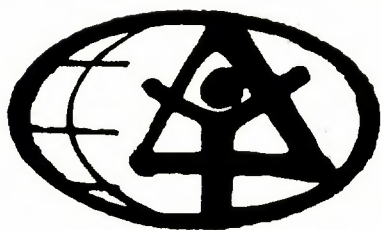


Forschungen zur Didaktik des Sachunterrichts, Band 5

Kay Spreckelsen / Kornelia Möller /  
Andreas Hartinger (Hrsg.)

# Ansätze und Methoden empirischer Forschung zum Sachunterricht



**KLINKHARDT**



ANSÄTZE UND METHODEN  
EMPIRISCHER FORSCHUNG  
ZUM SACHUNTERRICHT

FORSCHUNGEN ZUR DIDAKTIK DES  
SACHUNTERRICHTS

herausgegeben von  
Walter Köhnlein und Helmut Schreier

BAND 5



ANSÄTZE UND METHODEN  
EMPIRISCHER FORSCHUNG  
ZUM SACHUNTERRICHT

herausgegeben von  
Kay Spreckelsen, Kornelia Möller und Andreas Hartinger

**KLINKHARDT**

2002

---

VERLAG JULIUS KLINKHARDT • BAD HEILBRUNN / OBB.

## Schriftenreihe der Gesellschaft für Didaktik des Sachunterrichts e.V.

GD Die Gesellschaft für Didaktik des Sachunterrichts (GDSU) e.V. ist ein Zusammenschluß von  
SU Lehrenden aus Hochschule, Lehrerfort- und -weiterbildung und Schule. Ihre Aufgabe ist die  
Förderung der Didaktik des Sachunterrichts als wissenschaftlicher Disziplin in Forschung und  
Lehre sowie die Vertretung der Belange des Schulfaches Sachunterricht.

Die Deutsche Bibliothek – Cip-Einheitsaufnahme

Ein Titelsatz für diese Publikation ist bei  
Der Deutschen Bibliothek  
erhältlich

2002.3.r. © by Julius Klinkhardt.

Das Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt.  
Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung  
des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen,  
Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in  
elektronischen Systemen.

Druck und Bindung:  
WB-Druck, Rieden

Printed in Germany 2002

Gedruckt auf chlorfrei gebleichtem alterungsbeständigem Papier

ISBN 3-7815-1187-1

# Inhalt

Vorwort .....	7
Einleitung.....	9
<b>Zum Stand der Forschung</b>	
<i>Wolfgang Einsiedler</i> Empirische Forschung zum Sachunterricht – ein Überblick .....	17
<b>Forschungsmethoden und Erhebungsverfahren</b>	
<i>Hans-Günther Roßbach</i> Zum Design empirischer Untersuchungen in der Lehr-Lern-Forschung .....	41
<i>Philipp Mayring</i> Qualitative Analyseansätze in der Lehr-Lern-Forschung .....	59
<i>Jürgen Rost</i> Qualitative und Quantitative Methoden in der fachdidaktischen Forschung .....	71
<i>Richard White and Richard Gunstone</i> Assessment of Understanding of Science in Elementary School .....	91
<b>Empirische Befunde zum Lehren und Lernen im Sachunterricht</b>	
<i>Beate Sodian und Claudia Thoermer</i> Naturwissenschaftliches Denken im Grundschulalter. Die Koordination von Theorie und Evidenz .....	105
<i>Elsbeth Stern, Ilonca Hardy, und Susanne Koerber</i> Die Nutzung graphisch-visueller Repräsentationsformen im Sachunterricht .....	119
<i>Kay Spreckelsen</i> Nachdenken über physikalische Probleme Eine Untersuchung zum A/K-Wert in der Begegnung von Grundschulern mit physikalischen Phänomenen .....	133
<i>Hartmut Giest</i> Zur Entwicklung des begrifflichen Denkens im Grundschulalter .....	145

<i>Achim Engelen, Angela Jonen und Kornelia Möller</i> Lernfortschrittsdiagnosen durch Interviews – Ergebnisse einer Pilotstudie zum „Schwimmen und Sinken“ im Sachunterricht der Grundschule .....	155
<i>Andreas Hartinger</i> Selbstbestimmungsempfinden in offenen Lernsituationen Eine Pilotstudie zum Sachunterricht .....	174
<i>Hanns Perillon und Herrmann Laux</i> Soziale Beziehungen zwischen Grundschulkindern – empirische Befunde zu einem wichtigen Thema des Sachunterrichts .....	185
<i>Beate Blaseio</i> Inhaltsstruktur und Tendenzen der Inhalte im Sachunterricht Eine empirische Bestandsaufnahme der Inhalte des Sachunterrichts in den 70er, 80er und 90er Jahren anhand von Unterrichtslehrwerken .....	205
Autorinnen und Autoren .....	223

# Vorwort

Mit ihrer Buchreihe *Forschungen zur Didaktik des Sachunterrichts* legt die Gesellschaft für Didaktik des Sachunterrichts (GDSU) der pädagogisch interessierten Öffentlichkeit Ansätze und Erträge aus einer jungen akademischen Disziplin vor, die sich bemüht, das Lehren und Lernen in einem zentralen Bereich der Grundschularbeit nachhaltig zu verbessern. Gekennzeichnet ist diese Forschungsarbeit durch eine interdisziplinäre Kooperation mit Wissenschaftsbereichen, die ihrerseits zur Klärung von Voraussetzungen, Bedingungen und Möglichkeiten schulischer Bildungs- und Erziehungsarbeit beitragen. Besonderes Gewicht erhalten dabei neue Ergebnisse der Kognitionspsychologie und Theorien der Wissensrepräsentation.

Die bisher erschienenen vier Bände dieser Reihe haben ihre Schwerpunkte bei konzeptionellen und inhaltlichen Fragen des Sachunterrichts:

Kinder auf dem Wege zum Verstehen der Welt (1997)

Interessenförderung (1997)

Vielperspektivisches Denken im Sachunterricht (1999)

Innovation Sachunterricht – Befragung der Anfänge nach zukunftsfähigen Beständen (2001).

Der vorliegende Band über Ansätze und Methoden empirischer Forschung zum Sachunterricht nimmt die fachdidaktische Forschung selbst in den Blick. Den Herausgebern des Bandes ist es gelungen, namhafte Wissenschaftlerinnen und Wissenschaftler aus der Didaktik des Sachunterrichts selbst sowie aus benachbarten Disziplinen zu gewinnen, die über Bestandsaufnahmen hinaus weiterführende Perspektiven der Bildungsforschung eröffnen.

Die ersten Teile des Buches geben zunächst einen Überblick über empirische Untersuchungen zum Sachunterricht und informieren über den gegenwärtigen Forschungsstand. In mehreren Beiträgen wird das aktuell verfügbare forschungsmethodische Instrumentarium unter dem Gesichtspunkt seiner Ergiebigkeit für die spezifische fachdidaktische Arbeit dargestellt und reflektiert. Dabei ist die Spannweite von qualitativen und quantitativen Methoden berücksichtigt; Anregungen für die Anlage und Durchführung einschlägiger Untersuchungen werden gegeben. Im dritten Teil wird am Beispiel empirischer Studien zu verschiedenen Inhaltsbereichen des Sachunterrichts Einblick in Forschungswerkstätten gegeben und konkret aufgezeigt, wie sich eine verbesserte Kenntnis der Vorerfahrungen und des

Denkens von Kindern auf die Qualität und den Ertrag eines substantiellen Sachunterrichts auswirkt. Mit solchen Kenntnissen lassen sich dann auch Unterrichtskonzepte und -strategien bewerten und optimieren. Damit wendet sich das Buch auch an Lehrerinnen und Lehrer und unterstützt sie bei der Weiterentwicklung ihrer Arbeit in der Praxis.

Dank gebührt den Herausgebern dieses Bandes, den Autorinnen und Autoren sowie dem Verlag Klinkhardt für die Förderung der Sache des Sachunterrichts.

*Für die Reihenherausgeber*

*Walter Köhnlein*

## Einleitung

„Das Kennzeichen von wissenschaftlichen Aussagen ist die Methodisierung der Erfahrungsgewinnung“. Dieser Satz, mit dem Hans-Günther Roßbach seinen Beitrag für diesen 5. Forschungsband der Gesellschaft für Didaktik des Sachunterrichts einleitet, thematisiert das Anliegen des vorliegenden Forschungsbandes, Methoden der wissenschaftlichen Erkenntnisgewinnung zum Gegenstand der Reflektion zu machen. Die Didaktik des Sachunterrichts ist eine junge akademische Disziplin. Schon deshalb erscheint es uns erforderlich, neben der „Sache“ des Sachunterrichts auch Forschungsmethoden in der Didaktik des Sachunterrichts eingehender zu diskutieren.

Der vorliegende Forschungsband bemüht sich, einen Beitrag zu dieser Diskussion zu leisten, indem (im Unterschied zu den bislang vorliegenden vier Forschungsbänden), nicht mehr eine inhaltliche Fragestellung, sondern *eine* Form der wissenschaftlichen Erkenntnisgewinnung die Klammer für die einzelnen Beiträge bildet – die *empirische Forschung*. Es sollen an dieser Stelle nicht die Möglichkeiten und Grenzen empirischer Forschung diskutiert werden; die Empirie ist unserer Ansicht nach *eine* sinnvolle Forschungstradition unter anderen möglichen Traditionen. Will man Erkenntnisse über Prozesse des Lernens und Lehrens gewinnen, so sind empirische Forschungsmethoden unserer Ansicht nach unverzichtbare Instrumente, die angemessen und theoriegeleitet eingesetzt werden müssen.

Der Forschungsband gliedert sich in drei verschiedene Teile:

Zunächst legt Wolfgang Einsiedler eine Bestandsaufnahme über die aktuelle Forschung zum Sachunterricht vor; diese zeigt, dass sich auch in unserer jungen Disziplin bereits eine breite Vielfalt von Forschungsgegenständen und -methoden etablieren konnte. Im zweiten Teil werden unterschiedliche empirische Ansätze und ihr Verhältnis zueinander thematisiert; qualitative und quantitative Forschung werden dabei nicht als unversöhnliche Gegensätze, sondern als Ansätze mit unterschiedlicher Forschungsrichtung und -leistung dargestellt. Teil drei des Bandes beinhaltet Beiträge zu aktuellen empirischen Arbeiten.

Gezielt haben wir nicht nur Kolleginnen und Kollegen um Beiträge gebeten, deren Forschungsschwerpunkt in der Didaktik des Sachunterrichts im engeren Sinne liegt, sondern auch Wissenschaftlerinnen und Wissenschaftler aus benachbarten Disziplinen – in erster Linie aus der pädagogischen Psychologie. In diesen Nachbardisziplinen existieren elaborierte (empirische) Forschungstraditionen, die unserer Meinung nach Forschungen zur Didaktik des Sachunterrichts inhaltlich und

methodisch befruchten können. Allerdings – dies ist selbstverständlich – dürfen die Eigenheiten und die spezifischen Fragen und Problemstellungen des Sachunterrichts dabei nicht aus dem Blick geraten. Um nicht nur aus benachbarten Disziplinen, sondern auch aus benachbarten Ländern Anregungen zu erhalten, haben wir mit dem Artikel von Richard White und Richard Gunstone einen Beitrag aufgenommen, der aus der angloamerikanischen Forschungstradition stammt.

Zu den einzelnen Beiträgen:

Der Artikel von *Wolfgang Einsiedler* gibt einen Überblick über die aktuelle Forschungslage im Sachunterricht. Unter Rückgriff auf Studien aus der ‚Anfangszeit‘ der Sachunterrichtsforschung unterteilt Einsiedler in verschiedene Forschungsrichtungen und -traditionen (Unterrichtsmethodenforschung, Forschung zu Schülervoraussetzungen sowie Lehrstoff- und Lehrplananalysen). Die zentralen Ergebnisse werden berichtet; anschließend klärt Einsiedler die jeweilige Bedeutung der Forschungsansätze für die Weiterentwicklung des Faches Sachunterricht.

Der zweite – methodische – Teil unseres Bandes beginnt mit dem Beitrag von *Hans-Günther Roßbach*, der exemplarisch bewährte Formen der quantitativen Lehr-Lern-Forschung darstellt; dabei werden auch Schwierigkeiten und Grenzen quantitativer Forschungsarbeit aufgezeigt. Wie man mit solchen Problemen umgehen kann, wird zunächst allgemein geklärt und anschließend an einem Beispiel expliziert. Der Beitrag eignet sich in besonderer Weise als Anregung für eigene Forschungsarbeiten, ist jedoch auch von Interesse, wenn man einen Einblick in das Design empirischer (Vergleichs-)Studien gewinnen möchte.

Im Artikel von *Philipp Mayring* stehen die qualitativen Analyseverfahren – wie z. B. die Inhaltsanalyse – im Vordergrund. Gleichzeitig wird jedoch auch dargestellt, wie diese qualitativen Analyseansätze mit quantitativen Forschungsmethoden verknüpft werden können. Das Verfahren wird an einem konkreten Beispiel aus eigener Forschung illustriert.

Auch *Jürgen Rost* behandelt in seinem Beitrag die Notwendigkeit eines Zusammenspiels qualitativer und quantitativer Forschungsmethoden. Er stellt die Forderung auf, dass bei der Entwicklung qualitativer Untersuchungsdesigns dieselben Prinzipien bei der Versuchsplanung zum Tragen kommen müssen wie in der quantitativen Forschung (z.B. Vergleichsgruppen-Designs, Kontrolle von Störvariablen, Identifizierung von Moderatorvariablen).

Die Beiträge von Rost, Mayring und Roßbach basieren auf forschungsmethodischen Vorträgen, die auf Jahrestagungen der GDSU im Forum der Kommission „Forschung/Nachwuchsförderung“ in den letzten Jahren gehalten wurden. Die Auswahl der Vortragenden wurde dabei von der Idee geleitet, je einen Vortrag zu quantitativen Methoden (Roßbach), zu qualitativen Methoden (Mayring) und



zur Verknüpfung der Methoden (Rost) anzubieten. Über diese jeweiligen Schwerpunkte hinausgehend, betonen alle drei Autoren die Bedeutung und Notwendigkeit einer Integration von qualitativen und quantitativen Methoden in der empirischen Forschung.

Es ist eine zentrale Schwierigkeit vieler empirischer Studien, dass zwar einfache kognitive Effekte des Unterrichts leicht erhoben werden können, große Schwierigkeiten allerdings bestehen, wenn Verständnis oder Emotionen erfasst werden sollen. Gerade im Sachunterricht, in dem das Verstehen des Gelernten und die Fähigkeit zur Anwendung zentrale Ziele darstellen, kann sich die empirische Forschung diesem Problem aber nicht entziehen. *Richard White* und *Richard Gunstone* beschreiben in ihrem Artikel verschiedene Erhebungsverfahren, mit denen Verstehensprozesse erfasst werden können. Dieser Beitrag ist nicht nur für Forschergruppen interessant; auch für Lehrkräfte kann er von besonderem Interesse sein, da sich fast alle dargestellten Verfahren nicht nur für wissenschaftliche Studien, sondern auch für informelle schulische Leistungsevaluation eignen.

Im dritten Teil dieses Bandes werden einzelne empirische Forschungsvorhaben dargestellt. Dieser Teil beginnt mit dem Beitrag von *Beate Sodian* und *Claudia Thoermer*. In der hier dargestellten Studie wird aus entwicklungspsychologischer Perspektive die Frage untersucht, inwieweit Kinder im Grundschulalter schon in der Lage sind, wissenschaftlich zu denken, indem sie z. B. Kausalhypothesen erkennen und bewerten oder zwischen Hypothesen und der Evidenz eines Ereignisses unterscheiden. Die Ergebnisse zeigen, dass viele Schülerinnen und Schüler der vierten Jahrgangsstufe durchaus adäquate Verfahren für die Prüfung von Hypothesen identifizieren und Theoriekonflikte erkennen. Aufgabe der Sachunterrichtsforschung könnte es sein, Anlage und Methode dieser Untersuchung an sachunterrichtsspezifischen Inhalten zu replizieren. Aber auch für die Unterrichtspraxis lassen sich aus dieser Studie Anregungen gewinnen, da auf unterschiedliche Levels in den Argumentationsweisen von Kindern aufmerksam gemacht wird.

Der Artikel von *Elsbeth Stern*, *Ilonca Hardy* und *Susanne Koerber* beschäftigt sich mit dem Begriff des naturwissenschaftlichen ‚Verständnisses‘ aus der Sicht der Kognitionspsychologie. Aus dieser Sicht ist Verstehen dann gegeben, wenn ein Sachverhalt symbolisch, extern oder intern, repräsentiert werden kann. Wenn die symbolische Repräsentation von Sachverhalten für (naturwissenschaftliches bzw. mathematisches) Verstehen von Bedeutung ist, stellt sich die Frage, ob bereits in der Grundschule verstehendes Lernen durch den Einsatz geeigneter Repräsentationsformen gefördert werden kann. Das Interesse der Autorinnen richtet sich dabei speziell auf den Einsatz bisher nur wenig erforschter graphisch-visueller Repräsentationen. Es werden zwei Studien dargestellt, die Effekte verschiedener Repräsentationsformen auf das Verständnis der Schülerinnen und Schüler (für Pro-

portionen sowie für lineare Zusammenhänge) nachweisen konnten.

Es zeigte sich in verschiedenen Unterrichtsvorschlägen, dass das Ausgehen von Phänomenen ein fruchtbarer methodischer Ansatzpunkt im Sachunterricht der Grundschule ist. Im Unterricht ergeben sich jedoch hier auch Schwierigkeiten für Lehrerinnen und Lehrer. Wenn sich Kinder mit (naturwissenschaftlichen) Phänomenen beschäftigen, so ist es z. B. schwierig, anhand der verbalen Äußerungen zu erkennen, wie genau sie das Grundprinzip des Phänomens verstanden haben. Anhand der Beschreibung und Erklärung verschiedener physikalischer Phänomene wird von *Kay Spreckelsen* ein Verfahren vorgestellt, mit dem abstrakte und konkrete Äußerungen von Schülerinnen und Schüler miteinander „verrechnet“ werden können. In mehreren empirischen Untersuchungen wird dieses Verfahren eingesetzt und präzisiert, wobei sich u. a. interessante Subgruppenunterschiede zeigen. In der empirischen Studie von *Achim Engelen*, *Angela Jonen* und *Kornelia Möller* geht es um die Frage, inwieweit moderat-konstruktivistische Lehr-Lernumgebungen dazu beitragen können, qualitative Lernfortschritte hinsichtlich des Verstehens von Konzepten (hier Auftrieb bzw. Dichte) zu erzielen. Dabei ist insbesondere zu prüfen, ob es sich bei dem angeeigneten Wissen um wirklich verstandenes, anwendungsfähiges Wissen handelt und nicht nur um träges Wortwissen. Die Autoren verwenden zur Diagnose kognitiver Lernfortschritte eine Klassifizierung von Interviewäußerungen nach Konzeptebenen und zeigen, dass es ihnen damit gelingt, ein quantifizierbares Verfahren für die Auswertung inhaltsreicher, „weicher“ Daten zu gewinnen.

Die Untersuchung von *Hartmut Giest* leistet einen Beitrag zur entwicklungspsychologischen Fundierung des Wissenserwerbs im Sachunterricht, indem sie charakteristische Phasen in der Entwicklung für den Sachunterricht prototypischer Begriffe („Arbeit“ und „Pflanze“) im Rahmen zweier Längsschnittuntersuchungen behandelt. Wie in dem Beitrag von *Engelen et. al.* verwendet auch *Giest* eine Methode zur Ordnung der begrifflichen Struktur nach Konzeptebenen („begriffliche Stadien“). Die Analysen beziehen sich auf zwei Längsschnittuntersuchungen in den Jahren 1988-1991 sowie 1997-2000. Im Ergebnis werden Trends der begrifflichen Entwicklung in Richtung auf wissenschaftliche Begriffsbildungen hin beschrieben.

