über die entsprechenden Schulfächer und Berufe. Dieser Befund ist konsistent mit unserer Annahme, dass Stereotype und Prototypen in einer Kultur sozial geteilt werden und unabhängig von idiosynkratischen Bedingungen der Person ihre Wirkung entfalten.

Dafür sprechen unsere Befunde zum Selbst-Prototypen-Abgleich. Die von uns befragten Jugendlichen richteten ihre schulische Präferenzen und Wahlen danach aus, wie ähnlich oder unähnlich sie ihr Selbst zu prototypischen Personen sahen, die Physik präferieren bzw. den Beruf des Ingenieurs ergreifen. Erwartungsgemäß galt dieser Zusammenhang jedoch nur für solche Teilnehmer/innen, die ein klares Bild davon hatten, wer oder was sie selbst sind. Es scheint, dass diese Jugendlichen schulische Lernangebote zu ihrem Selbst in Beziehung setzen und Entscheidungen, z.B. in welchem Fach sie sich spezialisieren wollen, von der Ähnlichkeit ihres Selbst zum Prototypen abhängig machen, der für eine entsprechende Wahl steht.

Die Ergebnisse zur Evaluation der Schülerinnen&Schüler-Technik-Tage sprechen dafür, dass eine Beeinflussung von Stereotypen und Prototypen über Schulfächer und Berufsbilder möglich ist, indem Lernende in Kontakt mit einem individualisierten Vertreter der stereotypisierten Gruppen treten: Eine Gemeinsamkeit aller Kursangebote be-

stand daran, dass die Jugendlichen mit einer Person interagierten, die sich als im Bereich von Naturwissenschaft und Technik tätig vorgestellt hatte. Nach der Teilnahme an den Kursen war die Diskrepanz zwischen dem Selbstbild der Jugendlichen und den relevanten Prototypen gegenüber dem Vorhermesszeitpunkt reduziert. Dies bedeutet, die Jugendlichen erlebten eine größere Übereinstimmung zwischen dem Bild, das sie von sich selbst haben, und dem Bild, das sie von einem Physik-Liebhaber oder von einem Ingenieur/einer Ingenieurin haben.

Es kann jedoch eingewendet werden, dass ein solcher positiver Effekt der Konfrontation mit einem Mitglied aus der stereotypisierten Gruppe auch von dem Kontakt mit Lehrerinnen und Lehrern der Naturwissenschaften im Schulunterricht ausgehen könnte. Dass der Kontakt mit den Kursleiterinnen und Kursleitern der STT einen nachweisbar positiven Effekt hatte (obwohl die Schülerinnen und Schüler gleichzeitig einem potenziell positiven Einfluss durch schulische Modellpersonen ausgesetzt sind), erklärt sich möglicherweise damit, dass Lehrer/innen nicht in erster Linie als einzigartige Individuen, sondern als Repräsentanten der sozialen Kategorie „Physiklehrer“, „Mathelehrer“ usw. wahrgenommen werden. Zahlreiche Studien zeigen, dass der bloße Kontakt mit einem Mitglied einer stereotypisierten Gruppe nicht ausreichend ist, um das Stereotyp abzuschwächen oder gar aufzulösen. Vielmehr muss ein individualisierter, positiv erlebter Kontakt entstehen (vgl. z.B. Brewer 1996; Brewer/Miller 1988; Hewstone 1996). Dafür sind im Rahmen eines Freizeit-Kursangebots für Freiwillige vermutlich bessere Voraussetzungen geschaffen als im Kontext Schule, in dem Schüler/innen und Lehrer/innen gleichermaßen sich nicht nur als einzigartige Individuen, sondern auch als Vertreter ihrer unterschiedlichen sozialen Rollen begegnen.

Die Evaluation der STT stellt einen ersten Versuch dar, eine Intervention, die aus insgesamt 58 verschiedenen Kursangeboten bestand, in ihrer Gesamtheit zu evaluieren. Dabei haben wir uns die Logik des Experimentierens zu Nutze gemacht, nach der sich ein Effekt gemeinsamer systematischer Varianz dann, wenn er stark genug ist, auch gegen viele unsystematische Fehlervarianzquellen durchsetzen kann. Bezogen auf unser Beispiel spricht die Tatsache, dass obwohl jede/r Teilnehmer/in an den STT ein unterschiedliches „experimentelles Treatment“ erfahren hat, ein gemeinsamer Effekt – nämlich eine Verringerung der Distanz zwischen Selbstbild und Prototypen der Jugendlichen – sichtbar wurde, dafür, dass dieser Effekt vergleichsweise bedeutsam ist. Den über alle Kursangebote hinweg sichtbaren Effekt haben wir auf eine Gemeinsamkeit aller Angebote zurückgeführt, nämlich auf den personalisierten Kontakt mit einem Kursleiter oder einer Kursleiterin, die sich als im Bereich von Naturwissenschaft und Technik tätig ausgegeben hat. Allerdings haben wir in der vorliegenden Studie keine Daten erhoben, die eine direkte Prüfung dieser Kausalannahme ermöglichen würden. In zukünftigen Untersuchungen sollten deshalb ergänzend Quantität und Qualität des Kontakts zwischen den befragten Teilnehmern eines Freizeit-Kursangebots auf der einen Seite und den Kursleitern als Mitgliedern der stereotypisierten Gruppe von im Bereich von Naturwissenschaft und Technik Tätigen gemessen werden.

Kehren wir zum Ausgangspunkt unserer Überlegungen zurück, so bedeuten die Befunde unserer Studien, dass stereotype Vorstellungen über naturwissenschaftliche Schul-

fächer bestehen; u.a. dahingehend, dass sie im Vergleich zu anderen Fächern stark maskulin konnotiert wahrgenommen werden. Diese Stereotype stellen für Schülerinnen und Schüler Barrieren auf dem Weg zu einer Spezialisierung im Bereich von Naturwissenschaft und Technik dar. Weiter liefern unsere Befunde erste Evidenz dafür, dass obwohl das negative Stereotyp über die Naturwissenschaften sozial geteilt wird, die Diskrepanz zwischen dem Bild, das Jugendliche von sich selbst haben, und ihren Vorstellungen über den prototypischen Vertreter der naturwissenschaftlichen Schulfächer oder Berufe durch das Angebot von Freizeit-Kursen und durch individualisierten, positiven Kontakt mit Personen, die im Bereich von Naturwissenschaft und Technik tätig sind, abgebaut werden kann.

Literatur

- Aronson, E./Bridgeman, D. (1979): Jigsaw groups and the desegregated classroom: In pursuit of common goals. In: *Personality and Social Psychology Bulletin*, 5, S. 438–446.
- Baumert, J./Bos, W./Lehmann, R. (Hrsg.) (2000): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Pflichtschulzeit. (Bd. 1 + 2). Opladen: Leske + Budrich.
- Boehnke, K. (1996): Is intelligence negligible? The relationship of family climate and school behavior in a cross-cultural perspective. Münster: Waxmann.
- Brewer, M.B. (1996): When stereotypes lead to stereotyping: The use of stereotypes in person perception. In: Macrae, C.N./Stangor, C./Hewstone, M. (Hrsg.) *Stereotypes and stereotyping*. New York: Guilford, S. 254–275.
- Brewer, M.B./Miller, N. (1988): Contact and cooperation: When do they work? In: Katz, P./Taylor, D. (Hrsg.) *Eliminating racism: Means and controversies*. New York: Plenum, S. 315–326.
- Brewer, M.B./Miller, N. (1996): Intergroup relations. Buckingham: Open University Press.
- Campbell, J.D. (1990): Self-esteem and clarity of the self-concept. In: *Journal of Personality and Social Psychology*, 59, S. 538–549.
- Campbell, J.D./Trapnell, P.D./Heine, S. J./Katz, I.M./Lavalley, L. F./Lehman, D.R. (1996): Self-concept clarity: Measurement, personality correlates, and cultural boundaries. In: *Journal of Personality and Social Psychology*, 70, S. 141–156.
- Fuhrer, U./Josephs, I. (1999): *Persönliche Objekte, Identität und Entwicklung*. Göttingen: Vandenhoeck & Ruprecht.
- Habermas, T. (1999): *Geliebte Objekte: Symbole und Instrumente der Identitätsbildung*. Frankfurt: Suhrkamp.
- Hamilton, D.L./Trolie, T.K. (1986): Stereotypes and stereotyping: An overview of the cognitive approach. In: Dovidio, J.F./Gaertner, S.L. (Hrsg.) *Prejudice, discrimination, and racism*. Orlando: Academic Press, S. 127–163.
- Hannover, B. (1998): The development of self-concept and interests. In: Hoffmann, L./Krapp, A./Renninger, K.A./Baumert, J. (Hrsg.) *Interest and learning*. Kiel: IPN, S. 105–125.
- Hannover, B. (2002): *Kinder als Mädchen und Jungen*. In: Schröder, R. (Hrsg.) *Das LBS-Kinderbarometer. Was Kinder wünschen, hoffen und befürchten*. Opladen: Leske + Budrich, S. 299–325.
- Hannover, B. (im Druck): Vom biologischen zum psychologischen Geschlecht: Die Entwicklung von Geschlechtsunterschieden. In: Renkl, A. (Hrsg.) *Pädagogische Psychologie*. Bern: Huber.
- Hannover, B./Kessels, U. (eingereicht): Why German school students don't like math and sciences. A self-to-prototype matching approach.
- Hewstone, M. (1996): Contact and categorization: Social psychological interventions to change intergroup relations. In: Macrae, C.N./Stangor, C./Hewstone, M. (Hrsg.) *Stereotypes and stereotyping*. New York: Guilford, S. 323–368.

- Hoffmann, L./Häußler, P./Lehrke, M. (1998): Die IPN-Interessenstudie Physik. Kiel: IPN.
- Kessels, U. (2002): Undoing Gender in der Schule. Eine empirische Studie über Koedukation und Geschlechtsidentität im Physikunterricht. Weinheim: Juventa
- Kessels, U./Hannover, B. (in Vorbereitung) Das Image verschiedener Schulfächer bei Schülern, Schülerinnen und Studierenden.
- Krapp, A. (2000): Interest and human development during adolescence: An educational-psychological approach. Heckhausen, J. (Hrsg.): Motivational psychology of human development. Amsterdam: Elsevier (S. 109–128).
- Niedenthal, P.M./Cantor, N./Kihlstrom, J.F. (1985): Prototype matching: A strategy for social decision making. In: Journal of Personality and Social Psychology, 48, S. 575–584.
- Noack, P. (1990): Jugendentwicklung im Kontext. Zum aktiven Umgang mit sozialen Entwicklungsaufgaben in der Freizeit. München: Psychologie Verlags Union.
- Noack, P./Kracke, B./Wild, E./Hofer, M. (2001): Subjective experiences of social change in East and West Germany: Analyses of perceptions of adolescents and their parents. American Behavioral Scientist, 44, S. 1798–1817.
- Rosch, E. (1978): Principles of categorization. In: Rosch, E./Lloyd, B.B. (Hrsg.): Cognition and categorization. Hillsdale, NJ: Erlbaum.
- Seidel, T./Prenzel, M./Duit, R./Euler, M./Lehrke, M./Geiser, H./Hoffmann, L./Müller, C./Rimmele, R. (2002): „Jetzt bitte alle nach vorne schauen!“ – Lehr-Lernskripts im Physikunterricht und damit verbundene Bedingungen für individuelle Lernprozesse. Unterrichtswissenschaft, 30(1), S. 52–77.
- Setterlund, M.B./Niedenthal, P.M. (1993): „Who am I? Why am I here?“. Self-esteem, self-clarity, and prototype matching. In: Journal of Personality and Social Psychology, 65, S. 769–780.
- Stroebe, W./Hewstone, M./Stephenson, G.M. (1996): Sozialpsychologie. Eine Einführung. Berlin: Springer.
- Tajfel, H./Turner, J.C. (1986): The social identity theory of intergroup behavior. In: W. Austin, W./Worchel, S. (Hrsg.) The social psychology of intergroup relations. Monterey, CA.: Brooks/Cole, S. 7–24.
- Tajfel, H./Wilkes, A.J. (1963): Classification and quantitative judgement. In: British Journal of Psychology, 54, S. 101–114.
- Wild, E./Hofer, M. (1999): Familienbeziehungen in Zeiten sozialen Wandels. In: Walper, S./Pekrun, R. (Hrsg.): Familie und Entwicklung: Perspektiven der Familienpsychologie (S. 131–154). Göttingen: Hogrefe.
- Zwick, M.M./Renn, O. (2000): Die Attraktivität von technischen und ingenieurwissenschaftlichen Fächern bei der Studien- und Berufswahl junger Frauen und Männer. Stuttgart: Akademie für Technikfolgenabschätzung in Baden-Württemberg.

Anschrift der Autorinnen:

Prof. Dr. Bettina Hannover, Freie Universität Berlin, Fb Erziehungswissenschaft und Psychologie, Habelschwerdter Allee 45, 14195 Berlin.

Dr. Ursula Kessels, Freie Universität Berlin, Fb Erziehungswissenschaft und Psychologie, Habelschwerdter Allee 45, 14195 Berlin.

Juliane Strecker/Peter Noack

Wichtigkeit und Nützlichkeit von Mathematik aus Schülersicht¹

1. Einleitung

Die Ergebnisse der jüngeren internationalen Studien TIMSS und PISA (Baumert/Bos/Watermann 1997; OECD 2001) wiesen für manche unerwartet auf ein mäßiges Niveau der Kompetenzen deutscher Schüler – nicht nur – im Bereich Mathematik hin. Das Abschneiden im Ländervergleich, aber vielleicht mehr noch das absolute Leistungsniveau haben Fragen aufgeworfen, zu deren Beantwortung Untersuchungen, die im Rahmen des DFG-Schwerpunktprogramms „Bildungsqualität von Schule“ (Prenzel/Merkens/Noack 1999) durchgeführt werden, beitragen wollen. Dabei dient ein konzeptueller Rahmen als Orientierung, der den Erwerb domänenspezifischer Kompetenz als Resultat eines komplexen Zusammenspiels von individuellen Lernprozessen und schulischen sowie außerschulischen Kontextmerkmalen fasst.

Wenn es vor dem Hintergrund des erheblichen Abstands deutscher Schüler von der internationalen Leistungsspitze in Mathematik um eine Verbesserung der mathematisch-naturwissenschaftlichen Bildung geht, verdienen bei der Suche nach möglichen Einflussfaktoren grundlegende Bedingungen wie in der Gesellschaft verbreitete Einstellungen und Überzeugungen hinsichtlich des Wissensbereichs oder des Lernens in der Schule oder generelle organisatorisch-strukturelle Merkmale ein besonderes Interesse. Das eigene Projekt stellt erstere ins Zentrum. Es werden Überzeugungen von Schülern zu Mathematik untersucht, ihre Variation in Abhängigkeit von Einflüssen des familialen und schulischen Kontexts sowie ihre Effekte auf Lernverhalten und -erfolg. In diesem Beitrag werden erste empirische Ergebnisse zu den Überzeugungsmustern von Schülern im Übergang von der Grund- zur Oberschule sowie zu Zusammenhängen mit den von ihnen wahrgenommenen Überzeugungen im familialen Umfeld vorgestellt. Die ebenfalls erfassten Überzeugungen zu Deutsch und zu Allgemeinbildung dienen dabei als Vergleichsfolie. Zuvor gehen wir näher auf die fokussierten Überzeugungen sowie ihre Stellung in motivationalen Prozessen bei Schülern ein und diskutieren die Rolle der Eltern in diesem Zusammenhang.

1 Die Studie wird durch Mittel der Deutsche Forschungsgemeinschaft gefördert (NO 213/8-1) im Rahmen des Schwerpunktprogramms Bildungsqualität von Schule (SPP 1082).