Es gibt zwar inzwischen einige Studien über die Effekte offener Unterrichtsformen auf Lernergebnisse, Motivation und Interesse von Schülerinnen und Schüler im Sachunterricht der Grundschule. Allerdings existieren noch nahezu keine Ergebnisse darüber, wie diese Schülerinnen und Schüler offenen Unterricht wahrnehmen. *Andreas Hartinger* stellt eine Studie vor, in der die ‚objektiv beobachtbare‘ und die ‚subjektiv empfundene‘ Selbstbestimmung in Beziehung gebracht und ihre Auswirkungen auf Mitarbeit im Unterricht und Interesse dargestellt werden. Dabei zeigen sich zwar Überschneidungen, aber durchaus auch unabhängige Effekte.

Das Leben in einer Gemeinschaft ist ein zentrales Thema des Sachunterrichts. Gleichzeitig sind die sozialen Beziehungen in der Klasse für die Kinder ein – wenn nicht sogar der – zentrale Aspekt von Schule. In dem Beitrag von *Hanns Petillon* und *Hermann Laux* werden die wichtigsten Befunde (auch aus eigenen Forschungen) über soziale Beziehungen zwischen Grundschulkindern – gegliedert nach den verschiedenen Perspektiven ‚Kind‘, ‚Interaktion‘ und ‚Gruppe‘ – dargestellt und ihre Bedeutung für den Sachunterricht geklärt.

In dem Beitrag von *Beate Blaseio* geht es einerseits um die Erfassung der Inhaltsstruktur des Sachunterrichts, soweit sie aus gängigen Sachunterrichtswerken (Schulbüchern) erschlossen werden kann, und andererseits um die inhaltlichen Tendenzen des Faches, die sich seit der Ablösung der Heimatkunde vor dreißig Jahren abzeichnen. Auch wenn man nicht der Ansicht ist, dass Unterrichtswerke die tatsächliche Unterrichtswirklichkeit und darüber hinaus die gesamte Sachunterrichtsdiskussion widerspiegeln, so wird man doch die mit dieser Arbeit gewonnenen Befunde aufmerksam rezipieren müssen, zumal die Genehmigungspraxis von Schulbüchern sich allgemein an den jeweils seitens der Kultusministerien erlassenen Richtlinien orientiert.

Wenn hiermit seitens der GDSU im Rahmen der Reihe ihrer Forschungsbände bereits der fünfte Band vorgelegt wird, so mag dies auch als Ausweis der Bedeutung gelten, die die GDSU der Forschungsthematik im Rahmen ihrer Aktivitäten beimisst. Der Aufgabe der Didaktik des Sachunterrichts entsprechend, sich als universitäre Disziplin in der Doppelfunktion von Forschung und Lehre auszuweisen, hat die GDSU schon seit längerem eine ständige Kommission „Nachwuchsförderung/Forschung“ eingerichtet, von deren Mitgliedern der vorliegende Band als Herausgeber verantwortet wird. Unser Dank gilt den Kolleginnen und Kollegen, die zu diesem Band freundlich und geduldig beigetragen haben, dem Klinkhardt-Verlag für seine gewohnt kooperative Herstellung und der Universität/Gesamthochschule Kassel für einen sehr hilfreichen Druckkostenzuschuss.

*Kornelia Möller, Andreas Hartinger, Kay Spreckelsen*

Oktober 2001



# Zum Stand der Forschung



## Empirische Forschung zum Sachunterricht – ein Überblick

### 1. Einleitung

Sachunterricht der Grundschule ist ein junges Fach. Es wurde in der Bundesrepublik zwischen ca. 1969 und 1973 eingeführt. In den 70er Jahren konzentrierten sich die wissenschaftlichen Auseinandersetzungen mit dem Fach auf konzeptionelle Fragen. Eigene Professuren für Sachunterricht (und das relativ wenige) wurden erst ab ca. 1979 eingerichtet. So ist es zu verstehen, dass in den 70er Jahren trotz der erheblichen Innovationen und des daraus entstandenen Forschungsbedarfs nur wenig empirische Forschung zum Sachunterricht durchgeführt wurde. Die einzige größere Evaluationsstudie zum naturwissenschaftlichen Sachunterricht hat Spreckelsen initiiert (Adler, Cronjäger, Murkisch, Spreckelsen & Ulrich 1973); in den USA dagegen kam es zu Dutzenden von Untersuchungen vor allem zu den Science-Curricula, aber auch zum Bereich der „Social Studies“ (vgl. die Metaanalyse von Bredderman 1983).

Die später einsetzende Sachunterrichtsforschung ist sehr heterogen und es fällt schwer, für eine referierende Darstellung einen „roten Faden“ zu finden. Im Folgenden werden drei Schwerpunkte gesetzt:

- Unterrichtsmethodenforschung/Visualisierung: Die Forschungsfragen in diesem Bereich greifen Probleme auf, die schon im alten „Anschauungsunterricht“ erörtert wurden; und sie sind insofern hochaktuell geblieben, als auch in jüngster Zeit diskutiert wird, wie wichtig die konkrete Anschauung etwa von Naturphänomenen ist und welchen Stellenwert dabei die Herausarbeitung abstrakter Strukturen hat.
- Schülervorstellungen/Conceptual Change: Während der erste Problemkreis in die Allgemeine Didaktik hineinreicht, geht es bei der Analyse von Schülervorstellungen um genuin sachunterrichtliche Fragestellungen. Welche Konzepte bringen Schüler mit und wie verändern sie sich durch Unterricht? Mit dem „Conceptual Change“-Ansatz konnten dabei fruchtbare Anregungen aus der amerikanischen Forschung übernommen werden.
- Lehrstoff- und Lehrplananalysen: Hier handelt es sich nicht um Forschung im unterrichtlichen Feld, sondern um Dokumentenanalysen. Die Ergebnisse dieses Forschungsstranges bringen jedoch wichtige Informationen über Unterrichtspraxis und geben Hinweise zur Lehrplangestaltung.

## 2. Unterrichtsmethodenforschung/Visualisierung

Spätestens seit der Reformpädagogik war es eine durchgängige methodische Empfehlung, dass Grundschulunterricht auf die Anschauung der Wirklichkeit gerichtet sein müsse und die Selbsttätigkeit zu praktizieren habe. Üblicherweise basierten diese Empfehlungen auf den Erfahrungen von Lehrern und wurden ohne empirische Absicherung aus der Forschung in „Unterrichtslehren“ publiziert. Sogar in neueren Lehrbüchern findet man extrem vereinfachende Aussagen ohne Forschungsbasis, etwa wenn gesagt wird, Lesen habe 10 %, Sehen 30 % und Tun 90 % Lerneffektivität. Michael (1983) schätzt in einer ebenfalls vereinfachenden Darstellung ohne genügend Forschungsausweis den „Intensitätswert“ für das lernende Hören auf 1, für das Betrachten eines Bildes auf 4, für die Beobachtung eines realen Gegenstandes auf 12 und für den handelnden Umgang mit dem realen Gegenstand auf 16 (vgl. zur Kritik Weidenmann 1995).

Eine der wenigen empirischen Untersuchungen, auf die sich solche Aussagen zur Anschauung und zum handelnden Umgang stützen, stammt von Düker (1968). Düker hatte vier Lerngruppen zum Thema „Meerschweinchen“ gebildet. Die erste Gruppe lernte mit einem Tonbandtext und einem Bild des Tieres. Der zweiten Gruppe stand zum Tonbandtext ein „Modell“ (ausgestopftes Tier) zur Verfügung. Die dritte Gruppe hatte den Tonbandtext und in einem Käfig ein lebendiges Meerschweinchen. Eine Kontrollgruppe lernte nur mit Tonbandtext und ohne Anschauungshilfen. Ergebnis: Die Gruppe mit dem „realen“, lebendigen Meerschweinchen hatte im Nachtest die meisten Punkte, die Gruppe mit dem „Modell“ lag an zweiter Stelle, die Gruppe mit dem Bild an dritter Stelle, die Kontrollgruppe hatte am wenigsten Punkte.

Diese und ähnliche Untersuchungen führten zur Realismusthese, die besagt, dass umso besser gelernt und behalten werde, je realistischer ein Lerngegenstand sei bzw. je mehr ein Medium den optischen und akustischen Kanal anspreche. Aufgrund einer Vielzahl von Studien, in denen sich eher strukturierende Merkmale in Bildern als vorteilhaft erwiesen, gilt heute die Realismusthese als überholt (vgl. Weidenmann 1988, S. 131-133). Wesentlicher als empirische Belege sind jedoch theoretische Überlegungen zur kognitiven Repräsentation der zu lernenden Sachstrukturen. Solche Überlegungen konnten verständlicherweise vor dem Aufkommen der modernen Kognitionspsychologie in den 70er Jahren gar nicht angestellt werden. So ist zu fragen: Welche kognitiven Prozesse sollen im Sachunterricht der Grundschule gefördert werden? Muss es nicht darum gehen, symbolische Repräsentationen im Langzeitgedächtnis aufzubauen, die abstrakter und damit leichter transferierbar sind als bildhafte Vorstellungen? Welche Bedeutung hat dabei der reale Gegenstand oder ein Bild? Wirken sie unterstützend im Sinne der Doppelcodierung nach Paivio (1986)? Oder hatte das lebendige Meerschweinchen nur



motivierende Wirkung? Brauchen Kinder im 1./2. Schuljahr mehr Bilder und reale Gegenstände, während 3./4.-Klässler schon mehr „Anschauungen“ haben und zu abstrakteren Struktureinsichten geführt werden können? Weidenmann (1996) kritisiert sogar den Rückgriff auf die Typisierung nach optischem, akustischem und haptischem Sinneskanal als zu vereinfachend. Tatsächlich müssten die unterschiedlichen Codes beachtet werden, mit denen repräsentiert wird. So verbergen sich z. B. unter dem visuellen „Kanal“ ganz unterschiedliche Codes, etwa ein Text, eine Tabelle, ein Bild oder ein mit Begriffen und Erklärungen angereichertes Bild. Die Forschung zur medialen Unterstützung des Wissensaufbaues müsse nicht auf „Kanäle“, sondern auf Verarbeitungscodes gerichtet sein.

In einer Sachunterrichtsstudie am Nürnberger Institut für Grundschulforschung haben wir versucht, stärker theorieorientiert zu arbeiten als die Anschauungsforschung der 60er Jahre (Treinies & Einsiedler 1993). Wir bezogen uns auf Repräsentationstheorien, in denen man annimmt, dass unser Wissen relativ abstrakt in hierarchischen Strukturen und in Bedeutungsnetzwerken abgebildet ist. Außerdem war für uns die Theorie der mentalen Modelle maßgebend, der zufolge wir im Arbeitsgedächtnis die Informationsverarbeitung mit bildlichen Anteilen unterstützen können. Wir verglichen im 4. Schuljahr experimentell drei Methoden: Methode (1) war eine hierarchische Darstellung von Begriffen und Relationen eines biologischen Themas (s. Abb.1).

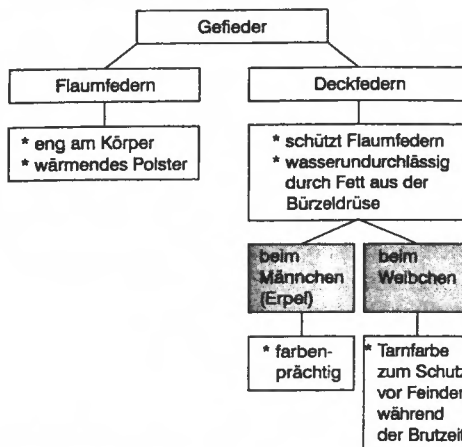


Abb. 1: Hierarchische Begriffsstruktur zum Thema „Die Anpassung der Stockente an ihre Umwelt“

Methode (2) umfasste netzwerkartige Darstellungen kausaler Zusammenhänge, die nach dem Konzept der mentalen Modelle mit bildlichen Anteilen angereichert waren (Abb.2).

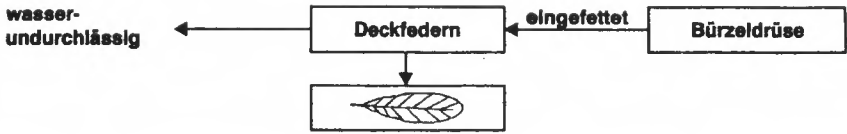


Abb. 2: Bedeutungsnetzartige Darstellung zum Thema „Die Anpassung der Stockente an ihre Umwelt“

Methode (3) war eine Kontrollmethode im Sinne traditioneller Lückentexte. In den Nachtests war hypothesenkonform ein (allerdings nicht signifikanter) Trend zugunsten der Methoden (1) und (2) sowohl beim Begriffslernen als auch bei höheren Denkopoperationen zu erkennen. Trotz fehlender signifikanter Überlegenheit der „neuen“ Methoden ist das Ergebnis ermutigend, weil es zeigt, dass auch schon Grundschulkinder auf relativ abstrakte Weise Ursache-Wirkungs-Wissen verarbeiten können; es ergeben sich differenziertere Möglichkeiten methodischer Zugänge und mentaler Verarbeitung.

In einer Arbeit von Martschinke (1996) wurde ebenfalls auf moderne Repräsentationstheorien Bezug genommen. Für sie sind mentale Modelle der Schlüsselbegriff; durch die Verwendung von Bildern und logischen Strukturen schlagen die mentalen Modelle sozusagen die Brücke zum Langzeitgedächtnis. Externe Modelle, die mentale Modelle unterstützen, müssen elaboriert sein, d.h. bezüglich des Sachunterrichts relativ viele und detaillierte bildliche Informationen enthalten; das

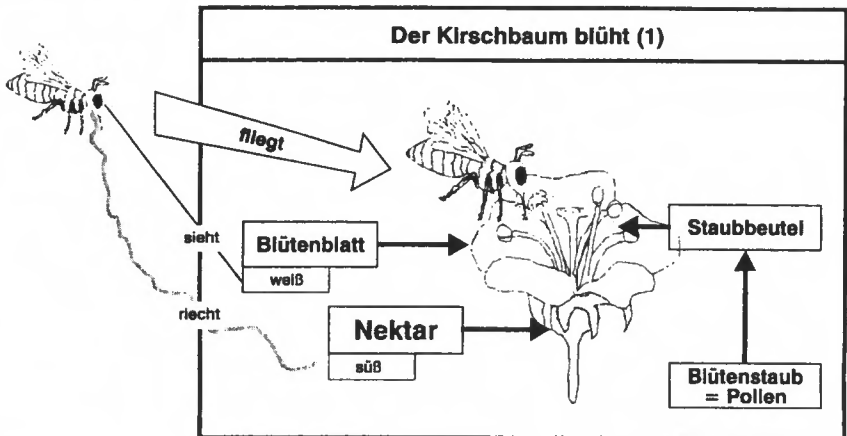


Abb. 3: Hochelaboriert-hochstrukturierte Darstellung (Ausschnitt aus dem Thema „Von der Blüte zur Frucht“). Bei niedrigelaborierter Darstellung würden bildliche Details fehlen, bei niedrigstrukturierter Darstellung Begriffe und Pfeile.

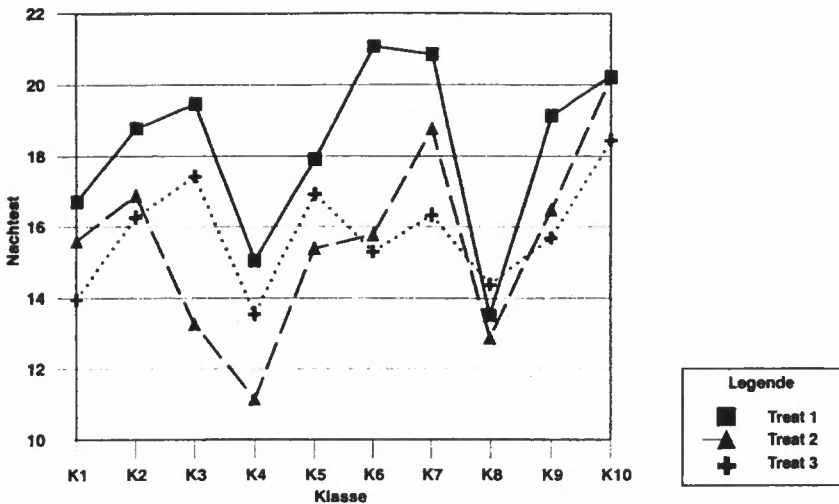


Abb. 4: Klassenmittelwerte in der Studie von Martschinke (1996). Treat 1: hochlaboriert-hochstrukturiert; Treat 2: hochstrukturiert-niedrigelaboriert; Treat 3: niedrigstrukturiert-hochelaboriert.

erleichtert den Kindern die Anknüpfung an Vorwissen und die Konstruktion individueller Modelle. Externe Modelle müssen aber auch hochstrukturiert sein, damit Ordnung im Wissen entsteht, die entscheidenden Begriffe hervorgehoben und die logischen Strukturen erkannt werden. Die Autorin verglich im 3. Schuljahr drei Lernmethoden: (1) hochelaboriert-hochstrukturiert: viele bildliche Details, viele Oberbegriffe und Relationsdarstellungen (Abb. 3);

(2) niedrigelaboriert-hochstrukturiert: keine bildlichen Details, nur strukturierende Abstraktionshilfen; (3) hochelaboriert-niedrigstrukturiert: viele Bilder, fast keine Strukturierungshilfen.

Methode (1) war den beiden anderen Methoden signifikant überlegen (Abb. 4). Die Autorin relativiert entsprechend diesem Ergebnis die Realismusthese in der Sachunterrichtsdidaktik und fordert mehr Strukturierungshilfen auch schon in der Grundschule. Realistische Abbilder müssten stärker zurücktreten, stattdessen sollten viel mehr „logische Bilder“ eingesetzt werden, mit denen die Struktureinsichten herausgearbeitet werden, die letztlich im Langzeitgedächtnis gespeichert werden.

Einsiedler & Martschinke (1998) führten in Anlehnung an die eben referierte Untersuchung eine Inhaltsanalyse von 77 Sachunterrichtsbüchern der 1. bis 4. Jahrgangsstufe durch. Dabei wurden insgesamt 8912 Illustrationen nach den Typen (Beschreibungen s. oben)

- (1) hochelaboriert-hochstrukturiert
  - (2) hochstrukturiert-niedrigelaboriert
  - (3) hochelaboriert-niedrigstrukturiert
- klassifiziert.

Ergebnisse: Von den 8912 Illustrationen waren 14,91 % dem Typ 1 zuzuordnen, 4,99 % dem Typ 2 und 80,09 % dem Typ 3. Es zeigt sich also ein enormes Übergewicht bei hochelaborierten-niedrigstrukturierten Illustrationen. Autoren von Sachunterrichtsbüchern präferieren in hohem Maße realistische und detailreiche Fotos und Zeichnungen ohne Strukturierungshilfen. Der für die kognitive Verarbeitung vorteilhafte Bildtyp hochelaboriert-hochstrukturiert (s. Abb. 3) wird nur wenig berücksichtigt. Nur jedes 20. Bild ist ein rein logisches Bild (z. B. Säulendiagramme, Pfeilbilder, etwa zur Energieerzeugung, Schaltbilder). Die Sachbuchautoren hängen überwiegend der Realismusthese an, sie wollen Grundschulkindern Abbilder vermitteln. Viele Bilder haben nur Dekorationsfunktion oder Unterhaltungswert (z. B. Fotos von Tätigkeiten in der Schule, Bilder von Advent und Weihnachten).

Die Annahme, dass der wichtige Illustrationstyp 1 von der 1. bis zur 4. Jahrgangsstufe zunehmend Verwendung findet, lässt sich kaum bestätigen. Der Wert dieses Typs steigt von der 1./2. Klasse (12,63 % bzw. 12,51 %) zur 3. Klasse an (17,04 %) und geht dann in der 4. Klasse auf 15,96 % zurück. D. h.: Die Schulbuchautoren berücksichtigen fast überhaupt nicht, dass viele Schüler der 3. und 4. Jahrgangsstufe logische Strukturdarstellungen verstehen und in Verbindung mit Bildanteilen Relationswissen besser verarbeiten können. Oder anders ausgedrückt: Die meisten Autoren sehen es nicht als Lehrziel oder Entwicklungsaufgabe der 3./4. Klasse an, allmählich zum Abstrahieren und zum Verstehen von Strukturdarstellungen anzuleiten.

Ganz auf der Linie der Förderung abstrahierend-symbolisierender Fähigkeiten von Grundschulkindern bewegt sich auch das Projekt von Stern, Koerber und Rode (1999) am Berliner Max-Planck-Institut für Bildungsforschung zu visuell-räumlichen Repräsentationen (s. auch den Beitrag von Stern, Hardy & Koerber in diesem Band). Es wird hier kurz referiert, weil es inhaltlich zu der bisher dargestellten Sachunterrichtsforschung gehört. Stern begründet das Vorgehen der Gruppe ebenfalls repräsentationstheoretisch und mit der Bedeutung von Symbolisierungen. Zu spezifisches deklaratives Wissen ist häufig auf zu spezielle Kontexte bezogen und deshalb oft nicht transferierbar. Symbolische Repräsentationen und vor allem visuell-räumliche Darstellungen (Diagramme, Graphiken) sind abstrakter und somit auf mehr Situationen transferierbar. „Je mehr sie von konkreten Handlungen losgelöst sind, d.h. je abstrakter sie sind, desto breiter können sie angewandt und auf andere Domänen übertragen werden“ (Stern et al. 1999, S. 7). Es wird auf Wygotski Bezug genommen, der es als Aufgabe der Schule ansieht, wissenschaftli-

che Begriffe anstelle von Alltagsbegriffen zu vermitteln und das symbolische Verständnis zu fördern. Symbole sind die Objekte des Denkens und sie erlauben die Konstruktion der kognitiven Strukturen.

Die Berliner Gruppe konzentriert sich jedoch mehr auf symbolisierende Verfahren (prozedurales Wissen) als auf das Abstrahieren inhaltlicher Struktureinsichten und nennt diese Verfahren „Tools“. Tools können Tabellen, Diagramme, graphische Darstellungen mit Pfeilen u. Ä. sein. Bisher wurde insbesondere der Umgang von Kindern mit Kartesischen Koordinatensystemen untersucht, das sind Diagramme mit x-Achse, y-Achse und Funktionsgeraden. In einer Pilotstudie wurde gefunden, dass schon 6-Jährige einfache Funktionen in solchen Diagrammen verstehen. In einer Studie mit konkreten Wasserumschüttaufgaben konnten 4.-Klässler gut mit Kartesischen Koordinatensystemen umgehen, wenn sie vorher beim Umschütten mit Tabellen gearbeitet hatten. Die Autoren sind der Auffassung, dass die Förderung visuell-räumlicher Repräsentationen bisher vernachlässigt wurde und die Ausbildung entsprechender Kompetenzen sehr hilfreich für das Problemlösen in den verschiedensten Wissensgebieten sei.

Nun ist nicht zu leugnen, dass die Empfehlungen aus der Forschung zu symbolisierenden Repräsentationen einem aktuellen Trend in der Sachunterrichtsdidaktik entgegenstehen, der gerade nicht abstrahiertem Strukturwissen das Wort redet, sondern die konkrete Anschauung, etwa von Naturphänomenen, favorisiert, dem sinnhaften Lernen großen Wert beimisst und insgesamt die Sachunterrichtsthemen in lebensweltliches Lernen eingebettet sehen möchte. Es ist nicht zu bestreiten, dass gerade im Grundschulalter keine abstrakten Struktureinsichten, etwa bei den Themen Wald oder Anpassung eines Tieres, aufgebaut werden können, wenn die Kinder nicht vorher konkrete Vorstellungen von den Sachverhalten erworben haben. Umgekehrt ist ein bloßes Betrachten des Naturphänomens unergiebig, etwa beim Thema Wald müssen abstrahierende Darstellungen hinzukommen, um das Verständnis des Natursystems anzubahnen. Die Einbettung in die Lebenswelt wirkt u.U. motivierend, aber nicht von selbst objektivierend. Die Widerlegung der Realismusthese macht es fast zwingend notwendig, die Anschauung von Konkretem mit Bildern, die die logischen Strukturen hervorheben, zu ergänzen. Auch weitere empirische Evidenz (z. B. im Bereich des Mathematikunterrichts) spricht – zumindest gegen Ende von Unterrichtssequenzen – für die Arbeit mit strukturierenden und symbolisierenden Darstellungen.

### 3. Schülervorstellungen/Conceptual Change

Die Forschung zum Wissenserwerb im Sachunterricht im Grundschulalter stagnierte lange Zeit, weil das Bild der konkret-operatorischen Strukturen nach Piaget dieses Stadium prägte. Man ging von einem relativ langsamen Lernfortschritt in

dieser Zeit aus, der erst wieder durch die Umbrüche im formal-operatorischen Stadium beschleunigt würde. Amerikanische Forscher wichen dagegen vom Stadienmodell ab und untersuchten, wie sich Schülervorstellungen von Jahr zu Jahr verändern. Dieses Paradigma des „Conceptual Change“ wurde Anfang der 80er Jahre von einer Gruppe an der Cornell University von Autoren wie Posner, Strike und Hewson mit Blick auf den naturwissenschaftlichen Unterricht der Sekundarstufe eingeführt (Posner, Strike, Hewson & Gertzog 1982). Ausgangspunkt war die Überlegung, dass Wissenserwerb nicht so sehr der Neuaufbau von einem Nullpunkt aus sei, sondern die Transformation vorhandener Wissensstrukturen, vor allem die häufig erforderliche Veränderung von Alltagsvorstellungen und falschen Vorstellungen (Misconceptions).

„Conceptual Change“ ist selbstverständlich auch auf das Grundschulalter anwendbar, hier kennt man seit längerem Misconceptions wie z. B. das Konzept der Sonnenbewegung um die Erde, die Annahme, Wolken seien Behälter oder die Vorstellung, im Stromkreis bräuchte man zwei Kabel zur ausreichenden Stromzuführung.

Für die Forschung im Grundschulalter wurde der Conceptual-Change-Ansatz von Susan Carey (1985) sehr bedeutsam. Der zentrale Begriff bei Carey ist der der Umstrukturierung; Novizen werden zu Experten, indem sie durch Umstrukturierung mehr Relationen und mehr abstrakte Schemata in ihren kognitiven Strukturen schaffen. Die Umstrukturierung kann „weich“ sein, z. B. bei Wissensausdifferenzierung, oder sie kann „radikal“ sein, z. B. bei der Aufgabe falscher Konzepte. Carey spricht von Theorien der Kinder, die sich allmählich oder rasch wandeln. Bei der Erforschung biologischer Begriffe im Grundschulalter (z. B. Ernährung, Verdauung, Atmung) kam Carey zu dem Ergebnis, dass für die Begriffsentwicklung die Annahme allgemeiner Stufenmodelle unzutreffend ist, dass vielmehr von einer sehr bereichsspezifischen kontinuierlichen Wissensveränderung auszugehen ist. 10-Jährige haben biologisch korrekt Begriffe integriert, z. B. Tiere und Pflanzen als lebendig, und sie haben korrekte Oberbegriffe gebildet, z. B. Lebenserhaltung für Essen und Verdauung. Carey geht auf starke Distanz zu Piaget und ist der Auffassung, dass Stadienmodelle zu generell angelegt seien, um die Veränderungen in Hunderten von verschiedenen Begriffsbereichen zu beschreiben. In einzelnen Begriffsfeldern könnten auch schon sehr junge Kinder formal-logisch schlussfolgern und umgekehrt seien Erwachsene in bestimmten Wissensdomänen dazu nicht in der Lage. Für Grundschulsachunterricht ist also entscheidend, den Prozess der Differenzierung und Integration des Wissens im Verlauf des konkret-operatorischen Stadiums zu studieren und nicht auf qualitative Sprünge im 11./12. Lebensjahr im formal-operatorischen Stadium zu warten.

Das ältere Conceptual Change-Modell von Posner et al. nahm rationale Überlegungen der Schüler an, die zum Konzeptwechsel führen. Dieses Modell wird inzwi-

schen als „rational unterkühlt“ angesehen. Pintrich, Marx und Boyle (1993) haben ihm ein „heißes“ Modell gegenübergestellt, das motivationale Faktoren und situationale Klassenzimmerbedingungen einbezieht. Danach ist Conceptual Change auch abhängig von individuellen Zielorientierungen, Überzeugungen und Interessen sowie von sozialen Strukturen und Dynamiken in der Klasse. Pintrich et al. empfehlen die Projektmethode, um für Conceptual Change intrinsische Motivations- und Selbstwirksamkeitsmotive ins Spiel zu bringen.

Einige Autoren bringen sowohl die Erfassung der Schülervorstellungen als auch ihre unterrichtliche Veränderung mit der Theorie der mentalen Modelle in Verbindung. So konnten Schnotz, Zink und Pfeiffer (1996) nachweisen, dass unterschiedliche externe Modelle im Unterricht zu unterschiedlichen internen Modellen und Wissensstrukturen führen. Gut gestaltete externe Modelle und logische Bilder sind also geeignete Methoden für Conceptual Change. Wenn man interne mentale Modelle erforscht, erhält man Anknüpfungspunkte für die Gestaltung unterrichtlicher Medien bzw. externer Modelle.

Vosniadou und Brewer (1992, 1994) haben solche internen Modelle am Beispiel der Form der Erde einschließlich der Schwerkraft und der Entstehung von Tag und Nacht untersucht. Die Autoren haben Grundschul Kinder mit raffinierten Methoden des Befragens und Zeichnenlassens ihre Vorstellungen von der Erde und der Schwerkraft darstellen lassen. Auf den ersten Blick zeigten die Ergebnisse inkonsistente mentale Modelle der Erde, etwa wenn die gleichen Kinder von der Erde als Kugel und von der Kante am Ende der Kugel sprachen. Weitere Auswertungen förderten dann aber zutage, dass die Kinder sehr wohl konsistente Modelle von der Erde haben, z. B. wenn das Problem des äußeren Eindrucks der flachen Erde mit einer abgeflachten Kugel („Pfannkuchenmodell“) oder mit der Vorstellung Kugel plus Fläche gelöst wird (Abb.5).

Am eindrucksvollsten ist wohl das Modell der Erde als hohle Kugel, das es erlaubt, vielseitige Inkonsistenzen, z. B. Nichtherunterfallen, Herunterfallen, Kugel ohne Ende und Erde mit Rand aufzuheben (Abb.6). Vosniadou und Brewer sprechen von überlegten kohärenten Synthesemodellen, die die Kinder entwickeln. Sie unterstützen die Annahme von Carey, dass Conceptual Change sehr bereichsspezifisch und nicht nach allgemeinen Stadienmodellen zu untersuchen ist und dass es sich um den allmählichen oder raschen Wandel theorieähnlicher Strukturen handelt (einige Autoren nehmen z. B. eine Entwicklungsreihe der Schülervorstellungen vom untersten falschen bis zum obersten richtigen Konzept in Abb. 5 an).

In Deutschland wurde relativ viel zu Schülervorstellungen im naturwissenschaftlichen Bereich gearbeitet, wobei es mehr Untersuchungen zur Sekundarstufe als zur Grundschule gibt. Exemplarisch seien zum physikalischen Bereich als Grundschulstudien aufgeführt:

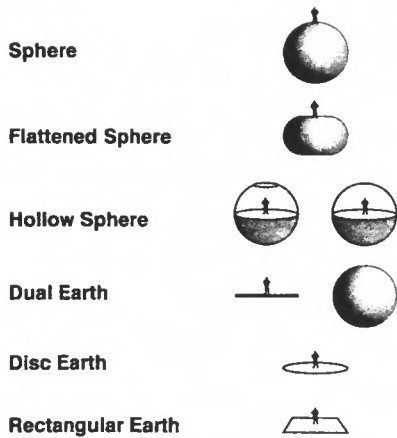


Abb. 5: Mentale Modelle der Erde von Grundschulkindern (Vosniadou & Brewer 1992)

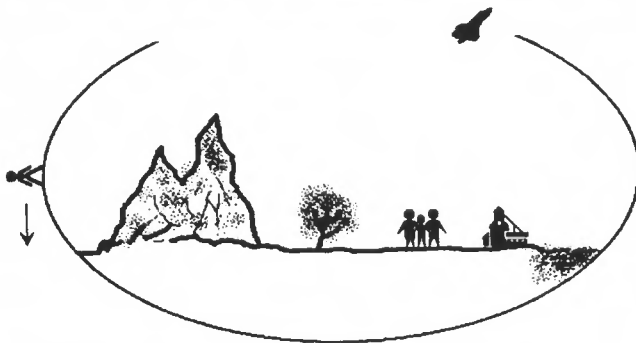


Abb. 6: Die Erde als hohle Kugel, „Synthese-Modell“ nach Vosniadou & Brewer 1992

- zum Schall (Kircher & Engel 1994, Wulf & Euler 1995),
  - zum Magnetismus (Kircher & Rohrer 1993),
  - zur Optik (Claus, Stork & Wiesner 1982),
  - zu Temperatur und Wärme (Wiesner & Stengl 1984),
  - zu Schatten und Licht (Wiesner & Claus 1985, Faust-Siehl 1993),
  - zu Elektrizität (Stork & Wiesner 1981),
  - zum Gleichgewicht am Hebel (Hagstedt & Spreckelsen 1986).
- (als Gesamtbibliographie s. Pfundt & Duit 1994, zu Grundschulfragen des Conceptual Change vgl. Duit 1997).
- Im soziokulturellen Bereich liegen wesentlich weniger Studien zu Schülervorstellungen vor. Als Beispiele seien hier genannt:



- soziale und moralische Vorstellungen (Faust-Siehl & Schweitzer 1991),
- Klassenregeln, Normen, Moralvorstellungen (Beck, Scholz & Walter 1991),
- Krieg und andere politische Themen (Richter 2000).

Die Erfassung der Schülervorstellungen erfolgt im Grundschulalter meist im Einzelverfahren mit einem Interviewleitfaden, bei naturwissenschaftlichen Themen werden häufig auch kleine Versuche eingeschoben. Beim Thema „Schall“ haben Kircher und Engel (1994) z. B. folgende Fragen gestellt: Hast du das Wort Schall schon einmal gehört? Weißt du, was Schall ist? Wie kann man Töne erzeugen? Beobachte ganz genau, was passiert, wenn man einen Ton erzeugt, bei der Stricknadel, beim Monochord... Kircher und Engel gehen hypothesengeleitet vor. In ihrer Schall-Studie lautet Hypothese 1: Kindern der Primarstufe ist in der Regel nicht bewusst, dass Töne oder Geräusche eine sich bewegende Schallquelle als Ursache haben. Diese Hypothese wurde bestätigt, die befragten Kinder gaben nur an, es müsse eine Tätigkeit ausgeübt werden, um Schall zu erzeugen, die schwingende Bewegung wurde nicht genannt. Danach konnten die Kinder die schwingende Stricknadel und die schwingende Saite beobachten. Hypothese 2 lautet deshalb: Kinder der Primarstufe können die sichtbare oder fühlbare Bewegung einer Schallquelle als Ursache der Tonentstehung erfassen. Jetzt waren die meisten Kinder in der Lage, den Zusammenhang zwischen der Bewegung und der Tonentstehung zu begreifen. Beim anschließenden Anschlagen eines Metallophons sollte geprüft werden, ob die Kinder das Konzept verstanden haben, auch wenn man die Schwingung nicht sieht. Hier war ein Entwicklungstrend feststellbar: Vor allem Kinder der 2. Klasse orientierten sich daran, was sie sehen, alle Zweitklässler waren der Meinung, dass sich beim Metallophon nichts bewege. Die meisten Kinder der 3. und 4. Klasse übertrugen dagegen das Bewegungskonzept auf das Metallophon und nannten Bewegung als Ursache für Schall. Weitere Hypothesen bezogen sich u.a. auf Luft als schwingenden „Körper“ und auf die Notwendigkeit eines Stoffes als leitfähiges Medium.

Der Physikdidaktiker Wiesner erfragte nicht nur Schülervorstellungen, sondern erforschte auch unterrichtliche Strategien, mit denen falsche Vorstellungen in Frage gestellt und richtige Konzepte eingeführt werden können. Mit einer Studie von Wiesner (1995) zum Stromkreis soll dieses Verfahren demonstriert werden. Wiesner hatte in einer früheren Untersuchung festgestellt, dass 75 % der befragten Grundschüler meinten, bei einem Stromkreis mit Batterie und Lämpchen reiche ein Verbindungskabel. Nach Versuchen mit zwei Kabeln waren 85 % der Auffassung, man brauche die zwei Kabel, um genügend Strom zuleiten zu können (Zweizuführungsvorstellung). Fast alle Schüler waren der Meinung, der Strom werde verbraucht, hinter dem Lämpchen sei weniger Elektrizität im Kabel als vor dem Lämpchen.

In der Studie zum Konzeptwechsel wurde eine „Akzeptanzbefragung“ durchge-

führt. Die Versuchsleiter stellten die Misconceptions in Frage und boten richtige physikalische Erklärungen an. Die Schüler sollten die Erklärungen mit eigenen Worten paraphrasieren und von einem Stromkreis mit Lämpchen auf einen Stromkreis mit Elektromotor transferieren. In einem ersten Schritt wurden richtige und falsche Stromkreisvorstellungen demonstriert, das Einzuführungskonzept widerlegt und die Kreisvorstellung mit entsprechenden Bildern besprochen. In einem zweiten Schritt wurde mit einer Batterie mit höherer Spannung das Lämpchen heller gemacht und die Schüler sollten selbstständig schlussfolgern, ob beim Motorversuch die Drehzahl des Motors bei höherer Spannung höher ist. Gegen die Vorstellung vom Stromverbrauch wurde eine Vorrichtung mit zwei Kompassnadeln eingesetzt, also die magnetische Wirkung des elektrischen Stroms genutzt (Abb. 7).

Die Kinder konnten so beobachten, dass die Nadel hinter dem Lämpchen genauso stark und in die gleiche Richtung ausschlug wie die Nadel vor dem Lämpchen. Das Gleiche wurde mit dem Austausch der Anschlüsse an der Batterie und mit einer stärkeren Voltzahl demonstriert.

Die Befragungen erbrachten, dass diese Interventionsstrategie erfolgreich war: 78 % der Schüler stimmten der Stromkreisvorstellung mit Zufluss und Rückfluss des Stroms zu, 15 % blieben resistent und hatten weiterhin eine Zweizuführungsvorstellung. 75 % der Schüler akzeptierten nach dem Magnetnadelversuch die Vorstellung von der Konstanz der Elektrizität im Stromkreis, 19 % äußerten weiterhin eine Verbrauchsvorstellung. Bei einer Reihenschaltung als Abschlusstest wandten 62 % der Grundschul Kinder recht sicher die Stromkreis- und die Stromerhaltungsvorstellung an.