1.1 *Überzeugungen von Schülern zu Mathematik*

Gruppenunterschiede im Lernerfolg, speziell auch das bessere Abschneiden asiatischer im Vergleich zu westlichen Schülern, werden verschiedentlich mit Verweis auf den kulturspezifisch variierenden Stellenwert von Bildung gedeutet (z.B. Fend 1998; Randel/Stevenson/Witruk 2000). So machen schon asiatische Kinder Klugheit eher an Mathematikkompetenzen fest als Gleichaltrige in den USA (Stevenson 1988), ein Beispiel für Aspekte übergreifender Überzeugungssysteme, die offenbar von Müttern und Lehrern geteilt werden. Wertungen von schulischer Bildung und Einzelfächern stehen ihrerseits in einem systematischen Zusammenhang mit Schulleistungen, wie in verschiedenen nationalen Kontexten Untersuchungen mit dem Fokus auf Mathematik zeigen konnten (z.B. Marjoribanks 1992; Rech/Stevens 1996; Robinson/Taylor/Piolat 1990). Nicht nur domänenspezifische Einschätzungen, sondern auch Urteile über die Schule allgemein scheinen sich auf Mathematikleistungen auszuwirken (Berndt/Miller 1990).

Studien, die Einstellungen deutscher Schüler untersuchen, zeichnen kein erfreuliches Bild. Die Schulverdrossenheit nimmt mit der Schulzeit zu, die Lernfreude sinkt, und aus der Sicht von Schülern spielt sich das eigentliche Leben offenbar jenseits der Schulmauern ab (Eckerle/Kraak 1993; Helmke 1997; Tillmann u.a. 1984). Mit Blick auf den mathematisch-naturwissenschaftlichen Bereich sind die Befunde gemischt, wobei die Haltung gegenüber Mathematik tendenziell positiv von gegenüber den naturwissenschaftlichen Fächern absticht (z.B. Gräber 1992; Jugendwerk der Deutschen Shell 1992). Daten der TIMS-Studie weisen zwar darauf hin, dass deutsche Schüler Mathematik wichtig finden und dasselbe auch über ihre Mütter äußern, die wahrgenommenen Urteile ihrer Freunde fallen im internationalen Vergleich hingegen extrem niedrig aus (Schmitz 1996). Angesichts des repräsentativen Charakters der Stichprobe könnte Letzteres als indirekter Ausdruck der eigenen Wertung gedeutet werden, die Diskrepanz dürfte dann auf Effekte sozialer Erwünschtheit zurückgehen.

Die Arbeitsgruppe von Eccles (z.B. Eccles u.a. 1983; Wigfield 1994) wählt einen differenzierten Zugang zu domänenspezifischen Einstellungen und Überzeugungen von Schülern und stellt sie in den Zusammenhang eines Erwartung x Wert-theoretischen Modells der Schülermotivation. Es werden die einem Fach zugeschriebene Wichtigkeit, seine subjektive Nützlichkeit mit Blick auf kurz- und längerfristige Ziele, Spaß bei entsprechenden Aktivitäten als intrinsischer Faktor sowie deren Kosten unterschieden, die zusammen den Wert des Gegenstands ausmachen und in Interaktion mit den eigenen Erfolgserwartungen leistungsbezogenes Verhalten motivieren und auf diesem Weg die Leistung beeinflussen sollen. Empirische Überprüfungen von Ausschnitten des Modells bestätigten diese Annahmen (z.B. Eccles 1984; Wigfield/Eccles 1992).

Die vorgeschlagene Differenzierung der Bewertung scheint sich etwa im Übergang zur Präadoleszenz einzustellen, während jüngere Schüler nur fachbezogenes Interesse einerseits und Wichtigkeit sowie Nützlichkeit als zweiten Aspekt unterscheiden (Eccles u.a. 1993). Ein weiteres altersgradiertes Muster betrifft die Ausprägung der Urteile. Dabei weist der Trend nach unten. Einen markanten Einbruch erfährt die Wertschätzung der Fächer nach dem Übergang zur weiterführenden Schule (Eccles u.a. 1983).

1.2 *Familiale Einflüsse*

Eltern nehmen in vielfältiger Weise Einfluss auf das Lernverhalten ihrer Kinder in der Schule und auf deren Bildungserfolg. Die untersuchten Merkmalen reichen von allgemeinen Charakteristika wie dem Familienklima und der elterlichen Empathie (Stecher 1996) über kulturelle Aktivitäten der Familie (Fend 1997) bis hin zu lernbezogenem Verhalten der Eltern im Rahmen der Hausaufgabenbetreuung (Helmke/Schrader/Lehn-eis-Klepper 1991). Auch wenn es nahe liegt, vor allem letztere Merkmale des Familienkontexts als Vermittlungsglieder zwischen elterlichen Einstellungen und motivationalen Orientierungen ihrer Kinder zu deuten (Eccles 1993), wurden die Einstellungen von Müttern und Vätern in der bisherigen Forschung eher vernachlässigt.

In Eccles Modell gehen sie als von den Söhnen und Töchtern wahrgenommene Überzeugungen und Erwartungen direkt in den Motivationsprozess ein. Dabei ist nicht davon auszugehen, dass solche Wahrnehmungen der Kinder die Einstellungen und Überzeugungen ihrer Eltern passiv spiegeln. Zwar belegen eine Reihe von Untersuchungen zu anderen Einstellungsbereichen eine systematische Transmission in der Familie (z.B. Georg/Hasenberg/Zinnecker 1996; Kracke u.a. 1993; Zinnecker/Georg 1996). Zumindest für politische Einstellungen lässt sich aber auch zeigen, dass, erstens, die Einstellungen von Kindern deutlich weniger mit den tatsächlichen Orientierungen der Eltern korrespondieren als mit den durch die Kinder wahrgenommenen elterlichen Orientierungen (Gniewosz u.a. 2001). Zweitens beeinflussen auch Kinder die Einstellungen ihrer Eltern (Hofer u.a. 1998). Inwieweit diese qualifizierenden Feststellungen ebenfalls für Einstellungen gegenüber Bildung und spezifischen Wissensdomänen wie Mathematik gelten, bleibt mangels entsprechender Forschungen derzeit eine offene Frage. In jedem Fall ist davon auszugehen, dass auch im hier interessierenden Bereich von den Einstellungen und Überzeugungen der Eltern, wie sie sich aus der Sicht von Schülern darstellen, stärkere Einflüsse ausgehen als von den tatsächlichen Einstellungen.

1.3 *Fragestellung*

Die im Folgenden vorgestellten Analysen beziehen sich auf Wichtigkeits- und Nützlichkeitsurteile von Schülern zu Mathematik. Sie werden im Kontext von Überzeugungen zu Deutsch als weiterer Domäne und zu Allgemeinbildung untersucht. Dabei ist das erste Anliegen deskriptiver Art. Es werden Urteile zu den drei betrachteten Bereichen verglichen sowie Variationen in Abhängigkeit vom Geschlecht und besuchten Schultyp geprüft.

Unsere zweite Frage richtet sich auf Zusammenhänge zwischen den Einschätzungen, die sich auf Mathematik, Deutsch und Allgemeinbildung beziehen. Mögliche Unterschiede im absoluten Niveau der subjektiven Wichtigkeit und Nützlichkeit schließen eine enge Assoziation nicht aus. Auch wenn Mathematik beispielsweise die höchste Wertschätzung erfahren sollte, könnten jene, die diesen Bereich positiver beurteilen dasselbe auch hinsichtlich Deutsch oder Allgemeinbildung tun.

Schließlich betrachten wir im dritten Schritt Zusammenhänge zwischen den von Schülern wahrgenommenen Elternurteilen und den Urteilen der Schüler selbst. Nur für diesen Teil der Auswertungen lassen sich auf der Basis der vorliegenden Literatur eindeutige Hypothesen formulieren. Die wahrgenommenen Elternurteile zu Mathematik sollten parallele Urteile der Schüler vorhersagen.

Hinsichtlich unserer ersten beiden Anliegen haben die Analysen einen eher explorativen Charakter. So ist beispielsweise denkbar, dass Überzeugungen zu Allgemeinbildung, die als Oberkategorie gesehen werden kann, enger mit den Überzeugungen zu den beiden spezifischen Domänen zusammen hängen als diese untereinander. Unter einer alltagsweltlichen Perspektive indessen könnte der gemeinsame Charakter von Mathematik und Deutsch als Schulfächern zu stark übereinstimmenden Urteilen führen, während die Assoziation mit den Urteilen zur Allgemeinbildung lockerer ausfallen mag. Vergleichbar ließen sich Argumente vortragen, die zu unterschiedlichen Voraussagen hinsichtlich Mittelwertsähnlichkeiten und -unterschieden führen würden.

2. Empirische Untersuchung

Angesichts der generell eher hohen Stabilität von Überzeugungssystemen wurden die Daten an einem Punkt der Bildungslaufbahn von Schülern gesammelt, dem Übergang zur weiterführenden Schule, der eine merkliche Zäsur und damit ein ökologisches Experiment im Sinne Bronfenbrenners (1981) darstellt, um die Wirkung von Kontextbedingungen besser analysieren zu können. Der Übergang geht, wie berichtet, im Mittel einher mit Veränderungen der Ausprägung von Schülereinstellungen wie auch von deren Differenziertheit. In solchen Phasen sollte sich die Wirkung von Kontexteinflüssen deutlicher zeigen. Die analysierten Daten entstammen einer umfangreichen Fragebogenerhebung, in der neben den betrachteten auch eine größere Zahl weiterer individueller sowie kontextueller Merkmale erfasst wurde und an der neben den Schülern auch deren Eltern und ein Geschwister teilnahmen. Dem Design der Untersuchung folgend wurde eine für das Bundesland Thüringen repräsentative Stichprobe von knapp 1200 Schülern einbezogen. Da die Datenerfassung aktuell noch nicht abgeschlossen ist, beschränken sich die folgenden Auswertungen auf eine größere Teilstichprobe.

2.1 Methode

Stichprobe

Aktuell stehen Fragebögen von einem guten Drittel der befragten Schüler für Analysen aufbereitet zur Verfügung. In den folgenden Auswertungen werden insgesamt 431 Teilnehmer berücksichtigt, von denen komplette Datensätze für die betrachteten Variablenbereiche vorliegen. Sie nahmen an der Befragung zu Beginn des fünften Schuljahrs teil und waren zum Erhebungszeitpunkt im Mittel 10,6 Jahre alt. Beide Geschlechter sind etwa hälftig vertreten (51% weiblich).

Das Stichprobendesign sieht Erhebungen in zwei fünften Klassen von jeweils 15 Regelschulen und Gymnasien vor, die die wesentlichen Formen weiterführender Schulen im Bundesland sind. Von den Schülern, über die hier berichtet wird, besuchten 38% eine Regelschule, die Übrigen ein Gymnasium. Für diese Teilstichprobe kann zwar keine Repräsentativität reklamiert werden. In ihr ist jedoch ein Querschnitt von Schülern aus den unterschiedlichen in Thüringen anzutreffenden Lebenslagen vertreten.

Fragebogen

Die Urteile der Schüler wurden durch kurze standardisierte Instrumente mit geschlossenem Antwortformat (jeweils: 1 = „stimmt gar nicht“, 4 = „stimmt genau“) erhoben, deren Items aus Befragungen der Arbeitsgruppe von Eccles (1993) adaptiert wurden. Die *Wichtigkeit von Mathematik* wurde durch zwei Items erfasst („Dieses Fach ist für das Leben jedes Menschen wichtig.“, „Alle sollten dieses Fach gut können.“ $r = .32$). Drei Items dienten zur Erhebung der subjektiven *Nützlichkeit von Mathematik* („In diesem Fach etwas zu können ist später wichtig, um viel Geld zu verdienen.“, „Für die spätere Ausbildung ist dieses Fach sehr wichtig.“, „Um später eine gute Arbeit zu finden, ist dieses Fach wichtig“, $\alpha = .70$). *Wichtigkeit und Nützlichkeit von Deutsch* wurden bis auf den Urteilgegenstand wortgleich erfasst ($r = .39$ bzw. $\alpha = .68$). Hinsichtlich *Allgemeinbildung* wurden beide Aspekte durch jeweils zwei Items angesprochen (*Wichtigkeit*, „Es ist nicht so wichtig, sich in vielen Bereichen gut auszukennen“ (umgepolt), „Jeder sollte eine gute Allgemeinbildung haben.“, $r = .27$; *Nützlichkeit*, „Nur wer viel weiß, kann im Leben etwas erreichen.“, „Wer eine gute Allgemeinbildung hat, verdient später einmal viel Geld.“, $r = .40$)

Die Urteile von Mutter und Vater aus Sicht der Schüler wurden mit denselben Items bestimmt, wobei die vorangestellte Instruktion Einschätzungen der jeweiligen Elternmeinung erfragte. Die bivariaten Korrelationen bzw. internen Konsistenzen entsprachen jenen, die für die parallelen Selbstauskünfte der Schüler ermittelt wurden (*Mathematik*, *Wichtigkeit*, Mutter: $r = .44$, Vater: $r = .52$; *Nützlichkeit*, Mutter: $\alpha = .77$, Vater: $\alpha = .82$; *Deutsch*, *Wichtigkeit*, Mutter: $r = .45$, Vater: $r = .54$; *Nützlichkeit*, Mutter: $\alpha = .78$, Vater: $\alpha = .82$; *Allgemeinbildung*, *Wichtigkeit*, Mutter: $r = .22$, Vater: $r = .23$; *Nützlichkeit*, Mutter: $\alpha = .51$, Vater: $\alpha = .56$).

Analyse

Mit Blick auf die erste verfolgte Fragestellung wurden die Daten varianzanalytisch ausgewertet. Zu diesem Zweck wurde jeweils der Mittelwert der Items einer gegebenen Skala bestimmt. Zusammenhänge zwischen den Urteilen hinsichtlich der drei Bereiche wurden durch Tests von Strukturgleichungsmodellen (AMOS, Arbuckle/Wothke 1999) separat für die Aspekte *Wichtigkeit* und *Nützlichkeit* analysiert. Dazu wurde zunächst jeweils ein Messmodell mit den Items der jeweiligen Skalen als Indikatoren geprüft und in einem zweiten Schritt das theoretisch postulierte Zusammenhangsmodell, das Korrelationen zwischen den je drei latenten Variablen vorsah. Es wurde schließlich empirisch gegen mögliche sparsamere und weniger sparsame Modelle abgesichert. Zur Optimie-

rung der Passung wurden solche Messfehlerkorrelationen zugelassen, die sich auf Paare wortgleicher Items bezogen und mithin als Hinweise auf itemspezifische Bedeutungsanteile interpretiert werden konnten.

Die Strategie der Analyse zur Rolle der elterlichen Urteile entsprach der hinsichtlich der zweiten Fragestellung eingesetzten. Ausgehend von der Annahme, dass die Überzeugungen beider Eltern aus Sicht der Schüler weitgehend übereinstimmen sollten, wurden im theoretischen Modell latente Variablen zweiter Ordnung postuliert, die das Elternurteil zur Wichtigkeit bzw. Nützlichkeit von Mathematik repräsentieren. Für diese wurde jeweils ein Pfad auf das korrespondierende Urteil der Schüler angenommen. Da weiterhin von Assoziationen zwischen Wichtigkeits- und Nützlichkeitsurteilen auszugehen war, wurden diese als Korrelation zwischen den beiden Elternvariablen zugelassen, aus technischen Gründen als Pfad vom Nützlichkeits- auf das Wichtigkeitsurteil der Schüler spezifiziert. Auch wenn diese Effektrichtung plausibel ist, sollte sie nicht als gerichteter Effekt i.e.S. interpretiert werden. Auch bei der Deutung der gerichteten Effekte vom wahrgenommenen Elternurteil auf die Schülerurteile ist daran zu denken, dass die zugrunde liegende Datenbasis querschnittlicher Art ist.

2.2 Ergebnisse

Mittelwertsunterschiede

Die Items zur Wichtigkeit und Nützlichkeit von Mathematik, Deutsch und Allgemeinbildung erhielten durchweg eine hohe Zustimmung seitens der Schüler. Dasselbe galt, wenn die Schüler die Perspektiven ihrer Mütter und Väter einnehmen sollten. Tabelle 1 zeigt die Skalenmittelwerte und Standardabweichungen.

Es wurde eine Varianzanalyse mit Schultyp und Geschlecht als unabhängigen Variablen sowie Urteilsgegenstand (Mathematik, Deutsch, Allgemeinbildung), Urteilsaspekt (Wichtigkeit, Nützlichkeit) und Referenzperson (Schüler, Wahrnehmung Mutter, Wahrnehmung Vater) als Messwiederholungsfaktoren durchgeführt. Angesichts des erheblichen Stichprobenumfangs werden nur Effekte mit einer Irrtumswahrscheinlichkeit über .01 berichtet und die jeweilige Varianzaufklärung (Eta^2) als zusätzliche Information, um die praktische Bedeutung einschätzen zu können.

Es ergab sich zunächst ein Effekt des Messwiederholungsfaktors Gegenstand ($F_{(2/422)} = 34,1, p < .001, \text{Eta}^2 = .14$) mit höheren Wertungen von Mathematik und Deutsch gegenüber Allgemeinbildung, der allerdings angesichts einer Wechselwirkung von Gegenstand und Schultyp ($F_{(2/422)} = 10.6, p < .001, \text{Eta}^2 = .05$) zu qualifizieren ist. Den genannten Unterschied machten im Wesentlichen die Regelschüler, die darüber hinaus generell höhere Zustimmungen äußerten, wie ein Haupteffekt des Schultyps belegt ($F_{(1/422)} = 7.7, p < .01, \text{Eta}^2 = .02$) In Abbildung 1 (S. 366) werden diese Ergebnisse grafisch veranschaulicht.

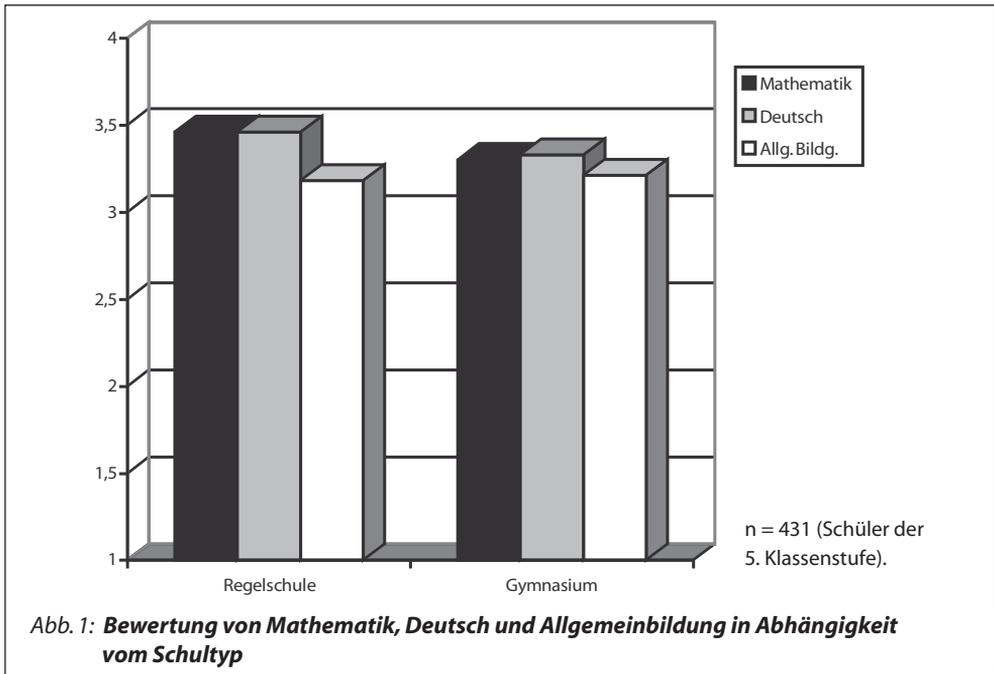
Gleichzeitig deutet eine Interaktion von Gegenstand und Urteilsaspekt ($F_{(1/422)} = 12.2, p < .001, \text{Eta}^2 = .06$) darauf hin, dass die angesprochene Differenzierung stärker bei den Nützlichkeitsurteilen ausgeprägt ist als bei den Wichtigkeitsurteilen, dies wiederum

vor allem unter den Regelschülern, während Gymnasiasten sogar Allgemeinbildung der Wichtigkeit nach mit leichtem Abstand den höchsten Rang zuordnen (Interaktion Gegenstand x Aspekt x Schultyp: $F_{(1/422)} = 20.13$, $p < .001$, $\text{Eta}^2 = .09$). Weiterhin gilt das Muster eher in der weiblichen Teilstichprobe als in jener der männlichen Schüler (Interaktion Gegenstand x Aspekt x Geschlecht: $F_{(1/422)} = 4.90$, $p < .01$, $\text{Eta}^2 = .02$).

Tab. 1: Wichtigkeits- und Nützlichkeitsurteile zu Mathematik, Deutsch und Allgemeinbildung; Mittelwerte und Standardabweichungen

Schultyp		Mathematik		Deutsch		Allgemeinbildung	
	Person*	Wicht.	Nützl.	Wicht.	Nützl.	Wicht.	Nützl.
Regelschule							
	Kind	3.5	3.48	3.52	3.41	3.32	3.3
		(.61)	(.68)	(.65)	(.70)	(.72)	(.76)
	Mutter	3.5	3.6	3.46	3.57	3.07	3.22
		(.66)	(.53)	(.70)	(.57)	(.77)	(.76)
Vater	3.48	3.5	3.49	3.46	3.21	3.3	
	(.68)	(.71)	(.66)	(.72)	(.71)	(.82)	
Gymnasium							
	Kind	3.32	3.31	3.38	3.36	3.4	3.03
		(.63)	(.66)	(.62)	(.58)	(.52)	(.71)
	Mutter	3.19	3.39	3.26	3.42	3.34	3.09
		(.68)	(.64)	(.67)	(.62)	(.57)	(.73)
	Vater	3.23	3.37	3.23	3.37	3.41	3.15
		(.69)	(.64)	(.68)	(.62)	(.58)	(.73)
Anmerkung: n = 431 (Schüler der 5. Klassenstufe); Skalenrange jeweils 1–4; Standardabweichungen in Klammern.							
* Urteile der Schüler und Elternurteile in der Wahrnehmung der Schüler.							

Bislang wurde von Unterschieden zwischen dem eigenen Urteil der Schüler und den wahrgenommenen Elternurteilen abgesehen. Für die drei Referenzpersonen ergaben sich jedoch Variationen sowohl mit Blick auf den Urteilsaspekt ($F_{(1/422)} = 16.87$, $p < .001$, $\text{Eta}^2 = .07$) als auch hinsichtlich des beurteilten Gegenstands ($F_{(1/422)} = 4.34$, $p < .01$, $\text{Eta}^2 = .04$). Während die Schüler die Wichtigkeit von Mathematik, Deutsch und Allgemeinbildung selbst höher beurteilten als das aus ihrer Sicht ihre Eltern tun, fielen die Antworten zur Nützlichkeit vergleichbar aus. Die Differenzierung zwischen den Fächern war am deutlichsten im Schülerurteil und dem wahrgenommenen Mutterurteil zu beobachten, dies wiederum vor allem innerhalb der Regelschülerfamilien (Interaktion Gegenstand x Person x Schultyp: $F_{(1/422)} = 4.81$, $p < .001$, $\text{Eta}^2 = .04$).



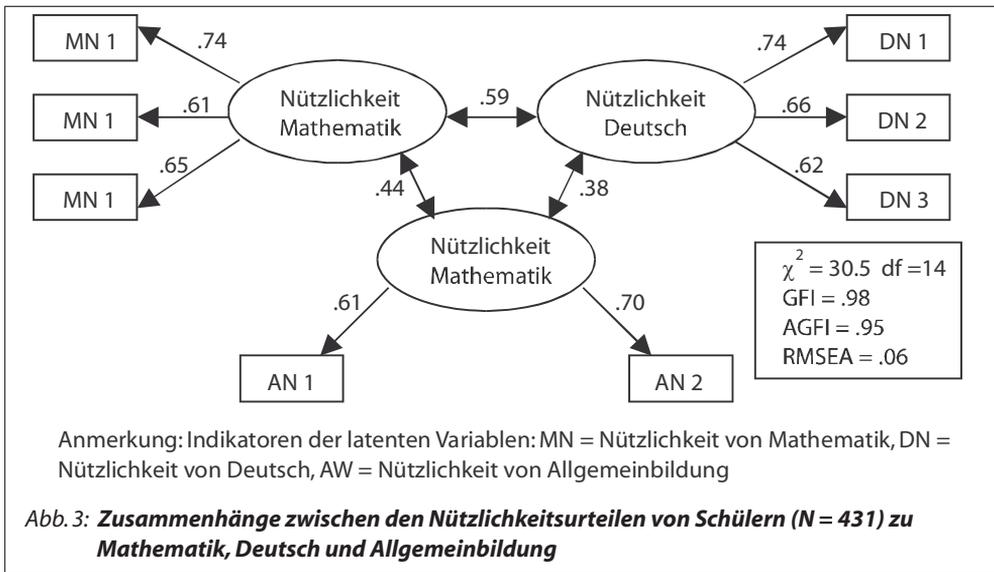
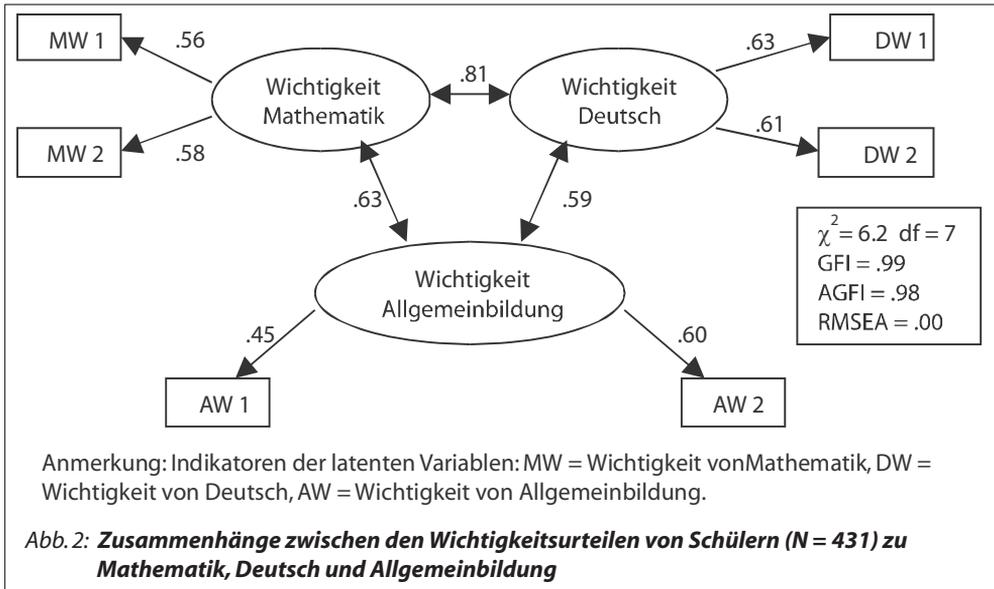
Zusammenhänge zwischen den Überzeugungen zu Mathematik, Deutsch und Allgemeinbildung

Das für die *Wichtigkeitsurteile* der Schüler postulierte Modell wies eine gute Übereinstimmung mit den empirischen Daten auf ($\text{Chi}^2(7) = 6.2, p > .05$; GFI = .99; AGFI = .98; RMSEA = .00). Alternative Modelle, in denen auf eine der angenommenen Korrelationen verzichtet wurde, zeigten eine jeweils signifikant schlechtere Passung ($\Delta \text{Chi}^2, p < .05$). Das akzeptierte Modell wird in Abbildung 2 gezeigt. Aus Gründen der Übersichtlichkeit wird in der Abbildung auf die Darstellung der Messfehler und Messfehlerkorrelationen verzichtet.

Die durchweg hohen Korrelationskoeffizienten weisen auf enge Zusammenhänge zwischen der subjektiven Wichtigkeit der drei Urteilsbereiche hin. Allerdings fällt die Assoziation der Urteile zu Mathematik und Deutsch etwas höher aus als jene zwischen Allgemeinbildung und den beiden spezifischen Domänen.

Auch für die *Nützlichkeitsurteile* der Schüler ist die Passung zwischen dem postulierten Modell und den Daten akzeptabel ($\text{Chi}^2(14) = 30.5, p < .01$; GFI = .98; AGFI = .95; RMSEA = .05). Zwar fällt der Chi^2 -Test signifikant aus, das Verhältnis der Chi^2 -Punkte zu Freiheitsgraden (2:1) wie auch die übrigen Kennwerte sprechen jedoch für das Modell. Aus unserer Sicht ist noch entscheidender, dass wiederum der Fit möglicher alternativer Modelle signifikant geringer ist. Abbildung 3 zeigt das akzeptierte Modell.

Das Muster der Korrelationen entspricht jenem, das schon für die *Wichtigkeitsurteile* der Schüler berichtet wurde. Während die subjektive Nützlichkeit von Mathematik und Deutsch einen engen Zusammenhang aufweist, sind die Assoziationen mit dem Nützlichkeitsurteil für Allgemeinbildung geringer aber ebenfalls statistisch bedeutsam.

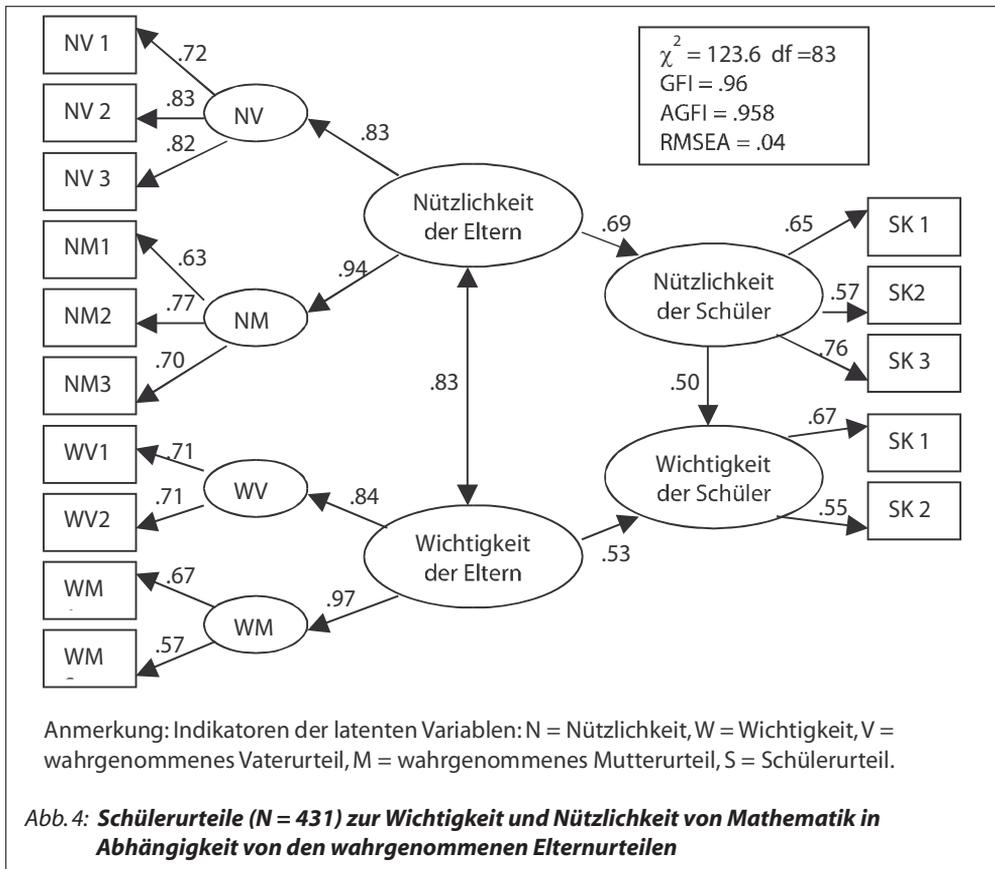


Schülerurteile zu Mathematik und wahrgenommene Elternurteile

Das spezifizierte Strukturgleichungsmodell stimmte angemessen mit den empirischen Daten überein ($\chi^2(83) = 123.6$, $p < .01$; $GFI = .96$; $AGFI = .95$; $RMSEA = .04$). Auch in diesem Fall weist der χ^2 -Test zwar eine signifikante Abweichung aus. Das Verhältnis von χ^2 -Punkten zu Freiheitsgraden (1.5:1) spricht jedoch wiederum zusammen mit den weiteren Koeffizienten für eine Annahme des Modells ebenso wie der Umstand, dass alternative Modelle, die beispielsweise auf einen der postulierten Pfade verzichten,

mit einer signifikant verschlechterten Passung einhergehen, und solche, die zusätzliche Pfade postulieren, keine bedeutsame Verbesserung des Fits erbringen. Das akzeptierte Modell wird in Abbildung 4 vorgestellt.