Die Arbeit von Wiesner zeigt, dass es mit allgemeinen Empfehlungen wie „In-Fragestellen“, „Konfliktstrategie“ oder Ähnlichem für Konzeptwechsel nicht getan ist. Vielmehr dürfte es sich als günstig erweisen, wenn Fachdidaktiker sehr themenspezifisch Unterrichtsvorschläge erarbeiten, die auf den Abbau von falschen und den Aufbau von richtigen Konzepten zielen. Ganz eindrucksvoll wird das mit dem Magnetnadelversuch zur Stromverbrauchsvorstellung deutlich.

Im Kontext der Wiesner-Studie drängt sich die Frage auf, ob die „Akzeptanzbefragung“ aus der Forschung auf den Unterricht übertragen werden sollte. Diese Methode ist ja gekennzeichnet durch eine Art Belehrung, Vorgabe von Erklärungen, Demonstration von Sachverhalten. Gegen die Übernahme dieser engführenden Strategie spricht u.a., dass die Schüler dadurch unselbstständig gehalten würden, keine Arbeitstechniken des jeweiligen Lernbereichs lernen würden und u. U. die Lernmotivation reduziert würde. Man wird also Wege finden müssen, forschendes und entdeckendes Lernen mit seinen pädagogischen Vorteilen für die jeweiligen Unterrichtseinheiten zu konzipieren und sorgfältig vorbereitete Versuchsmöglichkeiten anzubieten. Wiesner selbst schlägt ein gemischtes Vorgehen vor:

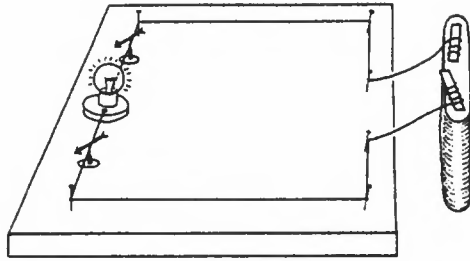


Abb. 7: Stromkreis mit zwei Magnetnadeln zur Verdeutlichung des „Kreis“-Konzepts und des Nicht-Verbrauchs

„Selbstverständlich müssen Elemente offenen Umgangs mit den Materialien, spielerische Zugänge usw. ausgewogen mit Phasen etwas engerer Führung bei der Diskussion der schwierigen physikalischen Ideen kombiniert werden“ (1995, S. 58).

### 3. Lehrstoff- und Lehrplananalysen

Schon eine gewisse Tradition in der Sachunterrichtsforschung haben Dokumentenanalysen, hier verstanden als quantitative Aufschlüsselungen von Inhaltsanteilen in Klassenbüchern/Lehrnachweisen, in Schülerarbeitsmappen, in Schulbüchern und in Lehrplänen. Sie erlauben Aussagen darüber, welche Lehrstoffe in der Sachunterrichtsrealität tatsächlich unterrichtet werden, wie sich die Gewichte über verschiedene Zeitabschnitte verschieben, und bei Lehrplananalysen, welchen Stellenwert man Fachanteilen und bestimmten Themen zumisst.

Den Anfang machte Höcker (1968) mit einer Inhaltsanalyse von Klassenbüchern der 4. Jahrgangsstufe in Kiel, damals noch stärker auf die alte Heimatkunde bezogen. Sie erbrachte, dass von 1.436 Heimatkundestunden 70,9 % der Erdkunde und allen anderen Fachanteilen jeweils nur unter 10 % der Stunden gewidmet waren. Es zeigte sich das nicht unerwartete Bild von der Heimatkunde als einer Art Propädeutik der Geographie (in Städten mit historisch bedeutsamen Themen überwog allerdings eher die Heimatgeschichte). Die Studie von Höcker trug zu der Intention bei, im wissenschaftsorientierten Sachunterricht ab 1970 Fachanteile wie Biologie und Physik/Technik stärker zu gewichten.

Schreier legte 1979 eine Auswertung von 279 Lehrberichten Kasseler Schulen mit insgesamt 22.563 Eintragungen vor. Schreier analysierte den Verlauf von 1968 bis 1974 und stellte fest, dass in diesem Zeitraum im 4. Schuljahr erdkundliche Themen von 61,6 % auf 28,8 % zurückgingen. Auch bei historischen Themen war ein Rückgang zu verzeichnen. Demgegenüber gab es Zunahmen bei sozialkundlichen Themen, bei Biologie von ca. 5 %/6 % (je nach Klassenstufe) auf ca. 19 % The-

menanteile und bei Physik/Technik von ca. 2 %/5 % auf ca. 20 % Themenanteile. Die Studie belegt gut den Übergang von der Heimatkunde zum Sachunterricht. In einer eigenen Untersuchung im Raum Nürnberg/Erlangen werteten wir Schülerarbeitsmappen im Längsschnitt von 1968 bis 1981 aus (Einsiedler & Schirmer 1986). Erdkunde hatte in unserer Arbeit 1968 nur einen Anteil von 26 % und ging bis 1981 auf ca. 20 % zurück. Der wissenschaftsorientierte Lehrplan Bayerns von 1971 bewirkte, dass der Anteil der Biologie von 15 % auf 25 % anstieg. Der Anteil der Geschichte sank zunächst leicht von 13 % auf 11 % und nahm dann nach 1976 auf 18 % zu. Erhellend ist der Verlauf in Physik/Chemie/Technik: Der Anteil lag im ersten Zeitraum bei 11 %, stieg dann durch den wissenschaftsorientierten Lehrplan auf 26 % an und ging nach einer Lehrplanrevision 1976 wieder auf 11 % zurück. In der Zeit um 1976/1978 kam es zu intensiven, häufig kontroversen Diskussionen über Physik/Chemie/Technik. Manche Schulräte und Seminarleiter schilderten subjektive Eindrücke von einer starken Überrepräsentation des naturwissenschaftlichen Sachunterrichts. In einer Lehrerbefragung 1978 in Bayern antworteten 44,5 % der Befragten, Physik/Chemie/Technik sei im Lehrplan übergewichtet (Bittlinger, Flügge, v. Alt-Stutterheim, Müller & Andreas 1980). In einer neuen Arbeit von Blaseio (1999) wurden Schulbücher zum Sachunterricht auf fachliche Anteile hin untersucht (vgl. Blaseio in diesem Band). Die Studie ist sehr gründlich und detailreich angelegt: Sie umfasst sechs Zeitblöcke 1970-1974 bis in die Gegenwart 1995-1999; analysiert wurden 3000 Buchseiten, wobei jede Seite in 10 Zehntel als Analyseeinheiten aufgeteilt wurde (eine sog. Raumanalyse). Hier soll für die folgende Diskussion nur ein Befund referiert werden: Der Fachanteil Physik/Chemie/Technik betrug 1970-1974 29 %, bereits 1975-1979 ist ein Rückgang auf 16,6 % festzustellen und 1995-1999 ist nur noch ein Anteil von 6,5 % zu verzeichnen.

Dieses Ergebnis korrespondiert mit Befunden von Strunck, Lück und Demuth (1998), die in umfassenden Analysen Lehrpläne, Schulbücher, Fachzeitschriften und Klassenbücher ausgewertet haben. Der Rückgang von Physik/Chemie in Lehrplänen zeigt sich hier z. B. in folgenden Zahlen: Lehrplan Nordrhein-Westfalen 1973 7,35 %, Schleswig-Holstein 1978 8,70 %, Mecklenburg-Vorpommern 1996 1,25 %, Schleswig-Holstein 1997 3,25 %. „Seit Beginn der 80er Jahre sind [...] physikalische und chemische Themen in den Lehrplänen meist nur noch eine Randerscheinung, bei allerdings großen Unterschieden zwischen den einzelnen Bundesländern. In Berlin, Sachsen-Anhalt, Thüringen und dem Saarland fehlt ein Bereich ‚Phänomene der unbelebten Natur‘ praktisch vollständig“ (a. a. O., S. 73). Die Autoren dokumentieren den Rückgang von Physik/Chemie/Technik auch für Schulbücher, Fachzeitschriften und Klassenbücher.

Der naturwissenschaftliche Sachunterricht war in den 70er Jahren mit großen Hoffnungen eingeführt worden. Jedoch wird auch nicht selten von großen Schwie-

rigkeiten berichtet, die diese Innovation begleiteten. In den 80er Jahren gelang es offensichtlich nicht, einen Neuanatz mit größerer Akzeptanz zu konzipieren. Es handelt sich um ein vielschichtiges Problem, bei dem u.a. folgende Faktoren eine Rolle spielen:

- Es herrscht Unsicherheit hinsichtlich der Unterrichtsgestaltung. In den 70er Jahren wurde häufig eine Unterrichtsmethodik im Sinne einer vorweggenommenen Sekundarstufenphysik praktiziert. Es gibt Kontroversen darüber, ob in der Grundschule das experimentierende Lernen mit allen Stufen wie Problemstellung, Hypothesenbilden, Versuchsdurchführung, Schlussfolgern usw. durchgeführt werden kann oder ob eine Art „kleines Experimentieren“ mit einer Beschränkung auf die Phänomenbetrachtung und ohne umfassende Problemerkörterung angebracht ist.
- Einige entwicklungspsychologische Fragen sind ungeklärt. So bezweifeln einzelne Autoren, ob bestimmte naturwissenschaftliche Themen wegen ihrer logischen Struktur bereits in der Grundschule unterrichtet werden können. Entzündet hat sich diese Debatte am Thema „Schwimmen und Sinken“, bei dem die Kinder gleichzeitig die Variablen Volumen und Gewicht beachten müssen. Janke (1995) hat mit einer geeigneten Versuchsanordnung nachgewiesen, dass die Kinder diese Zweidimensionalität bewältigen. Wiesner (1991) und Wilkening (1994) haben grundsätzlich die Gültigkeit der einschränkenden entwicklungspsychologischen Aussagen Piagets zum konkret-operatorischen Stadium bestritten und auf die Notwendigkeit einer veränderten Unterrichts- und Forschungsmethodik aufmerksam gemacht.
- Aus nichtsystematischen, eher subjektiven Unterrichtsbeobachtungen wird berichtet, dass in den 80er und 90er Jahren Sachunterrichts-Lehrkräfte soziokulturelle Inhalte präferierten. Dies ist jedoch nicht gut belegbar, in Schleswig-Holstein und Niedersachsen z. B. scheint die Erdkunde zu dominieren (Strunck et al. 1998, S. 76, Breitschuh 1997, S. 3). Gut belegbar ist jedoch die Zurückhaltung gegenüber technischen Themen und die Zuschreibung mangelnder Kompetenz in diesem Bereich. Dies zeigen auch die Ergebnisse einer Befragung zum Technikanteil des Sachunterrichts von Möller, Tenberge und Ziemann (1996): Auf die Frage „Haben Sie in diesem oder im letzten Schuljahr ein Technikthema behandelt?“ antworteten nur 18,2 % der Lehrkräfte mit „Ja“ (S. 16). Wenn die Antworten geschlechtsspezifisch ausgewertet wurden, ergab sich: 30,45 % der männlichen Lehrer, aber nur 15,04 % der Lehrerinnen antworteten mit „Ja“. Die Frage „Fühlen Sie sich kompetent hinsichtlich Unterricht in Technik?“ beantworteten 56,30 % der männlichen Lehrer positiv (S. 31). Nur 14,92 % der Lehrerinnen antworteten mit „trifft zu“. Über 80 % der Lehrkräfte in Nordrhein-Westfalen, wo die Befragung stattfand, ist weiblich.

Eine Replikation der Höcker-Studie von 1968 führte Breitschuh (1997) in einer niedersächsischen Kreisstadt durch. Er untersuchte Klassenbücher von 4. Jahr-

gangsstufen, also wieder Aufzeichnungen über tatsächlich gehaltene Unterrichtsstunden, auf Fachanteile hin (s. Abb.8).

Erdkundliche Themen stehen mit 28,6 % an erster Stelle, gefolgt von biologischen Themen mit 25,2 %. An dritter Stelle rangiert Verkehrserziehung einschließlich einer etwas unglücklich definierten Restkategorie (in Niedersachsen gehört Verkehrserziehung zum Sachunterricht und im 4. Schuljahr wird die Radfahrprüfung abgelegt). Erstaunlich niedrig fallen die Anteile für Sozialkunde und Geschichte aus, in anderen Bundesländern haben diese Lernfelder allein vom Lehrplan her mehr Gewicht. Aber offensichtlich wurden in den erfassten Klassen Lehrplanvorgaben nicht stark beachtet. Dazu betrachte man die Streuungswerte in Abb. 8: In einer Klasse ist der Erdkundeanteil 13,9 %, in einer anderen 61,8 %. Biologie streut zwischen 4,7 % und 46,5 %. In drei Klassen wurde der Lehrplan schlicht ignoriert: In einer Klasse ist der Anteil Physik/ Technik 0 %, in einer anderen hat Geschichte den Wert 0 % und in einer weiteren Klasse umfasst Sozialkunde 0,8 %. Breitschuh geht von 140 Stunden Sachunterricht im 4. Schuljahr aus, davon fielen in einer Klasse 18,1% aus, eine Zahl, die man nach Ansicht des Autors in Deutsch und Mathematik nicht akzeptiert hätte. Insgesamt zeigt die Tabelle, dass Fachanteile des Sachunterrichts fast beliebig unterrichtet werden, ausschlaggebend sind anscheinend subjektive Interessen der jeweiligen Lehrkräfte. Die Streuungszahlen verweisen auf eine gewisse Konzeptionslosigkeit und auf ziemlich willkürliche Einstellungen gegenüber dem Sachunterricht. Der Bildungsauftrag des Sachunterrichts hinsichtlich verschiedener Lernfelder wurde hier nicht ernst genommen.

An unserem Nürnberger Institut haben wir 1998 eine Analyse aller Sachunterrichtslehrpläne der 16 Bundesländer der Jahre 1979-1997 auf Einzelthemen hin vorgenommen (Einsiedler 1998). Ziel war eine Bestandsaufnahme vor allem zum naturwissenschaftlichen Sachunterricht im Hinblick auf eine internationale Tagung, die sich mit den Folgen von TIMSS für die Grundschule befasste. Es handelt sich um die Lehrpläne der 2. Generation nach der wissenschaftsorientierten Phase der 70er Jahre. Ausnahme ist der Teillehrplan Physik und Technik 1974 in Hamburg, der wegen der vielen dort enthaltenen Technikthemen zu einer Verzerrung im Lernfeld Technik führte und der seit längerem nicht mehr gültig ist. Die Lehrpläne der 2. Generation sind nicht mehr als Fachlehrpläne konzipiert, sie sind überwiegend fachübergreifend und lebensweltbezogen angelegt. Trotzdem haben wir wegen des Interesses an Science Education nach Fachanteilen zusammengefasst. Das Thema „Wetter“ war dabei schwer zu kategorisieren. Es enthält oft nur den Aspekt Wetterbeobachtung; häufig ist nicht klar, ob zum Wasserkreislauf Verdunstung und Kondensation tatsächlich physikalisch-experimentell behandelt werden sollen. Deshalb bildeten wir eine eigene Kategorie „Wetter“.

	Anteil	Streuung
Erdkunde	28,6 %	13,9 % - 61,8 %
Biologie	25,2 %	4,7 % - 46,5 %
Verkehrsunt./ Sonstiges	17,8 %	5,7 % - 25,4 %
Physik/Technik	9,6 %	0 % - 17,5 %
Sozialkunde	6,4 %	0,8 % - 14,8 %
Geschichte	4,6 %	0 % - 22,3 %
ausgefallen	7,7 %	2,9 % - 18,1 %

Abb. 8: Fachanteile in Klassenbüchern 4. Jahrgangsklassen (Breitschuh 1997)

In der 1./2. Klasse sind Soziales Lernen und Biologie die „Spitzenreiter“ (Abb.9). Beim Sozialen Lernen wird 60 x das Thema „Gemeinschaft/ Verständnis für andere“ genannt, gefolgt von 25 x „Zusammenleben in der Schule“. In Biologie kommt 44 x das Thema „Pflanzen“ und 42 x das Thema „Tiere“ vor, am meisten genannt werden „Tiere im Winter“, „Wiesenblumen“ und „Haltung eines Tieres im Klassenzimmer“. Technik weist ohne Hamburg 47 Themen auf. In Physik werden am häufigsten aufgeführt: Luft 8 x, Wasser 6 x, Thermometer 6 x (im 3./4. Schuljahr Luft 8 x, Wasser 12 x, Thermometer 3 x). In der Hälfte und mehr der 16 Bundesländer werden die genannten Themen im 1./2. Schuljahr nicht behandelt, man traut den 6-/7-Jährigen bei der Anbahnung physikalischen Denkens wenig zu (vgl. dagegen die erfolgreichen Unterweisungen mit 5-/6-Jährigen durch Lück 1999). Das Verhältnis „Social Studies“ zu „Science“ beträgt im 1./2. Schuljahr 59 % zu 41 %.

Von den 234 Science-Themen stammen 59,4 % aus der Biologie. In der 3./4. Klasse ist Biologie der umfangreichste Lernbereich (Abb. 10). Beim Sozialen Lernen geht es wieder um „Gemeinschaft/Verständnis für andere“, verstärkt durch Medienerziehung. Erdkunde ist an die 3. Stelle gerückt. Betrachten wir die Themen der Physik, stellen wir wieder fest, dass „Wasser“ und „Luft“ am häufigsten genannt werden; Beispiele sind „Verdunsten“, „Verdampfen“, „Luft ist elastisch“, „Luft kann bremsen“. Die Häufigkeiten der dann folgenden Themen sind sehr niedrig: Stromkreis 7 x (0 x in Klasse 1/2), Schwimmen und Sinken 3 x (1 x in Klasse 1/2), Magnetismus 2 x (2 x in Klasse 1/2). Das Verhältnis „Social Studies“ zu „Science“ beträgt im 3./4. Schuljahr 55 % zu 45 %, was auf den ersten Blick akzeptabel erscheint. Zu bedenken ist jedoch, dass Biologie an den 355 Science-Themen einen Anteil von 60 % hat. Von den insgesamt 790 Themen machen die 97 Themen in Physik und Technik (Technik ohne Hamburg) 12,3 % aus. Die Auswertung bestätigt im Großen und Ganzen die Ergebnisse von Strunck et al.