Pfadkoeffizienten erheblicher Höhe zeigen an, dass die von den Schülern vermuteten elterlichen Urteile zu Mathematik bedeutsame Prädiktoren der Schülerurteile sind. Das gilt sowohl für die Wichtigkeit als auch für die Nützlichkeit der Domäne. Gleichzeitig bestehen auch enge Verbindungen zwischen den beiden Urteilsaspekten hinsichtlich der wahrgenommenen Elternurteile wie hinsichtlich der subjektiven Wichtigkeit und Nützlichkeit, die die Schüler äußern. Ergänzende Modelltests, die der Möglichkeit nachgingen, die jeweiligen Urteile zu einer Gesamtdomänenbewertung zusammenzufassen, ergaben jedoch eine vergleichsweise schlechte Passung mit den Daten. Dasselbe galt für Modelle, die Gesamtbewertungen getrennt für beide Eltern spezifizierten.



3. Diskussion

Als Teil eines umfanglicheren Projekts, das den Einstellungen von Schülern gegenüber Mathematik und anderen Wissensbereichen sowie ihren familialen und schulischen Bedingungen nachgeht, wurden im hier vorgestellten ersten Schritt Urteile zur Wichtigkeit und Nützlichkeit analysiert. Im Zentrum standen Übereinstimmungen und Unterschiede zwischen den Einschätzungen von Mathematik, Deutsch und Bildung allgemein sowie deren Einbettung in die elterlichen Einschätzungen, wie sie von Schülern wahrgenommen werden.

Die erste Frage richtete sich auf Variationen der Urteile in Abhängigkeit von ihrem Gegenstand, der Bezugsperson sowie Geschlecht und besuchtem Schultyp. Bei durchweg hohen Bewertungen ergaben sich keine nennenswerten Unterschiede hinsichtlich der Bedeutung, die Mathematik und Deutsch zugeschrieben wurde, während Allgemeinbildung etwas hinter diesen beiden zurückblieb, insbesondere hinsichtlich der subjektiven Nützlichkeit, und im Wesentlichen unter den Regelschülern. Es ist nicht auszuschließen, dass diese Differenzierung zum Teil darauf zurückzuführen ist, dass angesichts des unterschiedlichen Charakters von Allgemeinbildung einerseits und der beiden Schulfächer andererseits die Itemformulierungen nur für die zwei letzteren wortidentisch gewählt werden konnten. Der Umstand, dass vor allem die Regelschüler den Unterschied machten, nicht hingegen die Gymnasiasten, lässt jedoch eher an eine inhaltliche Deutung denken, die auf eine Absetzung von Schul- gegenüber Allgemeinwissen verweist.

Die Unterschiede zwischen den eigenen Urteilen und den bei den Eltern vermuteten Einschätzungen fielen gering aus und betrafen eher den Wichtigkeitsaspekt. Es bestätigt sich die Vermutung einer deutlichen faktischen oder vermeintlichen Übereinstimmung der Schüler- und Elternsicht, wobei der Mittelwertsvergleich den Schluss auf eine intergenerationale Kontinuität der Überzeugungen in der Familie letztlich nicht zulässt. Falls eine solche Kontinuität vorliegt, wofür die weiter unten angesprochenen korrelativen Ergebnisse sprechen, bleibt offen, ob sie vor allem auf den Wahrnehmungen der Schüler beruht. Wenn die im Rahmen der eigenen Befragungen ebenfalls erhobenen Elterndaten vorliegen, wird sich prüfen lassen, in welchem Ausmaß die Wahrnehmungen der Söhne und Töchter die elterliche Perspektive realistisch spiegeln.

Die zweite hier verfolgte Frage bezog sich auf die Zusammenhänge zwischen den domänenspezifischen Urteilen, die Einblicke in die Struktur der bildungsbezogenen Überzeugungen geben können. Enge Zusammenhänge sowohl der Wichtigkeits- als auch der Nützlichkeitseinschätzungen für die drei betrachteten Bereiche kommen nicht unvermutet.

Aufschlussreicher ist die größere Nähe der Urteile zu Mathematik und Deutsch, mit denen jene zur Allgemeinbildung etwas schwächer verbunden waren. Das Muster, dass sich unter anderem Blickwinkel auch im Mittelwertsvergleich zeigte, kann als Hinweis darauf gesehen werden, dass die Schüler die angesprochenen Bereiche subjektiv weniger im Sinne einer hierarchischen Ordnung mit mathematischen und muttersprachlichen Kenntnissen als Teilbereichen von Allgemeinbildung organisieren, denn nach der äußere-

ren Einbindung in den schulischen Alltag. Mathematik und Deutsch kommt die eindeutige Rolle als Schulfächer zu, Allgemeinbildung mag diffuser sein oder weniger als schulisches Lernziel ausgemacht werden.

Es fällt weiterhin auf, dass die Zusammenhänge zwischen den Nützlichkeitsurteilen etwas geringer ausfallen als die Zusammenhänge zwischen den Wichtigkeitsurteilen, auch wenn der Unterschied ein geringes Ausmaß hat und sämtliche Koeffizienten signifikant sind. Der im Vergleich globalere Charakter der Wichtigkeitsüberzeugungen mag hierzu beigetragen haben. Die subjektive Nützlichkeit hingegen wurde mit Bezug auf bestimmte Ziele, zum Beispiel beruflicher und materieller Art erfasst.

Die dritte Teilfrage, die wir verfolgt haben, betraf Zusammenhänge zwischen den wahrgenommenen Überzeugungen der Eltern und jenen, die die Schüler selbst äußerten. Die beobachtete Kontinuität kann ein Hinweis sein für den Einfluss der elterlichen Haltungen gegenüber Mathematik, der über die Wahrnehmungen der Schüler wirksam wird. Eccles (1993) beobachtete in ihrer Studie einen erheblichen Effekt der elterlichen Sichtweisen, der sogar über jenen hinaus gehen kann, den Leistungsrückmeldungen direkt auf Schüler ausüben. Allerdings können anhand der uns derzeit verfügbaren Daten Projektionseffekte (vgl. Gniewosz u.a. 2001) nicht ausgeschlossen werden. In jedem Fall lässt sich festhalten, dass in den subjektiven Repräsentationen der Schüler die eigenen Überzeugungen bruchlos eingebettet sind in die Sichtweisen des familialen Umfelds, ein Rückhalt, der überzeugungsstabilisierend wirken dürfte.

Bei der Würdigung der Ergebnisse ist das junge Alter der teilnehmenden Schüler zu bedenken. Sie können, wie unsere Beobachtungen im Einklang mit Eccles und Kollegen (1993) zeigen, die betrachteten Urteilsaspekte zwar durchaus differenzieren. Die hohen Übereinstimmungen zwischen diesen ebenso wie zwischen der eigenen und der den Eltern zugeschriebenen Sicht mögen jedoch mit der fortschreitenden kognitiven Entwicklung (Kail/Park 1994; Keating 1990) zurückgehen.

Neben der Konzentration auf die Schülerperspektive und dem aktuell noch querschnittlichen Charakter des Datensatzes, die schon als Einschränkungen der vorgestellten Analysen angesprochen wurden, ist weiterhin darauf hinzuweisen, dass die berücksichtigten Schüler nur eine Teilgruppe der Gesamtstichprobe sind und mithin nicht deren repräsentativen Charakter für das Bundesland aufweisen. Umfang wie Heterogenität dieser Teilstichprobe lassen jedoch vermuten, dass die berichteten Ergebnisse Bestand haben dürften.

Nicht nur in dieser Hinsicht haben wir erste empirische Schritte im Rahmen des Projekts vorgestellt. In der Zukunft wird es zum einen darum gehen, die elterlichen Einflüsse differenzierter zu analysieren und ihr Zusammenwirken mit weiteren Kontextfaktoren, vor allem solchen im schulischen Umfeld zu prüfen. Zum anderen wird es gelten, die vermutete Bedeutung der bildungsbezogenen Überzeugungen von Schülern für ihr Lernverhalten und, letztlich, für die resultierenden mathematischen Fähigkeiten zu untersuchen.

Literatur

- Arbuckle, J.L./Wothke, W. (1999): AMOS 4.0 user's guide. Chicago: SmallWaters.
- Baumert, J./Bos, W./Watermann, R. (1997): TIMSS/III. Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse. Berlin MPI für Bildungsforschung.
- Berndt, T.J./Miller, K.E. (1990): Expectancies, values, and achievement in junior high school. In: *Journal of Educational Psychology* 15, S. 608–616.
- Bronfenbrenner, U. (1981). *Die Ökologie der menschlichen Entwicklung*. Stuttgart: Klett-Cotta.
- Eccles, J.S. (1984): Sex differences in achievement patterns. In: Sonderegger, T. (Hrsg.): *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, S. 97–132.
- Eccles, J.S. (1993): School and family effects on the ontogeny of children's interests, self-perceptions, and activity choices. In: Jacobs, J.E. (Hrsg.): *Nebraska Symposium on Motivation 1992*. Lincoln: University of Nebraska Press, S. 145–208.
- Eccles, J.S./Adler, T. F./Futtermann, R./Goff, S. B./Kaczala, C.M./Meece, J./Midgley, C. (1983): Expectancies, values and academic behaviors. In: J. T. Spence (Hrsg.): *Achievement and achievement motives*. San Francisco: Freeman, S. 75–146.
- Eccles, J.S./Wigfield, A./Harold, R./Blumenfeld, P. (1993): Age and gender differences in children's self- and task perceptions during elementary school. In: *Child Development* 64: S. 830–847.
- Eckerle, G.-A./Kraak, B. (1993): *Selbst- und Weltbilder von Schülern und Lehrern*. Göttingen: Hogrefe.
- Fend, H. (1997): *Soziale Entwicklung in der Adoleszenz*. Bern: Huber.
- Fend, H. (1998): *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistungen*. Weinheim: Juventa.
- Georg, W./Hasenberg, R./Zinnecker, J. (1996): Die Weitergabe der Sportkultur in der Familie. Söhne und Töchter im Vergleich. In: Zinnecker, J./Silbereisen, R.K. (Hrsg.): *Kindheit in Deutschland*. Weinheim: Juventa, S. 137–143.
- Gniewosz, B./Lange, M./Höhl, C./Kröger, C./Reinhardt, S./Noack, P. (2001): Fremdenfeindlichkeit Jugendlicher und wahrgenommene Fremdenfeindlichkeit im sozialen Umfeld. Poster auf der Tagung der Fachgruppe Entwicklungspsychologie, Postdam, September.
- Gräber, W. (1992): Untersuchung zum Schülerinteresse an Chemie und Chemieunterricht. In: *Chemie in der Schule* 39, S. 270–273.
- Helmke, A. (1997): Entwicklung lern- und leistungsbezogener Motive und Einstellungen: Ergebnisse. In: Weinert, F.E./Helmke, A. (Hrsg.): *Entwicklung im Grundschulalter*. Weinheim: Beltz, S. 59–76.
- Helmke, A./Schrader, F.-W./Lehneis-Klepper, G. (1991): Zur Rolle des Elternverhaltens für die Schulleistungsentwicklung. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 23, S. 1–22.
- Hofer, M./Noack, P./Oepke, M./Buhl, M./Wild, E. (1998): Reciprocal processes in political socialization among East and West German families. Vortrag auf dem 7th Biennial Meeting of the Society for Research on Adolescence, San Diego, Februar/März.
- Jugendwerk der Deutschen Shell (Hrsg.) (1992): *Jugend '92*. Opladen: Leske + Budrich
- Kail, R./Park, Y.-S. (1994): Processing time, articulation time, and memory span. In: *Journal of Experimental Child Psychology* 27, S. 281–291.
- Keating, D.P. (1990): Adolescent thinking. In: Feldman, S.S./Elliott, G.R. (Hrsg.): *At the threshold*. Cambridge: Harvard University Press, S. 54–89.
- Kracke, B./Noack, P./Hofer, M./Klein-Allermann, E. (1993): Die rechte Gesinnung: Zusammenhänge zwischen autoritären Orientierungen ost- und westdeutscher Jugendlicher und familialen Beziehungen. In: *Zeitschrift für Pädagogik* 39, S. 971–988.
- Marjoribanks, K. (1992): The predictive validity of an attitudes-to-school scale in relation to children's academic achievement. In: *Educational and Psychological Measurement* 52, S. 945–949.
- OECD (2001): *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD.

- Prenzel, M./Merkens, H./Noack, P. (1999): Die Bildungsqualität von Schule: Fachliches und fächerübergreifendes Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten. Antrag an die Deutsche Forschungsgemeinschaft. Kiel.
- Randel, B./Stevenson, H.W./Witruk, E. (2000): Attitudes, beliefs, and mathematics achievement of German and Japanese high school students. In: *International Journal of Behavioral Development* 24, S. 190–198.
- Rech, J.E./Stevens, D.J. (1996): Variables related to mathematics achievement among black students. In: *The Journal of Educational Research* 89, S. 346–350.
- Robinson, W.P./Tayler, C.A./Piolat, M. (1990): School attainment, self-esteem, and identity: France and England. In: *European Journal of Social Psychology* 20, 387–403.
- Schmitz, B. (1996): Anfertigung von Hausaufgaben und Schulleistung: Ergebnisse der TIMS-Studie. Vortrag auf der Tagung der AEPF, Berlin, Oktober.
- Stecher, L. (1996): Schulhabitus und soziales Kapital in der Familie. In: Zinnecker, J./Silbereisen, R.K. (Hrsg.): *Kindheit in Deutschland*. Weinheim: Juventa, S. 267–289.
- Stevenson, H.W. (1988): Culture and schooling: Influences on cognitive development. In: Hetherington, E.M./Lerner, R./Perlmutter, M. (Hrsg.): *Child development in life-span perspective*. Hillsdale: Erlbaum, S. 241–258.
- Tillmann, K.-J./Faulstich-Wieland, H./Horstkemper, M./Weissbach, B. (1984): Die Entwicklung von Schulverdrossenheit und Selbstvertrauen bei Schülern in der Sekundarstufe. In: *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie* 4, S. 231–249.
- Wigfield, A. (1994): The role of children's achievement values in the self-regulation of their learning outcomes. In: Schunk, D.H./Zimmerman, B.J. (Hrsg.): *Self-regulation of learning and performance*. Hillsdale: Erlbaum S. 101–124.
- Wigfield, A./Eccles, J.S. (1992): The development of achievement task values: A theoretical analysis. In: *Developmental Review* 12, S. 265–310.
- Zinnecker, J./Georg, W. (1996): Die Weitergabe kirchlich-religiöser Familienerziehung und Orientierung zwischen Eltern- und Kindergeneration. In: Zinnecker, J./Silbereisen, R.K. (Hrsg.): *Kindheit in Deutschland*. Weinheim: Juventa, S. 347–356.