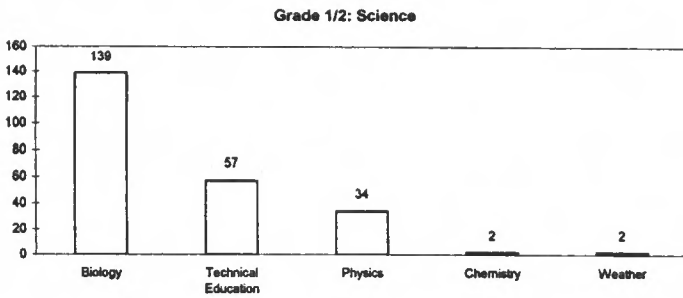
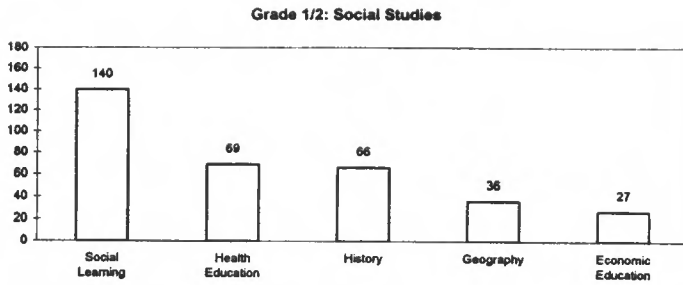


Abb. 9: Häufigkeit fachlicher Themen in 16 deutschen Lehrplänen (1./2. Klasse)

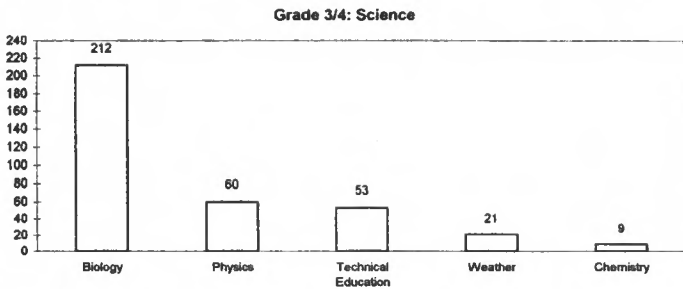
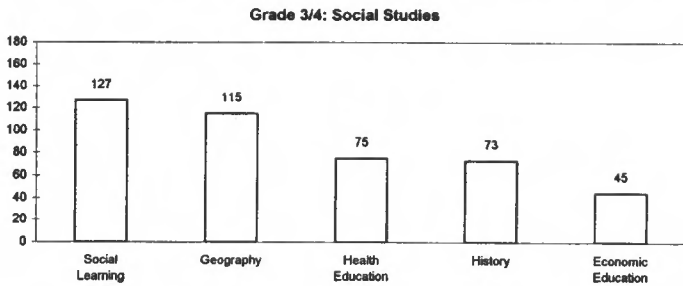


Abb. 10: Häufigkeit fachlicher Themen in 16 deutschen Lehrplänen (3./4. Klasse)



(1998). Die Themen aus Physik/Chemie/Technik sind in deutschen Sachunterrichtslehrplänen unterrepräsentiert. Besonders unsicher sind sich Lehrplanautoren offensichtlich bei der Didaktik und Methodik des naturwissenschaftlichen Unterrichts in Klasse 1/2. Dies gilt nicht für Biologie, die in den meisten Lehrplänen für alle Klassenstufen hohe Präferenz hat. Wie Strunck et al. (1998) haben wir große länderspezifische Unterschiede festgestellt. Es bleibt abzuwarten, ob sich die Diskussion um TIMSS auf den Sachunterricht der Grundschule und neue Lehrpläne der nächsten Jahre auswirkt.

#### 4. Schluss

(a) Die empirische Sachunterrichtsforschung ist in Deutschland u.a. wegen infrastruktureller Mängel nicht gut ausgebaut (wenige Professuren, kaum wissenschaftliche Mitarbeiter, geringe Finanzmittel). Dabei gäbe es dringenden Forschungsbedarf vor allem zu Neuansätzen einer Didaktik des naturwissenschaftlichen Sachunterrichts.

Einigermaßen befriedigend ist die Forschungslage im Bereich Methoden der Veranschaulichung/Visualisierung/Repräsentationsmodi. Außerdem kann man in diesem Feld auf medienpsychologische Forschungsüberblicke zurückgreifen, die häufig auf den Sachunterricht anwendbar sind (z. B. Weidenmann 1995, 1996; Schnotz 1994, 1998). Der Vorteil dieser Forschung (im Gegensatz zur alten Methodenforschung) ist ihre Rückbindung an die Kognitionspsychologie und an Theorien der Wissensrepräsentation. Leider werden die Ergebnisse dieser Forschung zur Bedeutung der Darstellung von Struktureinsichten in der Praxis noch wenig beachtet.

Es fehlt allerdings an empirischen Sachunterrichtsuntersuchungen zu Methoden wie Entdeckungslernen, problemorientiertes Vorgehen, offener Unterricht. Eine Ausnahme bildet die Studie von Hartinger (1997) zu Zusammenhängen zwischen stärker selbstgesteuertem Lernen und Interessenentwicklung.

(b) Die Forschung zu Schülervorstellungen ist z. T. als positiv, z. T. als defizitär zu kennzeichnen. Erfreulich ist, dass es eine relativ umfangreiche Forschung zur Beschreibung von Schülervorstellungen gibt, wobei Erhebungen zu Konzepten der Kinder im soziokulturellen Bereich zahlreicher sein dürften. Negativ zu beurteilen ist, dass viele Studien in diesem Bereich bei der Deskription stehen bleiben und nicht untersuchen, wie unterrichtliche Interventionen den Konzeptwandel beeinflussen. Jedoch ist das umfangreiche Forschungsprogramm von Kornelia Möller hervorzuheben, das genau hier ansetzt und das sie wie folgt beschreibt: „Die nach dem Schema Präkonzepterhebung, Lehr-Lernprozess-Dokumentation (mit Videoerfassung) und Postkonzepterhebung angelegten Untersuchungen bestehen aus der

inhaltsanalytischen Auswertung der transkribierten Schüleräußerungen im Hinblick auf Verstehen“ (Möller 1999, S. 1).

(c) Die Lehrplananalysen und die Erhebungen zu im Unterricht behandelten Lehrstoffen fördern ein verwirrendes Bild zutage. Sie bestätigen das Phänomen, das Fölling-Albers (1993, S. 10) als fehlende konsistente Identität des Sachunterrichts bezeichnet. Soziokulturelle Themen haben ein stärkeres Gewicht als naturwissenschaftliche, wobei allerdings biologische Fachanteile relativ umfangreich sind. Bei der Entwicklung des Sachunterrichtslehrplans Bayern 2001 wurde die öffentliche Diskussion um naturwissenschaftliche Bildung aufgegriffen; der Lehrplan enthält wieder mehr Themen aus Physik/Chemie/Technik.

Insgesamt ist die Forschung zur Ist-Lage des Sachunterrichts zu wenig ausgeprägt. Erforderlich wären repräsentative Beobachtungs- und Befragungsstudien, um das thematische Spektrum, das Anspruchsniveau und die praktizierten Unterrichtsmethoden in Erfahrung zu bringen.

## Literatur

- Adler, W., Cronjäger, H., Murkisch, G., Spreckelsen, K. & Ulrich, B. (1973). Einflüsse auf den Lernerfolg im naturwissenschaftlichen Unterricht der Grundschule. *Naturwissenschaften im Unterricht*, 21, 56-62.
- Beck, G., Scholz, G. & Walter, Ch. (1991). Szenen – Absichten – Deutungen. Zwei Jahre Auseinandersetzung mit moralischen Fragen. *Die Grundschulzeitschrift*, H. 50, 14-19.
- Bittinger, L., Flügge, K., v. Alt-Stutterheim, W., Müller, I. & Andreas, R. (1980). *Lehrer und Lehrplan in der Grundschule*. München: Ehrenwirth.
- Blaseio, B. (1999). *Entwicklungslinien der Unterrichtsinhalte des Sachunterrichts seit 1970*. Eine Inhaltsanalyse von Lehrwerken des Sachunterrichts. Vortrag auf der GDSU-Tagung 1999 in Bielefeld.
- Bredderman, T. (1983). Effects of activity-based elementary science on student outcomes: A quantitative synthesis. *Review of Educational Research*, 53, 499-518.
- Breitschuh, G. (1997). *Inhalte des Sachunterrichts im 4. Schuljahr*. Vortrag auf der GDSU-Tagung 1997 in Kiel.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: The MIT Press.
- Claus, J., Stork, E. & Wiesner, H. (1982). Optik im Sachunterricht. *Sachunterricht und Mathematik in der Primarstufe*, 10, 82-92.
- Düker, H. (1968). Veranschaulichung und Unterrichtserfolg. *Heilpädagogische Forschung*, 3. Beiheft, 99-106.
- Duit, R. (1997). Alltagsvorstellungen und Konzeptwechsel im naturwissenschaftlichen Unterricht – Forschungsstand und Perspektiven für den Sachunterricht in der Primarstufe. In W. Köhlein, B. Marquardt-Mau & H. Schreier (Hrsg.), *Kinder auf dem Wege zum Verstehen der Welt* (S. 233-246). Bad Heilbrunn: Klinkhardt.
- Einsiedler, W. (1998). The curricula of elementary science education in Germany. In *Mathematics and Elementary Science Education* (pp. 25-40). Berlin: Japanisch-Deutsches Zentrum.
- Einsiedler, W. & Martschinke, S. (1998). Elaboriertheit und Strukturiertheit in Schulbuchillustrationen des Grundschulsachunterrichts. In G. Dörr & K. L. Jüngst (Hrsg.), *Lernen mit Medien* (S. 45-65). Weinheim: Juventa.

- Einsiedler, W. & Schirmer, G. (1986). Sachunterrichtsreform und Unterrichtsgestaltung – Eine Analyse von Schülerarbeitsmappen. *Die Deutsche Schule*, 78, 316-326.
- Faust-Siehl, G. (1993). Mit Kindern Naturphänomene verstehen. Sachwissen, Kindervorstellungen und Unterrichtsbeobachtungen zum Thema „Schatten“. *Die Grundschulzeitschrift*, 7, H. 67, 8-16.
- Faust-Siehl, G. & Schweitzer, F. (1991). Rettung unter Einsatz des eigenen Lebens. Begründungsansätze und pädagogische Chancen in Moraldiskussionen mit Grundschulkindern. *Die Grundschulzeitschrift*, H. 50, 42-45.
- Fölling-Albers, M. (1993). Der Sachunterricht in der Grundschule – Auf der Suche nach einem Profil. In D. Richter (Hrsg.), *Grundlagen des Sachunterrichts* (S. 9-19). Oldenburg: Zentrum für pädagogische Berufspraxis.
- Hagstedt, H. & Spreckelsen, K. (1986). Wie Kinder physikalischen Phänomenen begegnen. *Sachunterricht und Mathematik in der Primarstufe*, 14, 318-323.
- Harteringer, A. (1997). *Interessensförderung. Eine Studie zum Sachunterricht*. Bad Heilbrunn: Klinkhardt.
- Höcker, G. (1968). Inhalte des Sachunterrichts im 4. Schuljahr. Eine kritische Analyse. *Die Grundschule*, 1, H. 3, 10-14.
- Janke, B. (1995). Entwicklung naiven Wissens über den physikalischen Auftrieb: Warum schwimmen Schiffe? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 122-138.
- Kircher, E. & Engel, Ch. (1994). Schülervorstellungen über Schall. *Sachunterricht und Mathematik in der Primarstufe*, 22, 53-67.
- Kircher, E. & Rohrer, H. (1993). Schülervorstellungen zum Magnetismus in der Primarstufe. *Sachunterricht und Mathematik in der Primarstufe*, 21, 336-342.
- Lück, G. (1999). *Naturwissenschaften im frühen Kindesalter. Untersuchungen zur Primärbegegnung von Kindern im Vorschulalter mit Phänomenen der unbelebten Natur*. Kiel, Habilitationsschrift.
- Martschinke, S. (1996). Der Aufbau mentaler Modelle durch bildliche Darstellungen. Eine experimentelle Studie über die Bedeutung der Merkmalsdimensionen Elaboriertheit und Strukturiertheit im Sachunterricht der Grundschule. *Zeitschrift für Pädagogik*, 42, 215-232.
- Michael, B. (1983). *Darbieten und Veranschaulichen*. Bad Heilbrunn: Klinkhardt.
- Möller, K. (1999). *Lehr-Lernprozessforschung im naturwissenschaftlich-technischen Bereich des Sachunterrichts*. Münster (Manuskript).
- Möller, K., Tenberge, C. & Ziemann, U. (1996). *Technische Bildung im Sachunterricht*. Münster: Veröffentlichungen der Abteilung Didaktik des Sachunterrichts, Bd. 2.
- Paivio, A. (1986). *Mental representations*. New York: Oxford University Press.
- Pfundt, H. & Duit, R. (1994). *Bibliographie Alltagsvorstellungen und naturwissenschaftlicher Unterricht*. Kiel: IPN (4. Aufl.) (aktueller Stand jeweils als Computerdatei).
- Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63, 167-199.
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accomodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211-228.
- Richter, D. (2000). Politisches Lernen in der Grundschule. *Grundschule*, 32, H. 4, 30-34.
- Schnotz, W. (1994). Wissenserwerb mit logischen Bildern. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern* (S. 95-147). Bern: Huber.
- Schnotz, W. (1998). Visuelles Lernen. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 556-560). Weinheim: Psychologie Verlags Union.
- Schnotz, W., Zink, Th. & Pfeiffer, M. (1996). Visualisierungen im Lehr-Lern-Prozess. *Zeitschrift für Pädagogik*, 42, 193-213.
- Schreier, H. (1979). *Sachunterricht. Themen und Tendenzen*. Paderborn: Schöningh.

- Stern, E., Koerber, S. & Rode, C. (1999). *Enhancing Knowledge Transfer and Efficient Reasoning by Practicing Representation in Science Education*. Paper presented at the Biennial Meeting of the Society for Research in Child Development. Albuquerque, USA.
- Stork, E. & Wiesner, H. (1981). Schülervorstellungen zur Elektrizitätslehre und Sachunterricht. *Sachunterricht und Mathematik in der Primarstufe*, 9, 218-230.
- 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. *Zeitschrift für Didaktik der Naturwissenschaften*, 4, 69-80.
- Treinius, G. & Einsiedler, W. (1993). Hierarchische und bedeutungsnetzartige Lehrstoffdarstellungen als Lernhilfen beim Wissenserwerb im Sachunterricht der Grundschule. *Psychologie in Erziehung und Unterricht*, 40, 263-277.
- Vosniadou, S. & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535-585.
- Vosniadou, S. & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18, 123-183.
- Weidenmann, B. (1988). *Psychische Prozesse beim Verstehen von Bildern*. Bern: Huber.
- Weidenmann, B. (1995). Multicodierung und Multimodalität im Lernprozeß. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 65-84). Weinheim: Psychologie Verlags Union.
- Weidenmann, B. (1996). Instruktionsmedien. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion*. Enzyklopädie der Psychologie D, I, 2 (S. 319-368). Göttingen: Hogrefe.
- Wiesner, H. (1991). Schwimmen und Sinken: Ist Piagets Theorie noch immer eine geeignete Interpretationshilfe für Lernvorgänge? *Sachunterricht und Mathematik in der Primarstufe*, 19, 2-6.
- Wiesner, H. (1995). Untersuchungen zu Lernschwierigkeiten von Grundschulern in der Elektrizitätslehre. *Sachunterricht und Mathematik in der Primarstufe*, 23, 50-58.
- Wiesner, H. & Claus, J. (1985). Vorstellungen zu Schatten und Licht bei Schülern der Primarstufe. *Sachunterricht und Mathematik in der Primarstufe*, 13, 318-322.
- Wiesner, H. & Stengl, D. (1984). Vorstellungen von Schülern der Primarstufe zu Temperatur und Wärme. *Sachunterricht und Mathematik in der Primarstufe*, 12, 445-452.
- Wilkening, F. (1994). Kinder sind schlauer als manche Pädagogen meinen. In F. Rösler & I. Florin (Hrsg.), *Psychologie und Gesellschaft* (S. 89-97). Stuttgart: Hirzel.
- Wulf, P. & Euler, M. (1995). Ein Ton fliegt durch die Luft – Vorstellungen von Primarstufenkindern zum Phänomen Schall. *Physik in der Schule*, 33, 7-8, 254-260.

# Forschungsmethoden und Erhebungsverfahren



## Zum Design empirischer Untersuchungen in der Lehr-Lernforschung

### 1. Einleitung

Das Kennzeichen von wissenschaftlichen Aussagen ist die Methodisierung der Erfahrungsgewinnung: Es sind die Methoden einer Disziplin, die Forschung von bloßer Beobachtung und Spekulation trennen. Eine wissenschaftliche Aussage soll nicht durch eloquente Formulierung oder oberflächliche Plausibilität überzeugen, sondern dadurch, dass die zu Grunde liegenden Daten, die Argumente und Schlussfolgerungen von anderen Mitgliedern der wissenschaftlichen Gemeinschaft sorgfältig überprüft und nachvollzogen werden können. In den Sozialwissenschaften gibt es verschiedenartige methodische Zugriffe. Eine Dogmatisierung eines methodischen Zugriffs – und damit der Ausschluss von anderen Zugriffen – ist unangemessen, da Methoden nicht isoliert existieren, sondern zusammenhängen mit bestimmten Fragestellungen, Blickrichtungen und theoretischen Ansätzen. In der Regel dürfte sich ein vollständiges Bild eines Gegenstandes erst dann ergeben, wenn Ergebnisse aus verschiedenen Perspektiven – und damit gewonnen aufgrund verschiedener methodischer Zugriffe – zusammengetragen werden. So sind weder „qualitative“ noch „quantitative“ Methoden die alleinigen Forschungsmethoden in der Erziehungswissenschaft; vielmehr wird eine Vielfalt der methodischen Zugriffe benötigt (vgl. Roßbach 1996; Shulman 1988).

Dieses Plädoyer für eine Methodenvielfalt und Mehrperspektivität in der erziehungswissenschaftlichen Forschung meint aber keine Beliebigkeit der methodischen Zugriffe. An alle Methoden bzw. methodischen Zugriffe müssen gemeinsame Grundanforderungen gestellt werden, wenn Erkenntnisse auf für andere nachvollziehbaren und nachprüfbareren Wegen gewonnen werden sollen. Zu nennen sind hier vor allem die Gütekriterien Objektivität, Zuverlässigkeit und Gültigkeit. Sicher wird es je nach methodischem Zugriff Unterschiede in der Art und Weise geben, wie diese Gütekriterien festgestellt werden. Die damit verbundenen Fragestellungen und Konzeptideen aber gelten für alle Methoden (vgl. von Saldern 1995).

Der folgende Beitrag beansprucht nicht, einen Überblick über quantitative Forschungsmethoden in der Lehr-Lernforschung im Sachunterricht und ihren besonderen Möglichkeiten – u.U. in Abgrenzung von qualitativen Methoden – zu geben. Vielmehr erfolgt eine Beschränkung auf Designfragen empirischer Untersuchungen, d.h. auf Aspekte der Anlage einer Untersuchung. Damit wird ein Problembereich thematisiert, der von größter Bedeutung für die Qualität einer Untersuchung ist. Auch die besten – quantitativen oder qualitativen – Erhebungsmethoden, die innerhalb eines Designs eingesetzt werden, können nicht die Schwächen eines unzureichenden Designs kompensieren. Je nach Untersuchungsziel werden an die Güte eines Designs unterschiedliche Anforderungen gestellt (Wellenreuther 2000):

- Besteht das Untersuchungsziel darin, Auftreten und Häufigkeiten eines Phänomens in der Realität zu beschreiben – wie z. B. die Einstellungen von Sachunterrichtslehrerinnen zu naturwissenschaftlichen Themen –, so wird eine repräsentative Stichprobe benötigt, d.h., die Stichprobe, an der die Untersuchung durchgeführt wird, muss weitgehend ein verkleinertes unverfälschtes Abbild der Grundgesamtheit sein, für die die Aussagen gemacht werden sollen. Im Beispiel interessiert dabei letztlich nicht, ob es Lehrerinnen mit naturwissenschaftsfeindlichen Einstellungen gibt, sondern wie häufig diese in der entsprechenden Grundgesamtheit sind. Realistische Abschätzungen der Verhältnisse in der Grundgesamtheit sind erforderlich, um nicht falschen Schlüssen für Reformmaßnahmen zu erliegen. *Repräsentativität als Gütekriterium* lässt sich über Zufallsstichproben aus der Grundgesamtheit erreichen, wobei größere Stichproben insofern von Vorteil sind, als damit der Stichprobenfehler reduziert werden kann. Zufallsstichproben sind in der Regel aber teuer und oftmals – gerade im Rahmen von Qualifikationsarbeiten – nur eingeschränkt oder gar nicht zu realisieren. Es gibt aber auch Möglichkeiten, jenseits von Zufallsstichproben durch eine bewusste und gezielte Auswahl der Stichprobenelemente anhand von ausgewählten und theoretisch begründbaren Merkmalen sich dem Ideal eines verkleinerten Abbildes der Grundgesamtheit anzunähern. Ein sorgfältiger Abgleich der Verteilung ausgewählter Merkmale in der Stichprobe mit deren Verteilung in der Grundgesamtheit ist dabei erforderlich – was allerdings voraussetzt, dass die Verteilung dieser Merkmale in der Grundgesamtheit bekannt ist.
- Besteht das Untersuchungsziel darin, eine Hypothese zu überprüfen – wie z. B., dass eine bestimmte Unterrichtseinheit oder Unterrichtsform anderen Einheiten oder Formen überlegen ist –, so ist Repräsentativität weniger wichtig bzw. unwichtig. Das Design ist dann gut, wenn durch die Untersuchungsanlage die Hypothese, dass die Variation einer unabhängigen Variablen zu bestimmten Veränderungen in einer abhängigen Variablen führt, geprüft und mögliche al-



ternative Erklärungen für die Veränderungen in der abhängigen Variablen ausgeschlossen werden können. Das *Gütekriterium für hypothesentestende Untersuchungen* ist somit der *Ausschluss alternativer Erklärungen*. Im Beispiel der Überprüfung der Effekte einer bestimmten Unterrichtseinheit sollen die Effekte tatsächlich auf diese Unterrichtseinheit zurückzuführen sein und nicht z.B. darauf, dass die Klasse, die die Unterrichtseinheit erhält, schon vorher bessere Leistungen, höhere Motivationen usw. aufzeigt als die Kontrollklasse, die einen anderen Unterricht erfährt, oder darauf, dass in der Versuchsklasse engagiertere, fähigere Lehrkräfte eingesetzt sind als in der Kontrollklasse. Ausschluss von Alternativverklärungen kann in der hypothesentestenden Forschung auf verschiedenen Wegen erreicht werden: über Zeitreihen- und Kontrollgruppenexperimente im Labor oder im Feld, über Interventionsstudien oder über Längsschnitt- und Querschnittuntersuchungen, bei denen die Stichprobe so konstruiert wird, dass theoretisch begründbare mögliche Alternativvariablen als Kontrollvariablen eingeführt oder über die verschiedenen Ausprägungen der unabhängigen Variablen konstant gehalten werden – was aber voraussetzt, dass solche mögliche Alternativvariablen bekannt sind.