Anschrift der Autoren:

Dipl.-Psych. Juliane Strecker, FSU Jena, Institut für Psychologie, Am Steiger 3/1, 07743 Jena.
Prof. Dr. Peter Noack, FSU Jena, Institut für Psychologie, Am Steiger 3/1, 07743 Jena.

Teil VI:

Schulforschung

Evaluation und Feedback auf Klassen- und Schulebene

Hartmut Ditton/Bettina Arnoldt/Eva Bornemann

Entwicklung und Implementation eines extern unterstützten Systems der Qualitätssicherung an Schulen – QuaSSU¹

1. Zielsetzung und Grundlagen

Die Grundlage des Projekts QuaSSU bildet ein Mehrebenenmodell zu Faktoren der Schul- und Unterrichtsqualität (Ditton 2000a, 2000b). Das Modell integriert Ergebnisse aus der Schulqualitäts- bzw. Schuleffektivitätsforschung (Scheerens/Bosker 1997; Teddlie/Reynolds 2000) mit Ergebnissen der Unterrichtsforschung (Einsiedler 1997a, 1997b; Helmke/Weinert 1997; Gruehn 2000). Ziel der Untersuchung ist es, Qualitätsmerkmale guter Schulen und guten Unterrichts aufeinander bezogen zu analysieren und die gewonnenen Ergebnisse zugleich systematisch für Entwicklungsprozesse an den beteiligten Schulen zu nutzen. In der Untersuchung werden Lehrkräfte und Schüler zur Schule und zum Unterricht – bislang in den Fächern Mathematik, Deutsch und Englisch – befragt. Zu den Ergebnissen erhalten die einzelnen Schulen und Lehrkräfte ein differenziertes Feedback in einer vergleichenden Form: In Abbildungen werden die individuell spezifischen Ergebnisse im Vergleich zu einer Referenzgruppe dargestellt und erläutert. Das Feedbackverfahren wurde bereits in vorangegangenen Untersuchungen eingesetzt und stößt bei den Teilnehmern, wie die bisherigen Erfahrungen zeigen, auf eine hohe Akzeptanz (Ditton 2001a, 2001b). Im Mittelpunkt der kommenden Projektphase stehen Analysen zu den durch das Feedback bewirkten Veränderungen, den Bedingungen der Weiterentwicklung des auf Befragungsdaten basierten Feedbacks zu einem praktikablen System der Qualitätssicherung an Schulen.

Kennzeichnend für den Untersuchungsansatz ist die Verbindung von Schulqualitäts- und Schulentwicklungsforschung bzw. die Frage nach den Möglichkeiten und Grenzen des *Transfers* wissenschaftlicher Erkenntnisse und Empfehlungen in das schulische Anwendungsfeld. Da aus dem Anwendungsbezug zugleich wieder Impulse für die Weiterentwicklung des Forschungsdesigns gewonnen werden, bestehen gewisse Parallelen zu Ansätzen der sog. *Handlungsforschung* (Heinze 1975).

Zu Merkmalen *guter Schulen* liegt eine Vielzahl an Untersuchungen und Systematisierungen vor (Scheerens/Bosker 1997), die es erlauben, als relevant anzusehende Faktoren differenziert anzugeben. Eine hohe Bedeutung kommt in einer zusammenfassenden Übersicht folgenden Dimensionen zu (Ditton 2000a): eine entwickelte *Schulkultur* bzw.

¹ Die Studie wurde gefördert durch Mittel der DFG (Geschäftszeichen DI 767/2-1) im Rahmen des DFG-Schwerpunktprogramms BIQUA.

ein Schulethos (im Sinne eines gemeinsam geteilten Aufgabenverständnisses im Kollegium), ein effektives *Management* der Schule, eine enge (auf den Unterricht bezogene) *Kooperation* sowie eine gezielte *Personalpolitik*. Eine hohe Schulqualität ist nicht als Selbstzweck anzusehen, sondern als eine bedeutsame Voraussetzung für guten Unterricht. Methodisch angemessen überprüft wurden die Beziehungen zwischen den Ebenen Schule und Unterricht allerdings nur selten (Mortimore u.a. 1988; Teddlie/ Stringfield 1993); auch für Deutschland stehen entsprechende Untersuchungen aus (Ditton/Krecker 1995; Ditton 2000a, 2000b). Die verbreitete These, dass mit einer höheren Schulqualität eine höhere Qualität und Homogenität des Unterrichts einhergeht, wird in der laufenden Untersuchung in einem Design mit einer genügend großen Zahl an Schulklassen pro teilnehmender Schule mehrebenenanalytisch überprüft. Weitgehend ungeklärt ist außerdem die Frage, wieweit eine Übereinstimmung in den von Schülern und Lehrkräften wahrgenommenen Faktoren der Schulqualität besteht. Den inzwischen vorliegenden Ergebnisse zufolge bestehen teilweise erhebliche Divergenzen.

In den Befragungen zu den Aspekten der *Unterrichtsqualität* werden in QuaSSU Elemente aus der Forschungstradition des sog. Prozess-Produkt-Paradigmas mit der Schul- bzw. Klassenklimaforschung verbunden. Trotz der Kritik an beiden Paradigmen (Einsiedler 1997a, 1997b; Helmke/Weinert 1997, Gruehn 2000; Ditton i.Druck) ergeben sich aus beiden Forschungstraditionen relevante Hinweise auf Merkmale eines hinsichtlich der Leistungsentwicklung der Schüler wirksamen Unterrichts. Zu Grunde gelegt wird hierbei ein Vorschlag zur Systematisierung von Unterrichtsfaktoren nach Slavin (1996) mit den primären Dimensionen *Qualität*, *Angemessenheit*, *Anregung* und *Zeit*, der um die Faktoren *Leistungsangst* und *Sozialklima* (Schüler-Schüler und Lehrer-Schüler) ergänzt wird². Zu diesen Merkmalen werden die Schüler im Hinblick auf den Fachunterricht bei einer konkreten Lehrkraft befragt. Auf der anderen Seite geben die Lehrkräfte eine Einschätzung zur Klasse ab, in der die Befragung durchgeführt wird (u.a. zu Leistungsniveau und -streuung, Motivation, Disziplin). Zudem wird jede Lehrkraft um eine Einschätzung dazu gebeten, wie die Ergebnisse der Befragung zum Unterricht in der Klasse vermutlich ausfallen werden. Obwohl die Dynamik von Unterricht durch eine Befragung nicht zureichend abgebildet werden kann, sind damit Ansatzpunkte gegeben, um Aspekte der wechselseitigen Wahrnehmung im Unterricht zu untersuchen.

Bereits in einer vorangegangenen Untersuchung wurde der Wunsch der Schulen und Lehrkräfte aufgegriffen, über die Ergebnisse der Erhebungen informiert zu werden (Ditton 2001a, 2001b). Für diesen Zweck wurde inzwischen ein Feedbacksystem entwickelt, das darauf abzielt, die ermittelten schul- und unterrichtsspezifischen Ergebnisse für die Entwicklungsarbeit an den Schulen nutzbar zu machen. Die Möglichkeiten der Nutzung von Feedbackinformationen aus empirischen Erhebungen an Schulen werden derzeit vielfach diskutiert und in mehreren Studien erprobt bzw. sind dort vorgesehen (LAU, QUASUM, MARKUS und WALZER, TIMSS, PISA, IGLU; s. Klieme/Baumert/

2 Zwischen den von uns verwendeten Dimensionen und denen aus den Analysen der TIMS-Video-studie (Unterrichts- und Klassenführung, Schülerorientierung, kognitive Aktivierung; vgl. Klieme/Schümer/Knoll 2001) bestehen vielfach Übereinstimmungen.

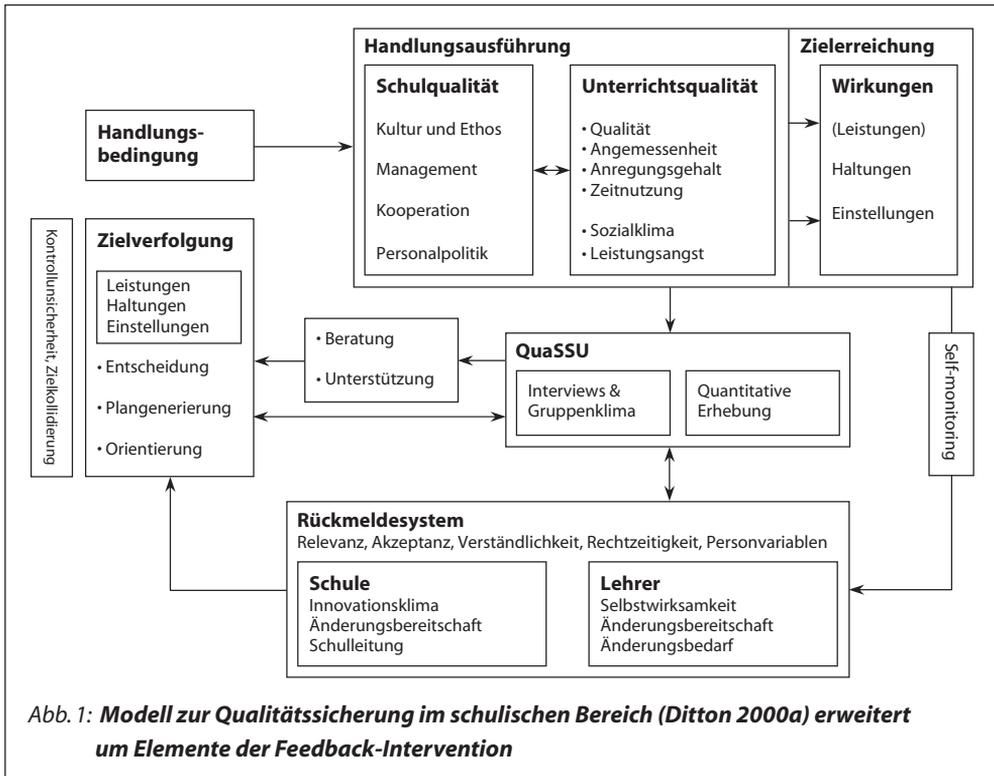
Schwippert 2000; vgl. auch Klieme/Baumert 2001). Analysen zu den Anwendungsbedingungen und damit erzielten Wirkungen liegen jedoch nicht vor. Im Folgenden wird darauf näher eingegangen.

Feedback ist allgemein jede Art von Information, die verfügbar ist, um das Ausmaß der Zielerreichung zu überprüfen. Rückmeldungen zum Handlungserfolg sind eine grundlegende Voraussetzung zur kontrollierten Handlungsregulierung und -optimierung. Allerdings führen Rückmeldungen nicht automatisch zu Veränderungen oder Verbesserungen im Handeln. In einer Metaanalyse zur Wirkung von Feedbackinterventionen konnten Kluger/DeNisi (1996) zwar zeigen, dass sich im Durchschnitt aller Studien die Leistungen verbesserten. In einem Drittel der analysierten Fälle ergaben sich jedoch auch negative Effekte. Als bedeutsame Einflussgrößen erwiesen sich dabei zum einen Merkmale der Rückmeldung selbst (Rechtzeitigkeit, Umfang, Art und Form; Kluger/DeNisi 1996) und zum anderen Merkmale der Adressaten (Balk 2000). Nach Kluger/DeNisi sind Rückmeldungen vor allem dann wirksam, wenn sie auf die *Task-motivation-Ebene* zielen, d.h. auf die im Mittelpunkt stehende Aufgabe. Rückmeldungen auf der *Meta-task-Ebene* (auf das Selbst bezogene Rückmeldungen) haben dagegen eine geringe Wirksamkeit. Für Rückmeldungen auf der *Task-learning-Ebene* (Details der Aufgabe) ergaben sich nur unter spezifischen Bedingungen positive Effekte (Kluger/DeNisi 1996). Zudem finden sich Hinweise dafür, dass Personen mit hoher Selbstwirksamkeit eher in der Lage sind, Rückmeldungen konstruktiv zu verwerten, als Personen mit niedriger Selbstwirksamkeit (Balk 2000). Für Letztere kann daher eine begleitende Beratung angezeigt sein. Über diese ersten Ergebnisse hinaus besteht die Notwendigkeit, die Anwendungsbedingungen und die Wirksamkeit von Rückmeldungen im schulischen Kontext im Einzelnen zu untersuchen.

In unserer Untersuchung wird ein Modell von Frese/Zapf (1994) in Verbindung mit den Ergebnissen aus den Analysen von Kluger/DeNisi (1996) als Werkzeug genutzt, um die Komponenten des Handlungsablaufs abzubilden und die Schnittstelle von *Handlungsausführung*, *Selbst-Monitoring* und *externem Feedback* aufzuzeigen. Dargestellt sind diese Bezüge in Abbildung 1, in der das bisher für die Untersuchung verwendete Modell (Ditton 2000a, 2000b) hinsichtlich der Intervention durch das Feedbackverfahren weiter differenziert und um relevante Untersuchungsgrößen erweitert wurde.

Wie die Abbildung zeigt, kann eine Rückmeldung auf der Basis von Daten einer externen Erhebung den Ausgangspunkt für die Reflexion der Handlungsplanung, Handlungsausführung und Zielsetzungen bilden, indem es die Prozesse des Self-Monitoring ergänzt. Dies dürfte vor allem dann der Fall sein, wenn Daten eines objektivierten sozialen Vergleichs³ zur Verfügung gestellt werden, wie es in QuaSSU der Fall ist. Diese Informationen stehen bei einem Self-Monitoring nicht zur Verfügung. Um allerdings die Feedbackinformationen mit den Entwicklungsprozessen an den Schulen eng zu verzahnen, muss ein Austausch über die Kompatibilität der Zielsetzungen und zu den Handlungsperspektiven zwischen der Forschergruppe und den Kollegien erfolgen. Von er

3 Vergleiche zu: eigene Schule/Stichprobe aller Schulen; eigener Unterricht/Unterricht einer Referenzgruppe von Lehrkräften



heblicher Bedeutung ist deshalb, dass im Projekt eine *längerfristige Zusammenarbeit* mit den teilnehmenden Schulen vorgesehen ist. In organisatorischer Hinsicht ist wesentlich, dass in Absprache mit den Teilnehmern die notwendigen Vorkehrungen getroffen wurden, um den *Datenschutz* zu gewährleisten. Im Rahmen der Untersuchung werden u.a. auch „sensible Daten“ zum Unterricht einzelner Lehrkräfte erhoben, die vertraulich behandelt werden müssen.