Im Folgenden wird aus Platzgründen nur noch auf das Kontrollgruppenexperiment und seine Möglichkeiten und Güte eingegangen. Dabei muss aber im Auge behalten werden, dass empirische Forschung nicht aus isolierten Einzeluntersuchungen besteht. Vielmehr muss stärker der Prozesscharakter von Forschung berücksichtigt werden: Feld- und Laborexperimente wechseln sich ab und kompensieren dabei ihre jeweiligen Nachteile; Experimente, Interventionsstudien sowie Längs- und Querschnittstudien ergänzen sich; in qualitativen Einzelfallstudien werden bestimmte Phänomene und Beziehungen zwischen Phänomenen und ihre Vermittlungsmechanismen in der Tiefe aufgedeckt und in nachfolgenden Repräsentativitätsuntersuchungen oder Experimenten überprüft usw. Die Beschränkung auf das Kontrollgruppenexperiment heißt nicht, dass ab nun nur noch solche strengen Experimente im Sachunterricht durchgeführt werden sollen. Vielmehr eignen sich die Überlegungen zur Güte von Kontrollgruppenexperimenten besonders gut, um zu verdeutlichen, was Ausschluss alternativer Erklärungen bedeuten kann.

## 2. Experimentelle Versuchspläne und ihre Güte

Die folgenden Ausführungen orientieren sich an Beiträgen, die als Standardliteratur zu diesem Themenkomplex zu betrachten sind (vgl. Schwarz 1973; Roberts & Rost 1974; Wellenreuther 2000; Zimmermann 1972).

In einem Experiment geht es um die Überprüfung der Hypothese, dass die Variation in einer unabhängigen Variablen (Ursache) zu einer bestimmten Variation in

einer abhängigen Variablen (Wirkung) führt. Im Experiment variieren (manipulieren) wir gezielt die unabhängige Variable, um die Effekte auf die abhängige Variable prüfen zu können. Bei einem Kontrollgruppenexperiment wird in der sogenannten Experimentalgruppe der experimentelle Stimulus – d.h. eine bestimmte Ausprägung der unabhängigen Variablen – eingeführt und in seinen Auswirkungen im Vergleich zu der Situation in der Kontrollgruppe – in der eine andere Ausprägung der unabhängigen Variablen gegeben ist – verglichen.<sup>1</sup> Zwei Fragen sind hierbei zentral:

- Wird die theoretische Hypothese angemessen in den unterschiedlichen Versuchsbedingungen operationalisiert?
- Werden mögliche Störfaktoren kontrolliert? Werden mögliche alternative Erklärungen ausgeschlossen?

## 2.1 Umsetzung der theoretischen Hypothese

Die theoretische Hypothese ist nicht mit dem Experiment identisch. Vielmehr muss sie in die unterschiedlichen Ausprägungen der unabhängigen Variablen umgesetzt werden. Hierzu ein Beispiel für ein Unterrichtsexperiment: Die theoretische und vor dem Hintergrund bisheriger Untersuchungen begründete Hypothese sei, dass ein Sachunterricht, der nach den Prinzipien A und B aufgebaut ist und in dem explizit Lernstrategien vermittelt werden, für den kognitiven Wissenserwerb der Schüler vorteilhafter ist als ein Sachunterricht, in dem dies nicht der Fall ist. Dazu plant der Forscher ein Experiment mit sechs Grundschulklassen, wobei in vier Klassen ein Sachunterricht nach den Prinzipien A und B und mit Vermittlung von Lernstrategien erteilt werden soll, während die beiden restlichen Klassen als Kontrollklassen betrachtet werden. Der Forscher trainiert vier Lehrer in den Prinzipien A und B und in der Vermittlung von Lernstrategien; diese Lehrer werden dann in den Experimentalklassen eingesetzt. Am Ende einer dreimonatigen Phase wird dann in beiden Gruppen ein Wissenstest eingesetzt.<sup>2</sup>

Zunächst muss hier gefragt werden, ob der in den Experimentalklassen realisierte Unterricht wirklich den theoretisch begründeten Prinzipien A und B und der angezielten Vermittlung von Lernstrategien entspricht. Wurden in dem Lehrertraining wirklich die theoretischen Ideen zu dem Unterricht thematisiert? Haben die Lehrer diese Prinzipien auch tatsächlich in ihrem Unterricht umgesetzt? Während die erste Frage nach theoretischen Analysen verlangt, ist die zweite Frage nur durch gezielte Beobachtungen des Unterrichts in den Experimentalklassen zu beantworten. Solche gezielten Beobachtungen der Umsetzung des experimentellen Stimulus werden oftmals vernachlässigt.

Die Frage nach der Umsetzung der theoretischen Hypothese bezieht sich aber nicht nur auf den Unterricht in den Experimentalklassen, sondern auch auf die

Kontrollklassen. Gegen welchen Unterricht will sich der Unterricht nach den Prinzipien A und B und mit Vermittlung von Lernstrategien absetzen? Oftmals werden als Kontrollklassen solche mit einem „herkömmlichen“ Unterricht herangezogen, ohne dass dieser genauer spezifiziert wird. Es ist durchaus zu fragen, ob ein solches Vorgehen angemessen ist. Vielleicht wären Kontrollklassen angemessener, in denen der Unterricht nicht herkömmlich, sondern nach einem spezifischen Alternativkonzept erteilt wird oder in denen zwar die Prinzipien A und B realisiert werden, nicht aber die Vermittlung der Lernstrategien. Solche Fragen können nur theoretisch und in Abhängigkeit des Wissensstandes über einen Bereich beantwortet werden. Der Forscher sollte es sich hier schwer machen, d. h., er sollte die Kontrollgruppe wählen, die den härtesten Test seiner Hypothese ermöglicht. Auf jeden Fall sollte auch der Unterricht in den Kontrollklassen sorgfältig beobachtet werden, um zu wissen, welcher Unterricht dort realisiert wird.

Ein besonderes Problem bei Unterrichtsexperimenten stellt immer die Auswahl der Lehrer dar (s.u.). Wir können in der Regel nicht oder schlecht die Lehrer per Zufall der Experimental- und Kontrollgruppe zuweisen. So könnte es in dem Beispiel sein, dass die Lehrer in den Experimentalklassen besonders fähige, erfahrene und engagierte Lehrer sind (die zudem noch ein Training in einem besonderen pädagogischen Konzept erfahren haben), während in den Kontrollklassen wenig erfahrene und demotivierte Lehrer unterrichten. Hier kommt es dann zu einer Konfundierung zwischen den Ausprägungen der unabhängigen Variablen und den Lehrermerkmalen, d. h., wir wissen nicht – Effekte in den abhängigen Variablen vorausgesetzt –, ob diese Effekte auf den besonderen Unterricht oder auf die Lehrermerkmale zurückzuführen sind. Der Forscher sollte hier nach einem möglichst fairen Test seiner Hypothese suchen und z. B. versuchen, in der Kontrollgruppe gleich fähige, erfahrene und engagierte Lehrer einzusetzen.

Eine Möglichkeit, die Umsetzung der theoretischen Hypothese in ein konkretes Experiment zu überprüfen, ist, das Experiment mit leicht abgewandelter Umsetzung der unabhängigen Variablen zu wiederholen. Bestätigt sich weiterhin die Hypothese, so kann dies als Hinweis auf die Gültigkeit der Hypothese gewertet werden. Bestätigt sich die Hypothese aber nicht, so kann die Hypothese falsch sein oder sie wurde unzureichend in dem konkreten Experiment operationalisiert.

## 2.2 Kontrolle von Störfaktoren

Das Ziel des Kontrollgruppenexperimentes ist es, den Zusammenhang zwischen der unabhängigen und der abhängigen Variablen klar aufzuweisen. Alternative Erklärungen für Veränderungen in den abhängigen Variablen sollen weitestgehend ausgeschlossen werden. Dies wird nicht immer explizit möglich sein. Das Nachdenken über alternative Erklärungsmöglichkeiten bzw. Störfaktoren ist dennoch aber

zentral, um die Tragweite von durch Experimente gewonnenen Aussagen abschätzen zu können und um möglicherweise für praktische pädagogische Maßnahmen nicht Fehlinterpretationen aufzusitzen. Dabei muss zwischen der internen und der externen Validität oder Gültigkeit unterschieden werden:<sup>3</sup>

- *Interne Validität* bezieht sich darauf, ob innerhalb des gegebenen Designs die Veränderung der abhängigen Variablen auf die Variation der unabhängigen Variablen zurückgeführt werden kann oder ob dafür andere Einflussgrößen verantwortlich sind. D.h., nicht die Variation der unabhängigen Variablen hat sich ausgewirkt, sondern eine andere Störgröße hat zu den beobachteten Wirkungen geführt. Man spricht hier auch von *Haupteffekten*.
- *Externe Validität* bezieht sich darauf, inwieweit die Bedingungen des Experiments von denen abweichen, die in der Realität gelten. Externe Validität betrifft damit die Generalisierbarkeit von – intern validen – Ergebnissen über die Untersuchungsstichprobe hinaus:<sup>4</sup> Gelten die gefundenen Ergebnisse nur für die im Experiment untersuchte Gruppe oder lassen sie sich darüber hinaus verallgemeinern? Man spricht hier auch von *Interaktionseffekten*, da die Störgrößen der externen Validität zusammen mit der Variation der unabhängigen Variablen zu den beobachteten Wirkungen geführt haben.

Campbell und Stanley (Schwarz 1973) haben in einem schon klassischen Artikel im Handbook of Research on Teaching verschiedene Faktoren aufgezählt, die die innere und äußere Validität beeinflussen. Im Folgenden sollen diese an zwei Beispielen erläutert werden, einem „vorexperimentellen“ Versuchsplan und einem „echten experimentellen“ Versuchsplan.

### *Ein vorexperimenteller Versuchsplan und seine Validität*

In diesem Design wird eine Experimentalgruppe – eine Gruppe von Schülern, Klassen usw. – einem Vortest (Beobachtung, Test) zum Zeitpunkt 1 unterzogen ( $O_1$ ), dann wird der experimentelle Stimulus (X) eingeführt. Nach Ablauf des Experiments wird eine weitere Messung zum Zeitpunkt 2 ( $O_2$ ) durchgeführt. Es existiert keine Kontrollgruppe. Schematisch lässt sich dieser Versuchsplan wie folgt darstellen:<sup>5</sup>

EG:  $O_1$     X     $O_2$

In der Auswertung wird dann der Unterschied bzw. die Veränderung zwischen  $O_1$  und  $O_2$  betrachtet. Dieser Versuchsplan zeigt im Hinblick auf die interne Validität, d.h. auf die Möglichkeit, die Variation der abhängigen Variablen im Experiment tatsächlich auf den experimentellen Stimulus zurückzuführen, deutliche Schwächen. Folgende acht Störfaktoren, Haupteffekte oder alternative Erklärungen können nach Campbell und Stanley betrachtet werden:

– *Zwischenzeitliches Geschehen*: Die Veränderung zwischen  $O_1$  und  $O_2$  kann mög-

licherweise auf Ereignisse zurückgeführt werden, die in keinem Zusammenhang mit der experimentellen Behandlung stehen. Besteht X z.B. aus einer speziellen Unterrichtseinheit, mit der die Schüler drei Monate unterrichtet werden, so kann in diesem Zeitraum im und außerhalb des Unterrichts vieles passieren, das die abhängige Variable beeinflusst, aber nicht in Zusammenhang mit X steht. Diese alternative Erklärung wird umso bedeutsamer, je größer der Zeitabstand zwischen den beiden Messungen ist.

- *Reifung/Entwicklung*: Versuchspersonen ändern sich im Verlaufe der Zeit in vielerlei Hinsichten; sie werden älter, erweitern ihr Wissen, vergrößern ihr Verhaltensrepertoire usw. Alle diese Reifungs-/Entwicklungsprozesse können die Veränderung zwischen den beiden Zeitpunkten bewirkt haben, ohne dass sich der experimentelle Stimulus ausgewirkt hat. Gerade bei längerandauernden Experimenten müssen solche Prozesse als mögliche Alternativerklärungen berücksichtigt werden.
- *Testungseffekt*: Das erste Testen selbst ( $O_1$ ) kann unabhängig von X einen Einfluss auf die zweite Messung ( $O_2$ ) haben. Schüler verhalten sich z. B. bei einer zweiten Testung anders als bei der ersten Untersuchung mit demselben Test. In der Regel erzielen sie bessere Ergebnisse, da sie den Test kennen und ihnen die Testsituation vertrauter ist oder sie sich an vorherige Lösungen erinnern. Einstellungsmessungen können z.B. auch alleine zu einem Einstellungswandel führen, da die Versuchspersonen durch die Konfrontation mit der ersten Erhebung zu Denkprozessen angeregt werden können. Es gibt allerdings auch Hinweise darauf, dass solche Effekte einer ersten Messung dann kaum von Bedeutung sind, wenn sich die Tests nicht wesentlich von den üblichen Prüfungsarbeiten in den Schulen unterscheiden.
- *Hilfsmittel*: Auch die Hilfsmittel bzw. Messinstrumente können sich zwischen den beiden Zeitpunkten verändern. Werden bei den beiden Messungen z.B. Videoaufnahmen der Schüler im Hinblick auf selbständiges Verhalten in Lernprozessen analysiert, so können die Auswerter, die möglicherweise aus der ersten Auswertungsphase gelernt haben, bei der zweiten Auswertung u. U. viel subtilere Verhaltensweisen schon als Anzeichen von selbständigem Verhalten kodieren als bei der ersten Messung. Diese alternative Erklärung für Veränderungen liegt dann besonders nahe, wenn der Forscher selbst die Beobachtungen kodiert oder der Datenauswerter die Hypothese der Untersuchung kennt.
- *Regression zur Mitte*: Jede Messung ist mit einem Fehler behaftet, so dass sich bei zwei Messungen beim gleichen Schüler aufgrund der Fehlerbehaftetheit unterschiedliche Ergebnisse zeigen können. Die Fehler können dazu führen, dass die wahren Werte der Schüler über- oder unterschätzt werden. Wählen wir nun z.B. aufgrund einer ersten Messung die Extremgruppe der besonders schlechten Schüler aus, so können wir erwarten, dass bei der zweiten Messung auch ohne

jeden experimentellen Stimulus die Messergebnisse sich in Richtung auf den Mittelwert der Gesamtschülergruppe hinbewegen. Der Grund dafür ist, dass es unwahrscheinlicher ist, bei der zweiten Messung den Fehler noch einmal in Richtung auf eine Unterschätzung der Schülerleistungen zu machen. Falls im vorliegenden Beispiel die Experimentalgruppe aufgrund von extremen Leistungen in  $O_1$  ausgewählt wurden, so geht zumindest ein Teil der Verbesserung zwischen  $O_1$  und  $O_2$  auf diese Regression zur Mitte zurück und nicht auf die Wirkung von X.

- *Auswahl der Versuchspersonen:* Dieser Störfaktor bezieht sich bei dem Vergleich von zwei Gruppen (Experimental- und Kontrollgruppe) darauf, dass die beiden Gruppen sich schon vor dem Experiment unterscheiden und deshalb die Unterschiede zwischen Experimental- und Kontrollgruppe nach dem Experiment nicht auf die Wirkung von X, sondern auf die schon vorher bestehenden Unterschiede zwischen den Gruppen zurückzuführen sind. Da bei dem vorexperimentellen Versuchsplan aber nur eine Gruppe betrachtet wird, kann diese Alternativerklärung nicht auftreten.
- *Einbußen von Versuchspersonen:* Eine experimentelle Behandlung kann bei verschiedenen Versuchspersonen zu unterschiedlichen Reaktionen führen, z.B. auch dazu, dass sie nicht mehr an dem Experiment teilnehmen wollen. Falls dies bei Experimenten im schulischen Kontext möglich ist, kann es sein, dass z.B. die schlechteren, unmotivierteren, uninteressierteren Schüler ausscheiden und sich allein dadurch der Mittelwert in diesen Variablen zum zweiten Zeitpunkt verbessert, ohne dass irgendein experimenteller Stimulus gewirkt hat. Ausscheidende Versuchspersonen müssen deshalb sorgfältig auf die Gründe ihres Ausscheidens hin analysiert werden.
- *Wechselwirkung zwischen Auswahl und Reifung/Entwicklung:* In Experimenten, in denen eine Experimentalgruppe mit einer Kontrollgruppe verglichen wird, kann es vorkommen, dass Reifungs-/Entwicklungsprozesse in den beiden Gruppen unterschiedlich ablaufen, z.B. wenn ein Unterrichtsexperiment mit Drittklässlern durchgeführt wird, die Kontrollgruppe aber aus Zweitklässlern besteht. Da es im vorliegenden Fall aber keine Kontrollgruppe gibt, ist dieses Problem hier irrelevant.

Es mag im konkreten Fall einer Untersuchung plausible theoretische Gründe geben, warum diese Störfaktoren oder einige von ihnen sich im Experiment nicht haben auswirken können, im Allgemeinen aber können diese Störfaktoren durch dieses vorexperimentelle Design nicht kontrolliert werden. Da interne Validität eine Voraussetzung für externe Validität ist und sich bei diesem Design erhebliche Probleme bei der internen Validität ergeben, soll auf Aspekte der externen Validität nicht eingegangen werden.