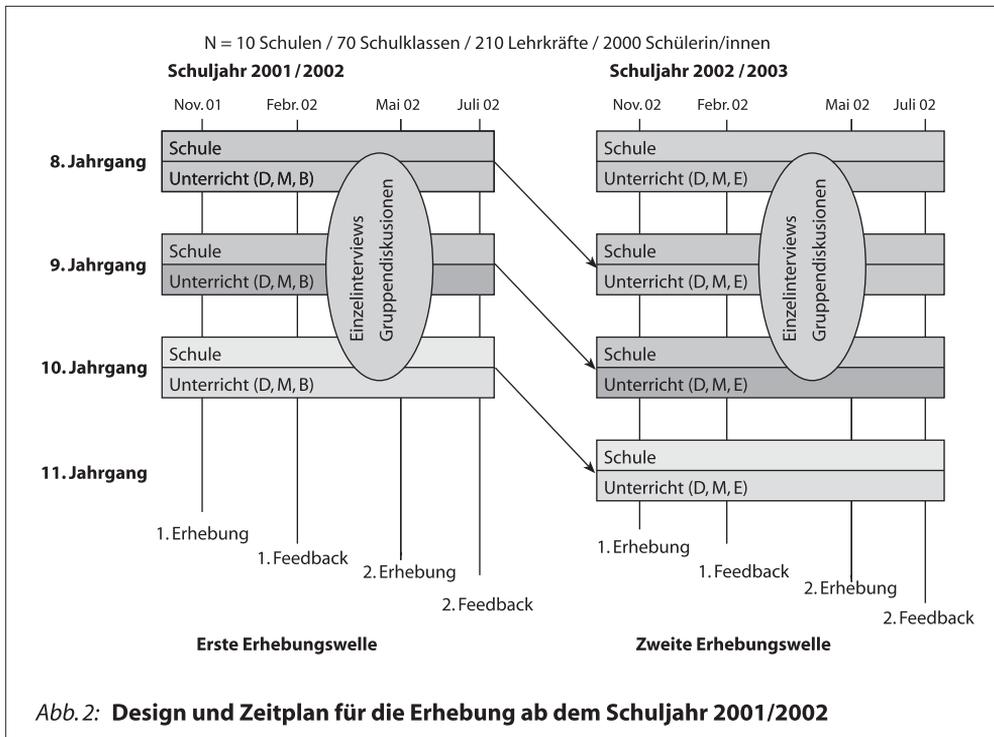
Dafür, dass extern organisierte bzw. unterstützte Feedbackverfahren wichtige Impulse für die Schulentwicklung geben können, sprechen die Erfahrungen aus den Projekten von Fitz-Gibbon (1996) an der University of Durham. Dort werden seit 1992 umfassende Feedback-Systeme entwickelt, die sich inzwischen einer äußerst regen Nachfrage seitens der Schulen erfreuen. Die externe Unterstützung erscheint nicht zuletzt auch deshalb geboten, weil an den Schulen häufig die Zeit und das erforderliche Wissen fehlen, um die „Datenbasis“ für weitere Entwicklungsprozesse zu schaffen. Zudem fehlen bei rein internen Evaluationen in aller Regel Vergleichs- bzw. Orientierungsmaßstäbe, die eine Einordnung individueller Ergebnisse erleichtern können.

2. Methode

2.1 Design und Vorgehen

Das Design der Untersuchung sieht in der kommenden Phase des Projekts zwei eng miteinander verzahnte Teilstudien vor (vgl. Abb. 2):

1. eine quantitative längsschnittliche Untersuchung zu Faktoren der Schul- und Unterrichtsqualität mit vier Erhebungszeitpunkten (2 Erhebungswellen mit jeweils 2 Messzeitpunkten), und
2. eine qualitative Studie zur Angemessenheit und Umsetzbarkeit der Feedbackinformation für Prozesse der Schul- und Unterrichtsentwicklung.



Die Erhebungen umfassen die wiederholte Befragung zur Wahrnehmung der Schüler und Lehrkräfte der Klassenstufen acht bis zehn zum Unterricht in den Fächern Mathematik, Deutsch und Englisch sowie zu Faktoren der Schulqualität. Für den Längsschnitt werden im kommenden Schuljahr die derzeit teilnehmenden Klassen in den Klassenstufen neun bis elf befragt und die Klassen der achten Klassenstufe werden neu hinzugenommen. Daraus ergeben sich zwei Erhebungswellen (derzeit: Klassenstufen 8–10; kommendes Schuljahr: Klassenstufen 8–11) mit jeweils zwei Erhebungszeitpunkten

zum Beginn und Ende des Schuljahres. Zwischen den beiden Erhebungen liegen das Feedback sowie Gruppendiskussionen in den Lehrerkollegien und einzelnen Schulklassen. Außerdem werden Einzelinterviews mit einer Teilstichprobe der Lehrkräfte durchgeführt.

Die erste Befragung findet jeweils ca. zwei Monate nach Schuljahresbeginn statt. Dieser Zeitraum wurde von den Schulen als ausreichend angesehen, um verlässliche Informationen der Schüler zum Unterricht der einzelnen Lehrkräfte zu erhalten.

Zu den Ergebnissen erhalten die Schulen und einzelnen Lehrkräfte eine schriftliche Rückmeldung. Die Rückmeldungen sind so aufgebaut, dass die Adressaten zur Schul- bzw. Unterrichtsbefragung das eigene Ergebnis im Vergleich zur jeweiligen Referenzgruppe der Schulen bzw. Lehrkräfte mitgeteilt bekommen. Die Rückmeldungen erfolgen in grafischer Form, d.h. zu den einzelnen Untersuchungsbereichen werden Abbildungen erstellt, die jeweils kommentiert sind.

Zusammen mit der Rückmeldung erhält jede Lehrkraft einen Fragebogen zu deren Verständlichkeit, Aussagekraft und Verwertbarkeit. Durch offene Fragen ist zudem die Möglichkeit für differenziertere Antworten gegeben.

Die zweite Erhebung findet ca. zwei Monate vor Schuljahresende statt. Sie beschränkt sich auf die Befragung zum Fachunterricht, da bezüglich der schulischen Faktoren bedeutsame Veränderungen in einem so kurzen Zeitraum nicht zu erwarten sind und eine Überbeanspruchung der Teilnehmer vermieden werden soll.

Um die Wirksamkeit der Feedbacks zuverlässig zu ermitteln, wäre ein experimentelles Design mit Experimental- und Kontrollgruppe erforderlich. Dies ist in einer Feldstudie, die auf eine längerfristige Kooperation mit den Untersuchungsteilnehmern angelegt ist, nicht zu realisieren. Stattdessen wird eine Kontrolle zumindest soweit realisiert, als ein Teil der Schulen erst zusammen mit der zweiten Erhebung zum Schuljahresende ein Feedback erhält. Auf diesem Weg lassen sich die Effekte der Feedbacks im Vergleich zur Kontrollgruppe zumindest näherungsweise abschätzen.

Wie die bisherigen Erfahrungen zeigen, bieten die schriftlichen Befragungen allein keine ausreichend breite Grundlage für die Weiterentwicklung des Feedback zu einem Qualitätssicherungssystem. Im Zeitraum zwischen den beiden quantitativen Erhebungen werden daher ergänzend *Gruppendiskussionen* und *Einzelinterviews* durchgeführt. In *Leitfadeninterviews* wird bei einer Teilstichprobe der Lehrkräfte die Fragestellung zur Rückmeldung aus den standardisierten Befragungen vertieft und ergänzt. Interviews werden sowohl mit Teilnehmern geführt als auch mit Lehrkräften, die nicht bereit waren, sich an der Untersuchung zu beteiligen. Bei den Verweigerern sollen in erster Linie die Gründe für die Weigerung ermittelt sowie der Frage nachgegangen werden, von welchen Voraussetzungen sie eine Teilnahme ggf. abhängig machen würden. Bei den Teilnehmern konzentrieren sich die Interviews darauf, wie die rückgemeldeten Informationen verarbeitet und welche Konsequenzen ggf. gezogen wurden. Weitere Themen sind das Unterrichtsverständnis bzw. der didaktische Ansatz, den die Lehrkräfte verfolgen, und die Frage nach den für erforderlich gehaltenen weiteren Beratungs- bzw. Unterstützungsleistungen, um Verbesserungen auf der Schul- und Unterrichtsebene zu erzielen. Die bisherigen Erfahrungen zeigen außerdem, dass von den Teilnehmern gewünscht

wird, die den Befragungen zu Grunde liegende theoretische Konzeption von Schul- und Unterrichtsqualität noch näher erläutert zu bekommen. Dafür soll nach Absprache mit den Kollegien der sog. *Pädagogische Tag* an den Schulen genutzt werden. Auf die Einbeziehung der Schüler in den schulischen Entwicklungsprozess zielen Gruppendiskussionen in einer Auswahl der an der Befragung teilnehmenden Schulklassen.

2.2 Organisation der Untersuchung und Stichprobe

Die Befragung an den Schulen erfolgt auf schriftlich-postalischem Weg. An jeder Schule hat sich eine Lehrkraft bereit erklärt, die Koordination der Erhebungen zu übernehmen und als Ansprechpartner für alle Beteiligten zur Verfügung zu stehen. Bei der Befragung der Schüler ist bislang jeweils eine Lehrkraft anwesend, die selbst nicht in der Klasse unterrichtet. Sofern die Schulen dem zustimmen, soll in ausgewählten Klassen erprobt werden, die Erhebung von den Schülern selbstverantwortlich organisieren zu lassen.

Eine grundsätzliche Bedingung und unabdingbare Voraussetzung, um die Schulen und Lehrkräfte für eine Teilnahme zu gewinnen, war die Sicherstellung des Datenschutzes. Dazu wurde in Absprache mit den Schulen ein Codiersystem entwickelt. Dieses erlaubt es, Datensätze aus mehreren Erhebungen im Längsschnitt bzw. die Befragungsergebnisse zum Unterricht einer Lehrkraft aus mehreren Klassen miteinander zu verknüpfen, ohne dass für die Forschergruppe die Einzelpersonen identifizierbar sind.

Tab. 1: Stichprobe Schüler/innen und Lehrkräfte

Befragte	Fach	N	
		1. Erhebung*	2. Erhebung**
Schüler zur Schule		1239	1992
Schüler zum Unterricht	Deutsch	1009	1751
	Englisch	1008	1490
	Mathematik	1057	1697
Lehrkräfte zur Schule		194	133
Lehrkräfte zum Unterricht bzw. Schulklassen	Deutsch	43	73
	Englisch	46	63
	Mathematik	39	71
* 7 Schulen; ** 10 Schulen, davon 5 aus der ersten Erhebung			

Die jetzige Untersuchung ist als Weiterführung eines vorausgegangenen Forschungsprojekts konzipiert, bei dem bereits ein Großteil der Instrumente entwickelt wurde. Aus der Stichprobe von N = 178 Schulen und 186 Schulklassen aus einer Erhebung im Jahr 2000 steht ein Datensatz zur Bildung von Vergleichswerten zur Verfügung. Die erste Erhebung im jetzigen Forschungsprojekt wurde im April/Mai 2001 durchgeführt, im Juni/

Juli 2001 erhielten die Schulen und Lehrkräfte die Rückmeldungen. Beteiligt waren in dieser ersten Phase 7 Schulen mit 141 Schulkassen, 1239 Schülern und 194 Lehrkräften. Die zweite Erhebung – als Beginn der Längsschnittstudie – fand im Dezember 2001 statt, die Rückmeldungen wurden im März 2002 versandt. Teilnehmer waren hier 10 Schulen (davon 5 aus der ersten Erhebung) mit 207 Schulklassen, 1992 Schüler/innen und 133 Lehrkräften (s. Tab. 1). Mit dieser Gruppe ist eine längerfristige Kooperation vorgesehen. Daher waren auch gezielt Schulen im Großraum München angesprochen worden sowie Schulen, die bereits an der ersten Erhebung teilgenommen hatten und nicht zu weit von München entfernt sind.

2.3 Übersicht der Bereiche und Skalen

Tab. 2: **Dimensionen und Faktoren des Unterrichts – erweitertes QUAIT-Modell nach Slavin (1996)**

Quality / Qualität	Klarheit
	Strukturiertheit
Appropriateness /Angemessenheit	Tempo
	Schwierigkeit
	Leistungserwartungen
Incentives / Motivierung	Interessantheit
	Übung
	Unterstützung
Time / Zeit	Zeitnutzung
	Klassenmanagement
Weitere Faktoren	Diagnostische Kompetenz der Lehrkraft
	Verhältnis Lehrkraft-Schüler
	Sozialklima in der Klasse
	Leistungsangst

Zu den im Untersuchungsmodell spezifizierten Dimensionen der Schul- und Unterrichtsqualität wurden bereits in der erwähnten vorausgegangenen Erhebung im Jahr 2000 Skalen entwickelt. Da die Teilnehmer die Beanspruchung durch die Befragung teilweise als hoch empfunden hatten, wurden die Skalen auf der Grundlage der Item- und Skalenkennwerte gekürzt. Berücksichtigt wurde dabei außerdem die faktorielle Struktur. Alle Skalen – auch in der gekürzten Fassung – bestehen aus wenigstens drei Items. Überwiegend liegen die Reliabilitätskoeffizienten bei ca. .70 (Cronbachs α), teilweise liegen sie deutlich höher. Insbesondere für die Skalen zur Unterrichtsbefragung werden gute bis sehr gute Kennwerte erreicht. Dies ist übereinstimmend für die Fächer Deutsch, Mathematik und Englisch der Fall. Aus Raumgründen geben die beiden Tabellen ledig-

lich einen zusammenfassenden Überblick. Eine Zusammenstellung der verwendeten Skalen und Items steht im Internet zur Verfügung und kann dort heruntergeladen werden (www.quassu.net).

Tab. 3: Dimensionen und Faktoren der Schulqualität

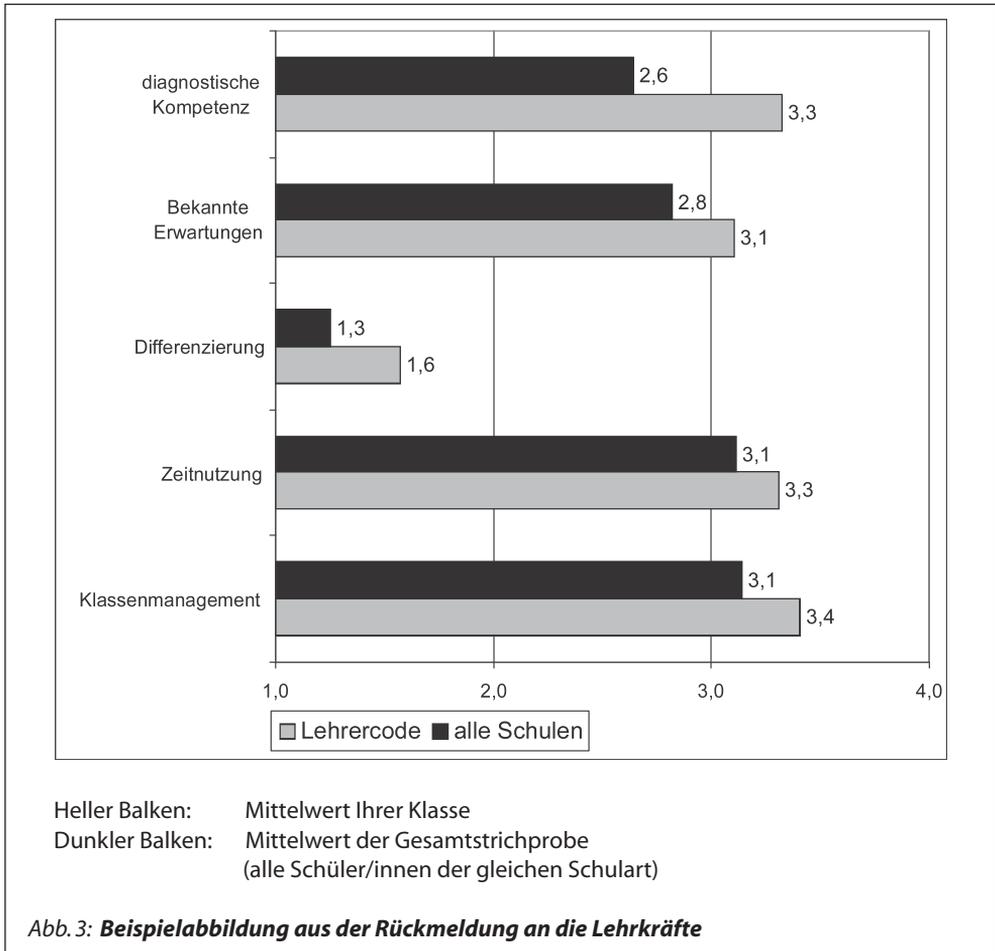
Schulkultur / Schulethos	Schulprofil
	Hohe Erwartungen
	Mitbestimmung der Schüler
Schulmanagement	Geordneter Schul-/Unterrichtsbetrieb
	geregelte Zuständigkeiten
	Entscheidungsprozesse
Kooperation und Koordination	Kooperation im Kollegium
	Absprachen, Kommunikation mit (Fach-)Kollegen
Personalpolitik- und Entwicklung	Fort- und Weiterbildung
	Erfahrungsaustausch

2.4 Rückmeldungen

Auf der Basis der Untersuchungsergebnisse werden zwei Formen der Rückmeldung (Schule/Unterricht) mit jeweils vier Abbildungen zu den Dimensionen des Untersuchungsmodells erstellt. Die Schulrückmeldung wird dabei dem Koordinator an der Schule zugestellt. Die Rückmeldungen zum Unterricht werden ebenfalls über den Koordinator zugestellt, allerdings in einzeln verschlossenen Umschlägen für jede Lehrkraft. Durch den aufgedruckten Lehrercode ist die korrekte Zuordnung zu den einzelnen Lehrkräften gewährleistet. Lehrkräfte, die mit mehr als einer Klasse an der Unterrichtsbefragung beteiligt waren, erhalten entsprechend mehrere Rückmeldeumschläge. Die Zuordnung zu der jeweiligen Klasse ist durch eine Kennung in den Abbildungen möglich.

Für die Rückmeldungen werden die Ergebnisse in grafischer Form, bislang als Balkendiagramme, dargestellt. Es handelt sich jeweils um Doppelblätter, auf deren linker Seite die Abbildung gezeigt wird und auf der gegenüberliegenden Seite die Erläuterung dazu. Alle Abbildungen beinhalten den Vergleich der individuellen Ergebnisse einer Schule bzw. zum Unterricht einer Lehrkraft mit der Gesamtstichprobe bzw. einer Referenzgruppe. Differenziert wird hierbei nach Schularten und bezogen auf den Unterricht zusätzlich nach Fächern. Jeweils eine Beispielabbildung (Abb. 3, Abb. 4) zu Schule und Unterricht ist nachfolgend wiedergegeben (s. auch www.quassu.net und Ditton 2001a, 2001b).

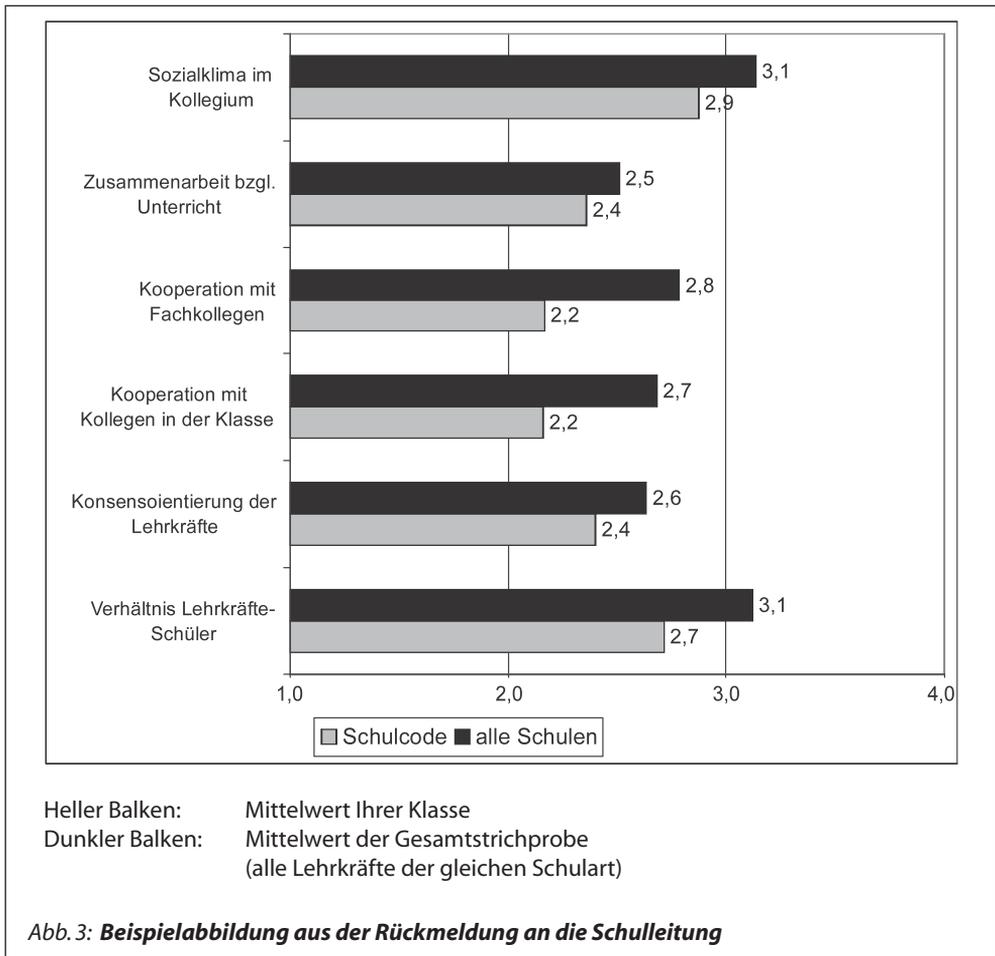
Die Darstellung von Mittelwerten kann sicherlich hinterfragt werden. Allerdings zeigen die bisherigen Erfahrungen, dass diese Form von den Lehrkräften als anschaulich und unmittelbar verständlich bezeichnet wird.



3. Überblick zu bisherigen Ergebnissen⁴

Analysen der Daten aus den aktuellen Erhebungen für die bisher vorliegende Teilstichprobe zeigen auf der *Schulebene*, dass die Varianz zwischen den Teilnehmerschulen in allen vier o.g. Dimensionen (Kultur, Management, Kooperation und Personalpolitik)

4 Ergebnisse zu Teilfragestellungen des vorangegangenen Forschungsprojekts, das mit QuaSSU weitergeführt wird, sind veröffentlicht zu: Schulprogrammarbeit (Ditton/Edelhäuser/Merz 2001), Einstellungen zu zentralen Testuntersuchungen (Ditton/Merz/Edelhäuser 2002), Feedbackverfahren (Ditton 2001a, 2001b); Übereinstimmung der Unterrichtswahrnehmung durch Schüler und Lehrkräfte im Fach Mathematik (Ditton 2002).



und den einzelnen Faktoren dieser Dimensionen sehr erheblich ist. Eine multivariate Analyse der Daten aus der Befragung der *Lehrkräfte* mit 21 Skalen führt zu einem Anteil von 41,5% der Varianz zwischen den Schulen. Überdies ergeben sich zwischen den Einzelfaktoren Beziehungen in der erwarteten Richtung: Die Berufszufriedenheit der Lehrkräfte korreliert mit dem Sozialklima im Kollegium und der Führungskompetenz der Schulleitung auf der Individualebene signifikant, wenn auch nicht sehr hoch ($r = .20$). Auf Schulebene aggregiert sind die Beziehungen wesentlich enger ($r = .50$).

Multivariate Analysen der Daten aus der *Schülerbefragung* (7 Skalen) ergeben einen Anteil von 12,1% an Varianz zwischen den Schulen in den untersuchten Dimensionen und bestätigen ebenfalls die erwarteten Beziehungen zwischen den einzelnen Faktoren. Enge Beziehungen auf der Individual- und Aggregatebene finden sich zwischen der wahrgenommenen Qualität der Schule und einer positiven Beurteilung des Schulgebäudes ($r = .56/.82$), einer wichtigen Rolle der Schülermitverantwortung an der Schule ($r = .30/.55$) sowie einem positiven Verhältnis zu den Lehrkräften ($r = .54/.59$).

Noch stärker als auf der Schulebene ist für die Erhebungen zum *Unterricht* (Jahrgangsstufen: 8–10; Fächer: Mathematik, Deutsch und Englisch) eine erhebliche Streuung in der Untersuchungsstichprobe gegeben. Vergleichsweise gering ist die Varianz zwischen den in die Untersuchung einbezogenen drei *Jahrgangsstufen* (multivariat für 18 Skalen: 5,1% Varianz). *Innerhalb eines Faches und einer Jahrgangsstufe* liegen die Varianzanteile *zwischen den einzelnen Schulklassen* multivariat dagegen jeweils über 20%. Diese gravierenden Differenzen in der von den Schülern wahrgenommenen Unterrichtsqualität im Vergleich zwischen den Schulklassen resp. den einzelnen Lehrkräften sind nicht nur eine günstige Bedingung für die weiterführenden Analysen und vorgesehenen Interventionen, sondern ein schon an sich bemerkenswertes Ergebnis. Bestätigt wird damit außerdem die These (Ditton/Krecker 1995), dass die Unterschiede in der Unterrichtsqualität innerhalb einer Schule (auch innerhalb einer Klassenstufe und eines Faches) beträchtlich sind. Dies wiederum verweist auf die Möglichkeiten, durch *schulische Entwicklungsprozesse*, die auf eine engere Kooperation und einen intensivierten Erfahrungsaustausch im Kollegium abzielen (sharing best practice), zu Verbesserungen des Unterrichts beizutragen.

Ein wichtiger Stellenwert in der Diskussion um die Aussagekraft von *Schülerurteilen über den Unterricht* kommt der Frage nach deren *prognostischer Validität für die Leistungsentwicklung* zu. Hierzu ist zunächst auf Ergebnisse aus Analysen von Gruehn (2000) und Clausen (2000) zu verweisen. Beide Untersuchungen bestätigen, dass den (klassenweise aggregierten, d.h. gemittelten) Schülerwahrnehmungen zum Unterricht ein vergleichsweise hoher Prognosewert zukommt. Für die von uns verwendeten Skalen zur Ermittlung der Unterrichtsqualität werden diese Befunde durch eine Untersuchung an Schulen in Hamburg⁵ bestätigt (Ditton/Babic/Zehme i.Vorb.). Für diese Untersuchung stehen zusätzlich zu den Erhebungen zum Fachunterricht die Daten zur Leistungsentwicklung in den Fächern Deutsch, Mathematik und Englisch zwischen der siebten und neunten Jahrgangsstufe aus der Untersuchung zur Lernausgangslage und Leistungsentwicklung (LAU – Lehmann u.a. 1998) zur Verfügung. Ersten Analysen zufolge erklären die durch die Schülerbefragung erhobenen Unterrichtsfaktoren einen bedeutsamen Anteil der Varianz in der Leistungsentwicklung (Lehmann u.a. 2002).

Wie bereits ausgeführt, werden nicht nur die Schüler zum Fachunterricht befragt, sondern auch die Lehrkräfte zu den von ihnen erwarteten Ergebnissen dieser Befragung. Es zeigt sich, dass die *Übereinstimmung* zwischen den Schülerwahrnehmungen und den Lehrererwartungen bezüglich der einzelnen Dimensionen des Unterrichts unterschiedlich hoch ist. Zu den Aspekten Klarheit des Unterrichts, Strukturiertheit, Schwierigkeit und Zeitnutzung ist die Übereinstimmung eher gering ($r < .20$). Vergleichsweise hohe Übereinstimmungen ($r > .30$) finden sich für die Aspekte diagnostische Kompetenz der Lehrkraft ($r = .44$) und Verhältnis zwischen der Lehrkraft und der Klasse ($r = .33$). In der vorhergehenden Untersuchung waren zum Teil höhere Übereinstimmungen zwischen Schülern und Lehrkräften bezüglich der einzelnen Dimensionen

5 Hamburger Untersuchung zur Schul- und Unterrichtsqualität (HAUS), im Auftrag der Behörde für Jugend, Schule und Berufsbildung der Hansestadt Hamburg.

ermittelt worden (Ditton, i.Druck). Allerdings war hierbei der Befragungsteil für die Lehrkräfte differenzierter – und von diesen als zu umfangreich und zeitaufwändig kritisiert worden.

3.3 Akzeptanz und Verwertbarkeit der Rückmeldungen

Zusammen mit den Rückmeldungen erhielten die Schulen (Koordinatoren) und einzelnen Lehrkräfte Fragebögen zugesandt, in denen sie gebeten wurden, ihre Einschätzung zu Form, Umfang, Verständlichkeit und Brauchbarkeit der Rückmeldung abzugeben.

Insgesamt werden die Rückmeldungen positiv aufgenommen (Ditton 2001a, 2001b). Die eindeutige Mehrheit hält die Rückmeldung für ausreichend erläutert, verständlich dargestellt, informativ und übersichtlich. Eine ebenfalls sehr deutliche Mehrheit ist außerdem bereit, den Schülerfragebogen zum Unterricht auch in anderen Klassen und ggf. auch regelmäßig einzusetzen. Weit überwiegend, wenn auch mit gewissen Abstrichen, werden die Schüler als kompetente Beurteiler des Unterrichts angesehen und ihre Einschätzungen als „fair“ oder zumindest „eher fair“ (im Gegensatz zu „eher unfair“ bzw. „unfair“) bezeichnet.

Zwar ebenfalls positiv, aber nicht ganz so günstig fallen die Antworten auf die Fragen nach der Verwertbarkeit der Rückmeldungen für die weitere schulische bzw. Unterrichtsarbeit aus. Nur ca. zwei Drittel der Lehrkräfte äußern sich hier zufrieden. Dies weist auf den weiteren Entwicklungsbedarf.

4. Perspektiven

Wichtig für die weitere Arbeit und ein Ansatzpunkt für die Schulentwicklung ist die Untersuchung von Schule und Unterricht aus den Perspektiven von Lehrkräften und Schülern. Teilweise bestehen hier erhebliche Abweichungen, die besonders hinsichtlich der wechselseitigen Wahrnehmung im Unterricht differenziert zu untersuchen sind. Auf der Basis der Daten der ersten Projektphase ist es außerdem bereits möglich, die Übereinstimmung in der Wahrnehmung des Unterrichts *einer* Lehrkraft durch *mehrere* Schulkassen zu ermitteln. Mit der Fortführung der Untersuchung wird es darüber hinaus möglich, die Stabilität der Unterrichtswahrnehmungen unter Berücksichtigung der zwischenzeitlichen Interventionen durch die Feedbacks zu ermitteln.

Um die Möglichkeiten einer Weiterentwicklung des Feedbackverfahrens zu einem praktikablen System der Qualitätssicherung an Schulen zu eruieren, sind eine intensivierte Kooperation und ein erweitertes Methodenrepertoire – Gruppendiskussionen, Leitfadeninterviews – gefordert. Zur Vorbereitung der weiteren Zusammenarbeit fand im September 2001 ein Treffen mit Vertretern der Schulen (Koordinator und weitere zwei bis drei Kollegen pro Schule) statt. Schon hier hat sich gezeigt, dass im persönlichen Kontakt vielfältigere Anregungen gewonnen werden können als durch ein standardisiertes Befragungsinstrumentarium. Zusammenfassend können folgende Anregungen besonders herausgestellt werden:

- Es wurde der Wunsch geäußert, den Rückmeldungen eine *Gegenüberstellung der Schülerwahrnehmungen und Lehrererwartungen* zum Unterricht beizufügen. Damit sei eine Selbstkontrolle der Erwartungen leichter möglich.
- Die Darstellung von *Mittelwerten* in den Abbildungen wurde als anschaulich und gut nachvollziehbar bezeichnet. Teilweise wurden aber auch differenziertere Informationen gewünscht – z.B. Angaben zur Verteilung bzw. Streuung der Antworten und eine Kennzeichnung signifikanter Differenzen zur Vergleichsgruppe. Allerdings waren die Meinungen diesbezüglich durchaus unterschiedlich.
- Wie oben bereits erwähnt wurde, ist die der Untersuchung zugrunde liegende Auffassung zu Schul- und Unterrichtsqualität noch deutlicher herauszustellen, sofern möglich im Rahmen des „Pädagogischen Tags“ für das ganze Kollegium. Des Weiteren wurde gewünscht, die einzelnen Skalen der Unterrichtsbefragung über die verwendeten Erklärungen hinaus noch besser zu veranschaulichen.