### *Ein echter experimenteller Versuchsplan und seine Validität*

Kontrolle von möglichen Störgrößen heißt in der Regel nicht, dass solche Störgrößen eliminiert werden. Schüler haben z.B. ein bestimmtes Geschlecht, eine bestimmte Sozialschichtzugehörigkeit, eine bestimmte Lernausgangslage usw. – Merkmale, in denen sich die Versuchspersonen unterscheiden und die in ihren Wirkungen auf die abhängige Variable nicht ausgeschaltet werden können. Kontrolle von Störgrößen heißt deshalb hier sicherzustellen, dass wir die Wirkungen von solchen Störgrößen „im Griff haben“ und sie uns nicht zu Fehlschlüssen verleiten. Als Paradigma der Kontrolle von Störgrößen kann die *Zufallsaufteilung* oder *Randomisierung* betrachtet werden. Die Untersuchungseinheiten werden per Zufall auf die Experimentalgruppe und auf die Kontrollgruppe aufgeteilt. Alle möglichen Variablen sollten dann – zufallsbedingte Unterschiede akzeptiert – in beiden Gruppen gleich verteilt sein. Es gibt keinen Grund, warum irgendeine Variable systematisch in einer der beiden Gruppen anders verteilt sein sollte. Die Störgrößen wirken dann zwar auch noch weiter, da sie aber in Experimental- und Kontrollgruppe gleich verteilt sind, kann ein Unterschied in einer abhängigen Variablen nach der Durchführung eines Experiments nicht auf sie, sondern nur auf den experimentellen Stimulus zurückgeführt werden – einen angemessenen Gruppenumfang vorausgesetzt.

Zufallsaufteilungen sind in der Unterrichtsforschung und bei der Durchführung von Unterrichtsexperimenten in Schulen nur schwierig oder oftmals praktisch nicht zu erreichen. Ebenfalls schwierig ist der zweite Weg, um zu vergleichbaren Gruppen vor dem Beginn des Experiments zu kommen: das *Matchen* oder die *Parallelisierung*. Hierbei werden die Untersuchungseinheiten nicht zufällig auf Experimental- und Kontrollgruppe aufgeteilt, vielmehr werden anhand von ausgewählten Variablen gleiche Paare gebildet, von denen dann je ein Element in der Experimentalgruppe und das andere in der Kontrollgruppe ist.<sup>6</sup> In der Unterrichtsforschung könnten z. B. gezielt zu Klassen, in denen eine bestimmte Unterrichtseinheit durchgeführt werden soll, andere Klassen als Kontrollklassen gesucht werden, die sich in Klassenstufe, Klassengröße, durchschnittlichem Leistungsniveau und Leistungsstreuung vor Beginn des Experiments sowie sozialer Zusammensetzung der Elternschaft und sozialem Umfeld der Schule nicht von den Experimentalklassen unterscheiden. Die Schwäche einer Parallelisierung ist, dass gleiche Gruppen nur im Hinblick auf die für die Parallelisierung ausgewählten Variablen erreicht werden und man nichts über die Verteilung anderer möglicher Störgrößen weiß.

Im folgenden Design werden Untersuchungseinheiten zufällig auf eine Experimental- und auf eine Kontrollgruppe aufgeteilt. Zum Zeitpunkt 1 findet in beiden Gruppen ein Vortest statt ( $O_1$  und  $O_3$ ). In der Experimentalgruppe wird dann der experimentelle Stimulus eingeführt, während dies in der Kontrollgruppe nicht ge-

schieht (zur Frage der Angemessenheit des Geschehens in der Kontrollgruppe s.o.). Nach Ablauf des Experiments wird in Experimental- und Kontrollgruppe eine weitere Messung zum Zeitpunkt 2 ( $O_2$  und  $O_4$ ) durchgeführt. Schematisch lässt sich dieser Versuchsplan wie folgt darstellen:

R EG:	$O_1$	X	$O_2$
R KG:	$O_3$		$O_4$

In der Auswertung wird dann die Differenz zwischen  $O_1$  und  $O_2$  im Vergleich mit jener zwischen  $O_3$  und  $O_4$  betrachtet. Wenn die Differenz zwischen  $O_1$  und  $O_2$  größer (oder je nach Fragestellung kleiner) als jene zwischen  $O_3$  und  $O_4$  ist, geht man von einem Effekt von X aus. Dieser Versuchsplan schneidet im Hinblick auf die interne Validität gut ab. Zwischenzeitliches Geschehen, Reifung/Entwicklung, Testungseffekte, Auswahl der Versuchspersonen und Wechselwirkung zwischen Auswahl und Reifung/Entwicklung können in diesem Beispiel nicht als alternative Erklärungen in Frage kommen, da durch die Zufallsaufteilung diese Störgrößen sich in Experimental- und Kontrollgruppe gleichermaßen auswirken. Die Störgröße Hilfsmittel ist ebenfalls kontrolliert, da sich Veränderungen in den Hilfsmitteln gleichermaßen in den beiden Gruppen auswirken müssten – zumindest dann, wenn vorausgesetzt werden kann, dass den Datenauswertern die Hypothese nicht bekannt ist und sie nicht wissen, ob sie Daten aus der Experimental- oder der Kontrollgruppe auswerten. Der Effekt der Regression zur Mitte ist hier ebenfalls kontrolliert, wenn die Einheiten der Extremgruppe (z. B. Extremgruppe der schlechten Schüler) zufällig auf Experimental- und Kontrollgruppe aufgeteilt werden, da der Regressionseffekt sich in beiden Gruppen gleichermaßen zeigt. Ein gewisses Problem könnte die Einbuße von Versuchspersonen darstellen. Scheiden z. B. schlechtere Schüler aufgrund der experimentellen Behandlung in der Experimentalgruppe, nicht aber in der Kontrollgruppe aus, so ergeben sich allein dadurch Veränderungen im Gruppenmittelwert zum zweiten Zeitpunkt, ohne dass die experimentelle Variable zu wirken braucht.

Die bisherigen Ausführungen bezogen sich auf die interne Validität, d.h. auf die Frage, inwieweit innerhalb des Designs die Veränderungen in der abhängigen Variablen auf die Variation der unabhängigen Variablen zurückgeführt werden kann. Die folgenden Störgrößen beziehen sich auf die externe Validität und damit auf die Generalisierbarkeit von – intern validen – Ergebnissen über die Untersuchungstichprobe hinaus. Campbell und Stanley sprechen vier Störgrößen, Interaktionseffekte oder alternative Erklärungen an:

- *Wechselwirkungen zwischen Vortest (O) und experimentellem Stimulus (X)*: Ein intern valides Ergebnis kann u. U. nur für Gruppen gelten, die vorgetestet wurden. Vortest ( $O_1$ ) und X treten in eine Interaktion, indem der Vortest für X sensibilisiert und X möglicherweise nur dann zu einer Veränderung in der abhängigen Variablen führt, wenn die Versuchspersonen auf diese Weise sensibili-



siert wurden. Sensibilisierung könnte hier auch bedeuten, dass durch den Vortest die Versuchspersonen hinsichtlich der Erwartungen des Experiments sensibilisiert werden. Dieser Interaktionseffekt ist durch den vorliegenden Versuchsplan nicht kontrolliert. Eine Kontrollmöglichkeit würde darin bestehen, – bei Zufallsaufteilung auf Experimental- und Kontrollgruppen – auf den Vortest zu verzichten

- *Wechselwirkung zwischen Auswahl und X*: Es kann sein, dass der experimentelle Stimulus X nur für die ausgewählte Gruppe gilt.<sup>7</sup> Wenn z.B. ein bestimmtes Unterrichtsmodell in einem Experiment mit Zweitklässlern untersucht wurde, kann es durchaus offen sein, ob ein vergleichbarer Effekt des Unterrichtsmodells auch bei Viertklässlern auftritt. Das Vertrauen in eine Generalisierbarkeit der Ergebnisse wächst, wenn das Experiment erfolgreich mit anderen Gruppen wiederholt werden kann. Es könnte hier auch zu Wechselwirkungen zwischen X und dem zwischenzeitlichen Geschehen kommen (z.B. jahreszeitliche Unterschiede; eine neue auf die Fragestellung der Untersuchung bezogene Fernsehsendung usw.), die ebenfalls nur durch Wiederholungen des Experiment zu anderen Zeitpunkten als Alternativerklärungen ausgeschlossen werden können.
- *Effekte der experimentellen Situation*: Die Teilnahme an einem Experiment ist eine besondere Situation. Die gefundenen Ergebnisse könnten deshalb u.U. nur für die Gruppen gelten, die (in Experimental- oder Kontrollgruppe) an dem Experiment teilgenommen haben. Dieser Interaktionseffekt ist durch den Versuchsplan nicht kontrolliert. Findet aber das Experiment z.B. integriert in einen „normalen“ Unterricht statt, so dürfte dieser Effekt eher gering ausfallen.
- *Beeinträchtigung durch mehrere experimentelle Stimuli  $X_1$  bis  $X_n$* : In manchen Designs werden die Versuchspersonen nacheinander mit mehreren experimentellen Stimuli  $X_1$  bis  $X_n$  konfrontiert. Es könnte dann sein, dass ein späterer Stimulus nur wirkt, weil vorher schon ein anderer Stimulus gewirkt hat. Da es in dem vorliegenden Design aber nur einen experimentellen Stimulus X gibt, ist dieser Interaktionseffekt hier irrelevant.

Die Beschreibung der Störgrößen für die interne und externe Validität eines experimentellen Designs ist nicht als erschöpfende Aufzählung zu verstehen. Es gibt weitere Faktoren, die die Gültigkeit von Kontrollgruppenexperimenten beeinflussen können und auf die in der Interpretation der Untersuchungsergebnisse geachtet werden muss. Einzelne Details des Experiments und/oder der Versuchsleiter können in der Experimentalgruppe den Versuchspersonen signalisieren, welches Verhalten von ihnen erwartet wird und damit die interne Validität beeinträchtigen. Neuigkeitseffekte können in der Experimentalgruppe eine Rolle spielen, indem dort mit den Versuchspersonen etwas Neues gemacht wird, was allein schon Veränderungen in den abhängigen Variablen stimulieren kann. Allerdings gibt es auch Hinweise darauf, dass in pädagogischen Untersuchungen solche Neuigkeitseffekte

bisher überschätzt wurden. Problematisch könnte in Experimenten im Unterricht auch sein, dass die Lehrer und/oder Schüler in den Kontrollgruppen merken, dass in anderen Klassen (den Experimentalklassen) etwas Neues, Interessantes, Attraktives geschieht und sie deshalb ihr Verhalten ändern (z.B. sich besonders anstrengen, um „es den Experimentalklassen zu zeigen“, oder im Gegenteil sich enttäuscht gerade weniger anstrengen). Schließlich arbeiten viele Experimente mit Täuschungen, indem z.B. die Versuchspersonen über die wahren Absichten eines Experiments getäuscht werden, was wiederum zu offenem oder verstecktem Widerstand führen kann (vgl. Henningsen 1964).

Die Kontrolle von Störfaktoren in Kontrollgruppenexperimenten und der Ausschluss von Alternativklärungen ist in der Praxis, besonders in Feldexperimenten in der Schule, schwierig. Das bedeutet aber nicht, dass Unterrichtsexperimente zu keinem Erkenntnisgewinn führen. Es wird kaum möglich sein, alle problematischen Aspekte in *einem einzigen* Experiment bearbeiten zu können. Vielmehr verspricht nur das Zusammendenken von verschiedenen Experimenten, die u.U. verschiedenste alternative Erklärungen ausschließen können, einen Erkenntnisgewinn. Gerade dies macht deutlich, wie wichtig das Durchdenken eines Experiments im Hinblick auf Störfaktoren der internen und externen Validität und die Suche nach theoretisch plausiblen Argumenten ist, die alternative Erklärungen ausschließen können.

### 3. Ein Beispiel für eine quasi-experimentelle Untersuchung

Im abschließenden Abschnitt soll an einem konkreten Beispiel einer Untersuchung verdeutlicht werden, wie experimentelles Denken – d. h. die Suche nach dem Ausschluss möglicher alternativer Erklärungen – in einem Feldexperiment in der Schule möglich ist. Ausgewählt wurde ein Beispiel aus dem Leseunterricht, da dieses Beispiel besonders geeignet ist, verschiedene Probleme von Experimenten im Unterricht zu verdeutlichen, *die ebenfalls auf Untersuchungen im Sachunterricht zutreffen*. Es handelt sich um „A Quasi-Experimental Validation of Transactional Strategies Instruction With Low-Achieving Second-Grade Readers“ von Brown, Pressley, Van Meter und Schuder (1996).

#### *Theoretischer Hintergrund und Ausgangspunkt für das Experiment*

Viele Untersuchungen haben aufgezeigt, dass Lesenlernen durch die Vermittlung von Lernstrategien unterstützt werden kann. Dabei scheint die Vermittlung einer einzigen Strategie weniger erfolgversprechend, vielmehr kommt es auf die Koordination mehrerer verschiedener Strategien an. Die Autoren gehen auf der Basis einer Analyse verschiedener pädagogischer Programme in der Grundschule davon aus,

dass effektives Verstehen durch drei Merkmale gekennzeichnet ist, die sie jeweils „transaktional“ nennen.

- Leser werden ermutigt, die Bedeutung eines Textes durch Strategien zu erschließen, mit denen der Inhalt des Textes mit ihrem Vorwissen verbunden wird.
- Ein Großteil des Strategielernens findet in Lesegruppen statt, indem die Gruppe gemeinsam Strategien nutzt, um den Sinn von Texten zu erschließen. Der dadurch erschlossene Sinn wird als transaktional bezeichnet und unterscheidet sich (möglicherweise) von dem Sinn, den ein Individuum sich alleine erarbeitet.
- In den Lesegruppen werden bei dem Erschließen eines Textes mit Hilfe von Strategien die Handlungen von jedem Gruppenmitglied durch die Handlungen der anderen beeinflusst. Die Gruppenmitglieder bestimmen gemeinsam ihr Denken über den Text.

Unterricht, in dem diese drei Formen transaktionalen Handelns vorkommen, wird auch als transaktionaler Strategieunterricht (transactional strategies instruction) bezeichnet. In der Vergangenheit wurden verschiedene qualitative Untersuchungen zu diesem Unterricht durchgeführt (z.B. ethnographische Studien, qualitative Interviews und Beobachtungen des Unterrichts, Langzeitfallstudien), die aufschlussreiche Informationen zu diesem Unterricht lieferten. Was allerdings nach Ansicht der Autoren fehlte, waren formalisierte Vergleiche der Leseleistungen (im weiteren Sinne) von Schülern, die den transaktionalen Strategieunterricht erfahren haben, mit denen von solchen Schülern, die konventionell unterrichtet wurden. Dazu wurde eine quasi-experimentelle Untersuchung mit einem spezifischen Modell des transaktionalen Strategieunterrichts durchgeführt: dem Students Achieving Independent Learning (SAIL) Programm. In diesem umfangreichen und komplexen Programm lernen die Schüler an Texten, kommende Ereignisse vorherzusagen, ihre Erwartungen im Fortgang des Textes zu ändern, während des Lesens Fragen und Interpretationen zu erzeugen, Ideen zu visualisieren, regelmäßig Zusammenfassungen anzufertigen und auf die wichtigsten Aspekte des Textes zu achten. Diese Strategien werden im Unterricht auf verschiedenen Wegen vermittelt: durch direkte Unterweisung, Modellierung von strategischem Verhalten durch den Lehrer, Coaching und Scaffolding (für eine weitergehende Beschreibung des Programms vgl. Brown, Pressley, Van Meter & Schuder 1996).

Mit diesem Unterrichtsexperiment – das hier in zweiten Klassen durchgeführt wurde – wird noch einmal deutlich, dass empirische Forschung nicht aus isolierten Einzeluntersuchungen besteht. Die verschiedenen vorgängigen qualitativen Untersuchungen haben viele tiefgehende Ergebnisse aufgedeckt, für die Funktion eines zusammenfassenden Vergleichs der Effekte des SAIL Programms sind aber formalisiertere und auch längerfristige – hier über ein Schuljahr – Vergleichsuntersuchungen nötig, ohne die eine breitere Implementation dieses Unterrichts in der Schulpraxis nicht zu verantworten ist.

### *Die Auswahl der Lehrer in Experimental- und Kontrollgruppe*

In den meisten Untersuchungen ist es nicht möglich, Lehrer zufällig auf die Experimentalklassen und auf die Kontrollklassen aufzuteilen und sie dann in dem jeweils benötigten Unterricht zu trainieren. In dieser Untersuchung wird davon ausgegangen, dass erst mehrjährige Erfahrungen mit dem transaktionalen Strategieunterricht zu effektiven SAIL Lehrern führt. Deshalb wurden in einem Schuldistrikt erfahrene SAIL Lehrer für die Experimentalklassen ausgewählt (durchschnittlich 10,4 Jahre Unterrichtspraxis). Um zu einem fairen Vergleich des SAIL Unterrichts zu kommen, wurden im gleichen Schuldistrikt für die Kontrollklassen mit Hilfe der Schulleiter und anderer Spezialisten Lehrer ausgewählt, die dem regulären LeseCurriculum des Distrikts folgten und die als besonders fähige Lehrer für den Leseunterricht in zweiten Klassen galten (durchschnittlich 23,4 Jahre Unterrichtspraxis). Obwohl man sich nicht sicher sein kann, dass die Fähigkeiten der Lehrer nicht doch in Experimental- und Kontrollgruppe unterschiedlich waren oder dass Unterschiede in den Lernergebnissen nicht auf die unterschiedlich lange Unterrichtspraxis der Lehrer zurückzuführen sind, so gibt es durch diese Auswahl der Lehrer für Experimental- und Kontrollgruppe doch gute Gründe, solche alternativen Erklärungen für unterschiedliche Lernergebnisse auf Schülerseite auszuschließen.

### *Die Bildung von Experimental- und Kontrollgruppen*

Um die Vergleichbarkeit der Experimentalklassen – mit SAIL Unterricht – und der Kontrollklassen – mit Unterricht entsprechend des regulären Curriculums – vor Beginn des Experiments sicherzustellen, wurde folgendes Verfahren angewandt:

- Anhand von schuldemographischen Aspekten (z. B. ethnische Zusammensetzung, Anteile von Minoritätengruppen, Größe und Ort der Schule, Durchschnittsleistungen in einem standardisierten Test) wurden fünf parallele Paare von je einer Experimentalklasse und einer Kontrollklasse gebildet, d.h. Paare von zweiten Klassen, die sich in diesen demographischen Variablen nicht unterscheiden. Es wurden somit insgesamt zehn zweite Klassen für die Untersuchung ausgewählt, fünf Experimentalklassen und fünf Kontrollklassen.
- Für die Auswahl der teilnehmenden Schüler in den Klassen wurde jedes Paar von einer Experimentalklasse und einer Kontrollklasse gesondert betrachtet. Sechs Schüler in der Experimentalklasse und sechs Schüler in der Kontrollklasse wurden so ausgewählt, dass sie in einem Leseverständnistest (Stanford Achievement Test, Comprehension Subtest) parallele Leistungen zeigten. Allerdings konnte dieser Leseverständnistest erst im Herbst – einige Zeit nach Beginn des Schuljahres und damit nach Beginn der Untersuchung – durchgeführt werden, d.h. zu einem Zeitpunkt, zu dem in den Experimentalklassen die Lehrer bereits einige Elemente des SAIL Programms eingeführt hatten.

Insgesamt nahmen damit zehn zweite Klassen und 60 Schüler an der Untersuchung teil. Bei den Schülern handelte es sich dabei um Schüler, deren Kenntnisstand zu Beginn des zweiten Schuljahres unter dem für die Klassenstufe üblichen lag. Vergleichs- und Analyseebene sind die Klassen. Diese Auswahl von Klassen und Schülern kann als ein gutes Beispiel betrachtet werden, bei den üblichen Restriktionen bei Untersuchungen in Schulen dennoch zu Gruppen mit vergleichbaren Ausgangslagen zu kommen.