Der Zentralaspekt ist allerdings die Frage, wie an den Schulen und von den einzelnen Lehrkräften mit der Rückmeldung umgegangen wird und wie die Ergebnisse konstruktiv genutzt werden können. Hier zeichnet sich nach den bisherigen Gesprächen eine breite Palette von Möglichkeiten ab, die in Abhängigkeit von individuellen Präferenzen äußerst differieren. Welche Kommunikations-, Beratungs- und Unterstützungssysteme an den Schulen gewünscht werden, scheint sehr unterschiedlich zu sein. Bezüglich der Unterrichtsergebnisse reicht die Palette der Verarbeitungsstrategien von der direkten Kommunikation der Ergebnisse mit den *Schülern* über die Besprechung der – mehr oder weniger anonymisierten Ergebnisse – in der *Fachgruppe* bzw. im ganzen *Kollegium* bis hin zum Wunsch nach der Beratung durch eine neutrale, *externe Instanz*. Hohe Akzeptanz scheint der Vorschlag zu finden, sich über die Ergebnisse in einem *Schulnetzwerk* (der Teilnehmerschulen) auszutauschen, weil hier bereits eine gemeinsame Gesprächsbasis gegeben ist und trotzdem genügend Distanz besteht, also sozusagen ein Mittelweg aus Transparenz und Anonymität gewählt werden kann. Dass sich damit im Gesamtergebnis vermutlich kein einheitliches Konzept für die weitere Entwicklung abzeichnen wird, mag auf den ersten Blick unbefriedigend erscheinen. Positiv gewendet zeigt sich in der Verweigerung von Einheitslösungen (*one size fits all*) möglicherweise aber auch das ohnehin vielfach geforderte plurale Selbstverständnis der Schulen sowie die Professionalität oder das Selbstbewusstsein einer Berufsgruppe.

Literatur

- Balk, M. (2000): Die Wirkung von Evaluationsfeedback und der Einfluss von Selbstwirksamkeitserwartungen auf Handlungsparameter von Lehrpersonen. Frankfurt am Main: Lang.
- Clausen, M. (2000): Wahrnehmung von Unterricht. Übereinstimmung, Konstruktvalidität und Kriteriumsvalidität in der Forschung zur Unterrichtsqualität. Dissertation: Freie Universität Berlin.
- Ditton, H. (2000a): Qualitätskontrolle und -sicherung in Schule und Unterricht – ein Überblick zum Stand der empirischen Forschung. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualitätssicherung im Bildungsbereich. Beiheft Nr. 41 der Zeitschrift für Pädagogik. Beltz: Weinheim, S. 73–92.

- Ditton, H. (2000b): Elemente eines Systems der Qualitätssicherung im schulischen Bereich. In: Weis-
haupt, H. (Hrsg.): Qualitätssicherung im Bildungswesen. (Reihe: Erfurter Studien zur Entwicklung
des Bildungswesens). Erfurt: Pädagogische Hochschule, S. 13–35.
- Ditton, H. (2001a): Der Beitrag der empirischen Schulforschung zur Qualitätssicherung an Schulen.
Forschungslö-
gische Aspekte und Anwendungsbezüge. In: Hansel, T. (Hrsg.): Schulprofil und
Schulqualität. Perspektiven der aktuellen Schulreformdebatte. Herbolzheim: Centaurus-Verlag, S.
70–98.
- Ditton, H. (2001b): Implikationen des Begriffs Schulqualität. In: Dokumentation der BLK-Früh-
jahrstagung zur QuiSS-Frühlingsakademie, Bonn, 2001, S. 1–16.
- Ditton, H. (2002): Lehrkräfte und Unterricht aus Schülersicht. Ergebnisse einer Untersuchung im Fach
Mathematik. In: Zeitschrift für Pädagogik, 48, S. 262–286.
- Ditton, H. (i.Druck): Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspekti-
ven. Erscheint in: Unterrichtswissenschaft.
- Ditton, H./Babic, B./Zehme, M. (i.Vorb.): HAUS – Hamburger Untersuchung zur Schulqualität. Be-
hörde für Schule, Jugend und Berufsbildung (BSJB). Freie Hansestadt Hamburg.
- Ditton, H./Edelhäuser, T./Merz, D. (2001): Erweiterte Selbstverantwortung im Urteil von Lehrkräften
und Schulleitungen. In: Die Deutsche Schule, 93, S. 210–222.
- Ditton, H./Krecker, L. (1995): Qualität von Schule und Unterricht – Empirische Befunde zu Fragestel-
lungen und Aufgaben der Forschung. In: Zeitschrift für Pädagogik, 41, S. 507–529.
- Ditton, H./Merz, D./Edelhäuser, T. (i.Druck): Einstellungen von Lehrkräften und Schulleiter/innen zu
zentralen Testuntersuchungen an Schulen. Erscheint in: Empirische Pädagogik, Frühjahr 2002
- Einsiedler, W. (1997a): Unterrichtsqualität und Leistungsentwicklung: Literaturüberblick. In: Weinert,
F.E./Helmke, A. (Hrsg.): Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union,
S. 225–240.
- Einsiedler, W. (1997b): Research on Instructional Methods: A European Perspective. In: Tennyson,
R.D./Schott, F./Seel, N./Dijkstra, S. (Hrsg.): Instructional design: International perspectives, Vol. 1:
Theory, research and models. London: Lawrence Erlbaum 1997, S. 269–291.
- Fitz-Gibbon, C.T. (1996): Monitoring Education. Indicators, Quality and Effectiveness. London: Cas-
sell.
- Frese, M.; Zapf, D. (1994): Action as the Core of Work Psychology: A German Approach. In: Triandis,
H.C./Dunnette, M.D./Hough, L.M. (Hrsg.): Handbook of Industrial and Organizational Psycholo-
gy, Vol. 4 (2nd ed., pp. 271–340). Palo Alto, Ca: Consulting Psychologists Press.
- Gruehn, S. (2000): Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschrei-
bung. Münster: Waxmann.
- Heinze, T. (1975): Handlungsforschung im pädagogischen Feld. München: Juventa.
- Helmke, A./Weinert, F.E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Enzyklopädie der
Psychologie, Serie I, Band 3 (Pädagogische Psychologie). Göttingen, Bern u.a.
- Klieme, E./Baumert, J. (2001): TIMMS als Startpunkt für Qualitätssicherung und Qualitätsentwick-
lung im Bildungswesen. In: BMBF (Hrsg.): TIMSS – Impulse für Schule und Unterricht. Bonn:
BMBF Publik. S. 5–9.
- Klieme, E./Baumert, J./Schwippert, K. (2000): Schulbezogene Evaluation und Schulleistungsvergleiche:
eine Studie im Anschluss an TIMSS. In: Rolff, H.G./Bos, W./Klemm, K./Pfeiffer, H./Schulz-Zander,
R. (Hrsg.): Jahrbuch der Schulentwicklung (Bd.11, S.387–420). Weinheim: Juventa.
- Klieme, E./Schümer, G./Knoll, S. (2001): Mathematikunterricht in der Sekundarstufe I: „Aufgabenkul-
tur“ und Unterrichtsgestaltung. In: BMBF (Hrsg.): TIMSS – Impulse für Schule und Unterricht.
Bonn: BMBF Publik. S. 43–57.
- Kluger, A.N./DeNisi, A. (1996): The Effects of Feedback Interventions on Performance: A Historical
Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. In: Psychological Bulle-
tin, Vol. 119, No. 2, S. 254–284.

- Lehmann, R.H./Gänsfuß, R./Peek, R. (1998): Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern an Hamburger Schulen. Bericht über die Untersuchung im September 1998. Hamburg: Behörde für Schule, Jugend und Berufsbildung, Amt für Schule Hamburg.
- Lehmann, R.H./Peek, R./Gänsfuß, R./Husfeldt, V. (2002): Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9. Ergebnisse einer Längsschnittuntersuchung in Hamburg. Hamburg: Behörde für Bildung und Sport.
- Mortimore, P./Sammons, P./Stoll, L./Lewis, D./Ecob, R. (1988): *School Matters. The Junior Years*. Somerset: Open Books.
- Scheerens, J./Bosker, R. (1997): *The Foundations of Educational Effectiveness*. Oxford: Pergamon.
- Slavin, R.E. (1996): *Education for all*. Lisse: Swets & Zeitlinger.
- Teddlie, C./Reynolds, D. (2000): *The International Handbook of School Effectiveness Research*. London; New York: Falmer Press.
- Teddlie, T.E./Stringfield, S. (1993): *Schools make a difference. Lessons learned from a 10-year study of school effects*. New York: Teachers College Press.

Anschrift der Autoren:

Prof. Dr. Hartmut Ditton, Lehrstuhl für Allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung, Institut für Pädagogik, Universität München, Leopoldstraße 13, 80802 München.

Dipl.-Päd. Bettina Arnoldt, Lehrstuhl für Allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung, Institut für Pädagogik, Universität München, Leopoldstraße 13, 80802 München.

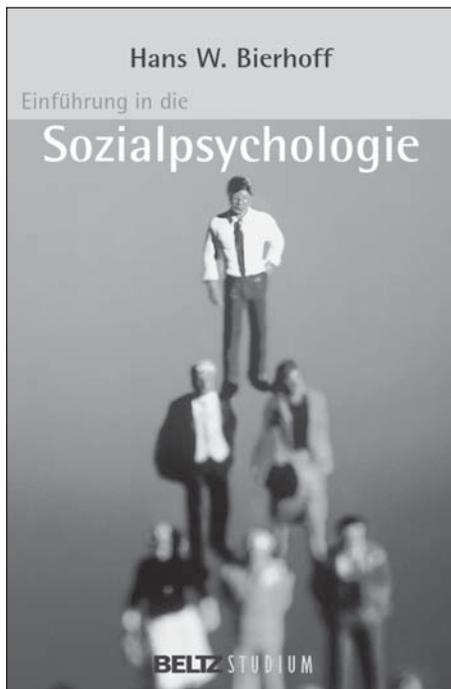
Eva Bornemann, M.A., Lehrstuhl für Allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung, Institut für Pädagogik, Universität München, Leopoldstraße 13, 80802 München.

Reihe »Beltz Studium«



Wolfgang Frindte
**Einführung in die
Kommunikationspsychologie**
2002. 232 Seiten. Broschiert.
ISBN 3-407-25254-4

Kommunikationspsychologie befasst sich mit der Kommunikation zwischen unterschiedlich komplexen personalen und sozialen Systemen. Dabei geht es z.B. um den sprachlichen und nicht-sprachlichen Austausch in Zweierbeziehungen, um Gruppenkommunikation, um interkulturelle und medial vermittelte (medierte) Kommunikation.



Hans W. Bierhoff
**Einführung in die
Sozialpsychologie**
2001. 208 Seiten. Broschiert.
ISBN 3-407-25251-X

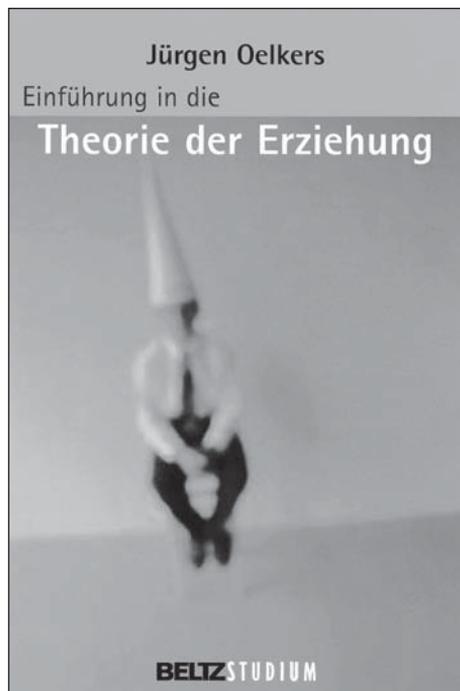
Eine problemorientierte Einführung in die Fragestellungen der Sozialpsychologie. Zur Beantwortung wird auf soziale Eindrucksbildung und soziales Erinnern, Vorurteile gegenüber sozialen Gruppen und Geschlechtern eingegangen.

Infos und Ladenpreise: www.beltz.de

BELTZ

Beltz Verlag · Postfach 100154 · 69441 Weinheim

Reihe »Beltz Studium«



Jürgen Oelkers
**Einführung in die
Theorie der Erziehung**

2001. 294 Seiten. Broschiert.
ISBN 3-407-25236-6

Bis weit ins 20. Jahrhundert hinein sollte mit »Erziehung« der neue Mensch hervorgebracht werden, die utopische Gesellschaft entstehen oder allgemeine Emanzipation auf den Weg gebracht werden. Die Einführung analysiert thematische Koppelungen, die die heutige Diskussion bestimmen.



Ralf Vollbrecht
**Einführung in die
Medienpädagogik**

2001. 239 Seiten. Broschiert.
ISBN 3-407-25234-X

Eine Einführung in medienpädagogisches Denken, in medienwissenschaftliche Theorien und grundlegende empirische Erkenntnisse über Medienutzung.

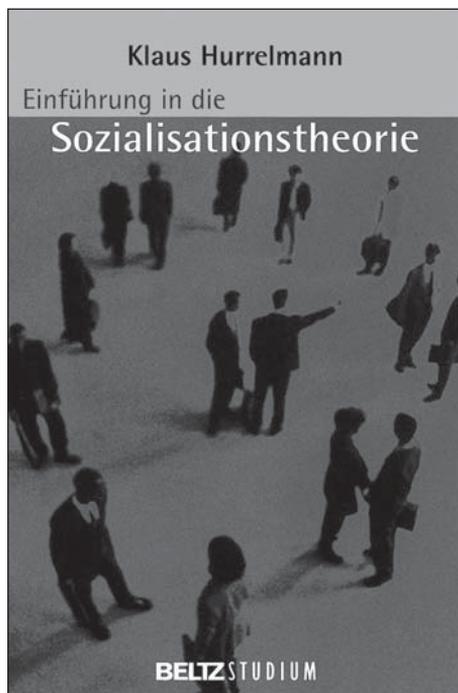
Infos und Ladenpreise: www.beltz.de

BELTZ

Beltz Verlag · Postfach 100154 · 69441 Weinheim

F0128

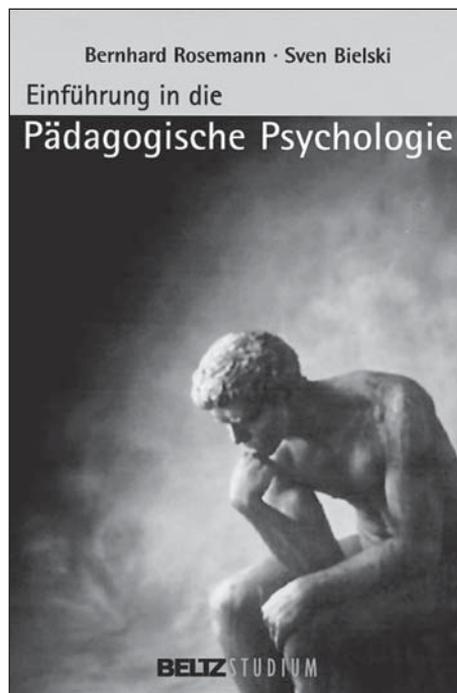
Reihe »Beltz Studium«



Klaus Hurrelmann
**Einführung in die
Sozialisationsforschung**

8., völlig überarb. Auflage 2002.
301 Seiten. Broschiert.
ISBN 3-407-25235-8

Der Autor zeichnet die wichtigsten Theorien und Methoden der Sozialisationsforschung in leicht verständlicher Form nach und stellt die wesentlichen Untersuchungsergebnisse zur Sozialisation in Familien, Erziehungs- und Bildungssystemen, Gleichaltrigen- und Medien zusammen.



Bernhard Rosemann / Sven Bielski
**Einführung in die
Pädagogische Psychologie**

2001. 207 Seiten. Broschiert.
ISBN 3-407-25238-2

Der Band bietet eine an der Praxis orientierte Einführung in relevante Teilgebiete der Pädagogischen Psychologie.

Infos und Ladenpreise: www.beltz.de

BELTZ

Beltz Verlag · Postfach 100154 · 69441 Weinheim