### *Überprüfung des Unterrichts in Experimental- und Kontrollklassen*

In Feldexperimenten ist man nie völlig sicher, ob sich der Unterricht in den Experimentalklassen von jenem in den Kontrollklassen in der angezielten Weise unterscheidet. Dies wurde in der vorliegenden Untersuchung umfangreich überprüft. Zunächst wurden zu Beginn des Experiments Orientierungen zum Leseunterricht bei Experimental- und Kontrollklassenlehrern erhoben. Die beiden Gruppen unterscheiden sich klar und in der erwarteten Richtung voneinander. In der zweiten Hälfte des Schuljahres wurden dann in allen zehn Klassen zwei Geschichten gelesen. Der Unterricht und die Interaktionen während des Vorlesens dieser Geschichten wurden auf Video aufgenommen und transkribiert. Vier Rater, die nicht wussten, aus welchen Klassen die Transkripte stammten, ordneten sehr treffsicher die Transkripte zu den Experimental- und den Kontrollklassen zu, so dass sich klare Unterschiede im Unterricht in den SAIL Klassen und den Kontrollklassen zeigten. Zwei Rater beurteilten weiterhin die Verwendung von Strategien in den transkribierten Unterrichtsstunden, wobei sich in den Experimentalklassen deutlich häufiger ein Strategieunterricht zeigte. Die Autoren diskutieren dabei, dass für die Überprüfung der Beurteilerübereinstimmung der beiden Rater diese eigentlich nicht wissen sollten, ob es sich um Transkripte aus den Experimental- oder Kontrollklassen handelte, und diese Rater auch dem angezielten Unterricht in den Experimentalklassen eher „entfernt“ bzw. „neutral“ gegenüberstehen sollten. Andererseits wird von qualitativen Forschern in Frage gestellt, ob solche „blinden“ Rater wirklich angemessen die Daten analysieren können. Als Kompromiss haben die Autoren deshalb nur einen „blinden“ Rater ausgewählt, der gewissermaßen den Interpretationen des „nicht-blinden“ Raters – hier einer der Forscher – Glaubwürdigkeit verleihen kann. Im Übrigen zeigte sich bei der Analyse der Unterrichtsstunden, dass die Unterrichtseinheiten, in denen die beiden Geschichten bearbeitet wurden, in den Experimentalklassen länger ausfielen als in den Kontrollklassen. Die Autoren führen dies darauf zurück, dass das Aushandeln von Interpretationen, das Erklären und Vormachen von Strategien und lautes Denken eben zeitintensive Aktivitäten sind.

Diese sehr sorgfältige Überprüfung des tatsächlich realisierten Unterrichts in den Experimental- und den Kontrollklassen stellt die angemessene Operationalisierung

der theoretischen Hypothese sicher – ein Punkt der für jedes Unterrichtsexperiment bedeutsam ist.

### *Verbindung von qualitativen und quantitativen Erhebungsverfahren*

Eine Hauptintention des Experiments war, die Auswirkungen des SAIL Unterrichts auf einen allgemeinen Leistungstest nachzuweisen – speziell im Hinblick auf eine breitere Implementation des SAIL Programms. Deshalb wurden am Ende des Experiments zwei standardisierte Tests (Subtests Reading Comprehension und Word Skills des Stanford Achievement Tests) eingesetzt. Gleichzeitig wurden verschiedene qualitative Verfahren verwendet: Mit den Schülern wurde zum einen individuell ein halbstrukturiertes Strategieinterview durchgeführt. Zum anderen wurden im Anschluss an die Unterrichtsstunden, in denen die beiden Geschichten präsentiert wurden, die Schüler individuell aufgefordert, die Geschichten zu erinnern und selber zu erzählen. Schließlich wurde mit den Schülern individuell eine Geschichte gelesen, wobei die Schüler an vier Stellen unterbrochen und gebeten wurden, laut über die Geschichte nachzudenken. Das Ziel des Einsatzes dieser qualitativen Verfahren bestand darin, über die standardisierten Verfahren hinaus tiefergreifende Informationen über die Strategienutzung der Schüler zu erhalten. In der Tat zeigten die qualitativen Verfahren große Unterschiede zwischen Experimental- und Kontrollgruppe in der Breite und im Umfang der verwendeten Strategien auf. Der Einsatz von quantitativen und qualitativen Verfahren innerhalb des experimentellen Designs wird hier nicht als Alternative, sondern als gegenseitige Ergänzung betrachtet.

### **Auswertung**

Das Design dieser Untersuchung erlaubte für die Hauptfragestellungen eine relativ einfache statistische Auswertung. Es wurden jeweils die Mittelwerte der sechs Schüler innerhalb der fünf Experimentalklassen und der fünf Kontrollklassen gebildet. Über t-Tests für abhängige Stichproben wurde dann z.B. geprüft, ob sich in den interessierenden abhängigen Variablen die Mittelwerte zwischen Experimentalgruppe (d. h. der Mittelwert über die Mittelwerte der fünf Experimentalklassen) und Kontrollgruppe (d. h. der Mittelwert über die Mittelwerte der fünf Kontrollklassen) unterscheiden.<sup>8</sup> Es zeigten sich hier signifikante Unterschiede, die die Hypothese der Untersuchung bestätigte, nämlich dass der Unterricht in den SAIL Klassen im Vergleich zum Unterricht in den Kontrollklassen zu besseren Schülerleistungen in den standardisierten Tests sowie in der Strategienutzung führt.

Insgesamt kann diese Untersuchung von Brown, Pressley, Van Meter und Schuder

von 1996 als ein gutes Beispiel für eine hypothesentestende Untersuchung unter den Bedingungen des Schulalltags betrachtet werden, ein Beispiel, das aufgrund seiner Anlage auch für künftige Forschung im Sachunterricht hilfreich sein könnte.

## Anmerkungen

- <sup>1</sup> Dabei muss der experimentelle Stimulus nicht zwingend durch den Experimentator selbst explizit eingeführt werden. Es kann auch sein, dass im natürlichen Geschehensablauf eine Situation entsteht, die zu unterschiedlichen Ausprägungen in der unabhängigen Variablen führt und bei der mögliche Störfaktoren (s.u.) kontrolliert werden können (Naturexperiment).
- <sup>2</sup> Auf die Frage der Operationalisierung der abhängigen Variablen soll hier nicht weiter eingegangen werden.
- <sup>3</sup> Es wird auch unterschieden zwischen zufälligen bzw. unsystematischen Fehlern, wie sie z.B. durch eine Zufallsaufteilung auf Experimental- und Kontrollgruppe entstehen, internen Fehlern (die sich auf die interne Validität auswirken) und externen Fehlern (die sich auf die externe Validität auswirken).
- <sup>4</sup> Streng logisch kann nicht von Ergebnissen, die an einer Stichprobe gewonnen wurden, auf die Population geschlossen werden. Wohl aber kann bei extern validen Experimenten – die evtl. auch noch bei gewissen Modifikationen am Experiment repliziert wurden – die Berechtigung erhöht werden, eine Bewährung des gefundenen Zusammenhangs auch außerhalb des Experiments anzunehmen.
- <sup>5</sup> Die im Weiteren verwendeten Zeichen haben folgende Bedeutung: O = Beobachtung, Tests; X = experimentelle Behandlung; R = Randomisierung, Zufallsaufteilung; M = Matching, Parallelisierung; EG = Experimentalgruppe; KG = Kontrollgruppe.
- <sup>6</sup> Es ist auch eine Verbindung von Parallelisierung und Zufallsaufteilung möglich, wenn nach bekannten möglichen Störgrößen parallele Paare gebildet werden, die dann aber zufällig auf Experimental- und Kontrollgruppe verteilt werden. Solche kombinierten Auswahlen sind reinen Zufallsaufteilungen überlegen, die wieder besser sind als einfache Parallelisierungen.
- <sup>7</sup> Dieser Störfaktor bezieht sich nicht auf die interne Validität, denn durch die Randomisierung wirken sich Auswahleffekte in Experimental- und Kontrollgruppe gleichermaßen aus.
- <sup>8</sup> Da gewissermaßen über die Anlage der Untersuchung Zwillingspaare aus Experimental- und Kontrollklassen gebildet wurden, wird hier von abhängigen Stichproben ausgegangen.

## Literatur

- Brown, R., Pressley, M., Van Meter, P. & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology*, 88, 18-37.
- Henningsen, J. (1964). *Test, Experiment, Befragung: ein kritisches Plädoyer*. 2. Aufl. Essen: Neue Deutsche Schule.
- Roberts, K. H. & Rost, D. H. (1974). *Analyse und Bewertung empirischer Untersuchungen*. Weinheim: Beltz.
- Roßbach, H. G. (1996). Lage und Perspektiven der empirischen Grundschulforschung. *Empirische Pädagogik*, 10, 167-191.
- Schwarz, E. (1973). Experimentelle und quasi-experimentelle Anordnungen in der Unterrichtsforschung. Deutsche Bearbeitung des Kapitels 5: Experimental and quasi-experimental designs for

- research on teaching von Donald T. Campbell und Julia C. Stanley. In H. Feger, C. Holzkamp, E. Schwarz & E. v. Trotsenburg (Hrsg.), Strategien der Unterrichtsforschung. Teilausgabe des Handbuchs der Unterrichtsforschung (S. 99-193). Weinheim: Beltz.
- Saldern, M. v. (1995). Zum Verhältnis von qualitativen und quantitativen Methoden. In E. König & P. Zedler (Hrsg.), Bilanz qualitativer Forschung. Band I: Grundlagen qualitativer Forschung (S. 331-371). Weinheim: Deutscher Studien Verlag.
- Shulman, L. S. (1988). Disciplines of inquiry in education: An overview. In R.M. Jaeger (Ed.), Complementary methods for research in education (S. 3-17). Washington: American Educational Research Association.
- Wellenreuther, M. (2000). Quantitative Forschungsmethoden in der Erziehungswissenschaft. Eine Einführung. Weinheim: Juventa.
- Zimmermann, E. (1972). Das Experiment in den Sozialwissenschaften. Stuttgart: Teubner.



## Qualitative Analyseansätze in der Lehr-Lern-Forschung

### 1. Einleitung: Ein Trend zur qualitativen Forschung?

Qualitativ orientierte Ansätze spielen im sozialwissenschaftlichen Methodenspektrum heute eher eine Nebenrolle. Ein Blick in Handbücher und methodologische Einführungswerke zeigt dies sehr deutlich.

So finden sich in der Enzyklopädie der Psychologie im Band zu den methodologischen Grundlagen (Herrmann & Tack 1994) unter 16 Beiträgen nur zwei qualitativ orientierte Beiträge (zu den Themen Interpretation und Induktion), im Band zur Datenerhebung nur die Themen Exploration und Biografie-forschung, in den Bänden zu Messen und Testen und zur Strukturierung und Reduzierung von Daten jeweils kein einziger qualitativ orientierter Beitrag. Das Handwörterbuch der Pädagogischen Psychologie (Rost 1998) führt unter 98 Beiträgen keinen qualitativ orientierten Beitrag; nicht einmal im Sachregister ist qualitative Forschung aufgeführt. Auch das Handwörterbuch Psychologie (Asanger & Weninger 1999) enthält kein qualitativ orientiertes Stichwort (unter 159 Stichwörtern); das Sachregister verweist immerhin auf die Begriffe „qualitativ“ und „offen“ im Abschnitt zur Wissenschaftstheorie, dort werden die Stichworte allerdings nur kurz und in Anführungszeichen erwähnt. Das einschlägige Methodenwerk von Bortz & Döring (1995) entfaltet ein variablenorientiertes, hypothesentestendes, statistisch auswertendes Vorgehen als Standard, enthält aber immerhin ein Kapitel zu qualitativen Methoden (Qualitative Befragung, Beobachtung, Grounded Theory, Qualitative Inhaltsanalyse, Feldforschung, Aktionsforschung, Frauenforschung, Biografie-forschung), allerdings in einer sehr dichotomisierenden Art und Weise. Betrachtet man den Bereich der Soziologie und Erziehungswissenschaft, so ändert sich das Bild ein wenig. Atteslander (2000), das vielleicht meistverkaufte sozialwissenschaftliche Methodenbuch (über 100.000 seit der Erstauflage 1968) geht in den neueren Auflagen auf qualitative Forschung ein; teilnehmende Beobachtung, offene Befragung und Qualitative Inhaltsanalyse werden besprochen, allerdings fehlen weitere qualitativ orientierte Auswertungsansätze. Die Enzyklopädie Erziehungswissenschaft dagegen geht in ihrem Methodenband (Haft & Kordes 1984) ausführlich auf qualitative Forschung ein (Aktionsforschung, Objektive Hermeneu-

tik, Exploration, Introspektion, verstehende Methoden, biografische Methoden, Interpretation, Einzelfallanalyse, narratives Interview etc.).

Die für Lehr-Lernforschung zentrale Zeitschrift „Unterrichtswissenschaft“ widmete der qualitativen Forschung zweimal einen eigenen Themenschwerpunkt (Unterrichtswissenschaft 1984, Heft 3 und 1999, Heft 4). Die Lehr-Lern-Forschung zeigt sich besonders dann qualitativen Ansätzen offen, wenn sie sich stärker auf erziehungswissenschaftliche Traditionen als auf psychologische Traditionen bezieht (vgl. Mayring 1999b). Denn in der Erziehungswissenschaft haben sich, vielleicht wegen der bis heute hier stärker vertretenen geisteswissenschaftlichen Traditionen mit hermeneutisch-interpretativen Vorgehensweisen, qualitative Ansätze stark entwickelt (vgl. zum Überblick König & Zedler 1995 oder für den englischsprachigen Raum Jaeger 1988; Eisner & Peshkin 1990). Dabei spielt die einzelfallanalytische Vorgehensweise eine besondere Rolle (Baacke & Schulze 1993; Shulman 1992). Die moderne fachdidaktische Unterrichtsforschung hat an verschiedenen Stellen solche qualitativen Ansätze aufgegriffen (vgl. Punkt 4), gerade auch in der Forschung zum Sachunterricht (vgl. Richter 2000).

Kann man in der heutigen Lehr-Lern-Forschung von einem Trend zur qualitativen Forschung sprechen? Ich denke, Ansatzpunkte in diese Richtung gibt es, zumindest zu einer größeren methodischen Offenheit. Allerdings wirkt hier eine übertriebene Frontstellung, ein oft formuliertes entweder – oder erschwerend.

## 2. Von der Unsinnigkeit einer Qualitativ-Quantitativ-Dichotomisierung

Bei den Versuchen einer Definition qualitativer Analyse stößt man immer wieder auf strikte Dichotomisierungen zur quantitativen Analyse. „Hinter ... forschungspraktischen Differenzen stehen bei manchen Autoren weitreichende Diskrepanzen in der Auffassung von Wissenschaft, wenn nicht gar im Menschenbild“ (Bortz & Döring 1995, S. 274). Am eindrucksvollsten gerät die Gegenüberstellung quantitativ – qualitativ in der Darstellung von Lamnek (1988).

Fast alle diese Gegenüberstellungen (vgl. Tab. 1) erweisen sich beim genaueren Hinsehen auf die Praxis konkreter Forschungsprojekte als schwer aufrechtzuerhalten. So steckt hinter der Dichotomie Erklären vs. Verstehen zwar eine alte, auf Wilhelm Dilthey zurückgehende wissenschaftstheoretische Debatte (vgl. v. Wright 1974; Riedel 1978). Allerdings beanspruchen qualitativ orientierte Ansätze oft auch, Kausalerklärungen oder Überprüfungen in Bezug auf Allgemeinaussagen (nomothetisch, theorieprüfend) vornehmen zu können (besonders mit Hilfe von Einzelfallanalysen, aber auch z. B. im Rahmen von Aktionsforschungsprojekten). Andererseits gibt es auch idiographische, einzelfallanalytische Ansätze innerhalb eines quantitativ orientierten Forschungsparadigmas (vgl. z. B. das zeitreihenanalytische Vorgehen, Schmitz 1996). Im Rahmen der Qualitativen Inhaltsanalyse

Quantitative Sozialforschung	Qualitative Sozialforschung
Erklären	Verstehen
nomothetisch	idiographisch
theorieprüfend	theorieentwickelnd
deduktiv	induktiv
objektiv	subjektiv
ätiologisch	interpretativ
ahistorisch	historisierend
geschlossen	offen
Prädetermination des Forschers	Relevanzsysteme der Betroffenen
Distanz	Identifikation
statisch	dynamisch-prozesshaft
starrs Vorgehen	flexibles Vorgehen
partikularistisch	holistisch
Zufallsstichprobe	theoretical sampling
Datennähe	Datenferne
Unterschiede	Gemeinsamkeiten
reduktive Datenanalyse	explikative Datenanalyse
hohes Messniveau	niedriges Messniveau

Tab. 1: Gegenüberstellung quantitativer und qualitativer Sozialforschung (Lamnek 1988, S. 228)

(vgl. Mayring 2000) werden induktive (Kategorienentwicklung) und deduktive (Kategorienanwendung) Verfahrensweisen beschrieben. Der Anspruch objektiver Wissenschaft (vgl. das fünfte Begriffspaar bei Lamnek) wird heute in quantitativ (orientierter Methodologie nicht mehr unhinterfragt aufrechterhalten. So ließe sich die Kritik der weiteren Begriffspaare in Tab. 1 problemlos fortsetzen. Dadurch wird klar, dass eine strikte Dichotomie qualitativ versus quantitativ heute nicht mehr zu vertreten ist. Auf die verschiedenen Verbindungslinien, Mischmöglichkeiten und Integrationsmodelle ist immer wieder hingewiesen worden (vgl. Buer 1984; Saldern 1992; Mayring 2001). Dabei ist an anderer Stelle ausgearbeitet worden (Mayring 2001), dass eine Integration qualitativer und quantitativer Methoden

- bereits auf technischer Ebene durch die Verwendung computergestützter qualitativer Programme (vgl. Weitzman & Miles 1995) nahegelegt wird,
- durch das Arbeiten mit Kategorien oder Codes (im Rahmen von Qualitativer Inhaltsanalyse oder Grounded Theory) vorbereitet wird,
- durch eine schrittweise Verbreiterung der Fallbasis in einzelfallorientierten Forschungsplänen umgesetzt werden kann,
- durch Verschränkung verschiedener qualitativer und quantitativer Analyse-schritte in einem übergeordneten Forschungsdesign bewerkstelligt werden kann und schließlich

- durch ein gemeinsames Ablaufmodell empirischer Forschung für qualitativ und quantitativ orientierte Projekte am besten veranschaulicht werden kann.

Trotz dieser Verbindungslinien ist die Frage berechtigt, was nun eigentlich qualitativ orientierte Forschung auszeichnet. Wir gehen davon aus, dass keine strikte Trennungslinie zu ziehen ist, dass aber qualitativ orientierte Projekte einen spezifischen Schwerpunkt haben, der sich in fünf Aspekten ausdrückt (vgl. auch Mayring 1999a).

### *Auch deskriptive Herangehensweise an den Gegenstand*

Qualitativ orientierte Forschungsprojekte werden sich ihrem Gegenstand immer erst offen beschreibend nähern. Der deskriptive Zugang zum Untersuchungsfeld wird heute in quantitativer Forschung oft vernachlässigt, und hier liegt eine Stärke qualitativer Forschung. Das schließt allerdings nicht aus, dass im weiteren Verlauf auch Theoriekonzepte und Variableneinengungen an das Material herangetragen werden.

### *Interpretative Behandlung des erhobenen Materials*

Das in der Regel relativ offen erhobene Material kann in qualitativ orientierten Forschungsprojekten nicht automatisch wie Testrohwerte oder Fragebogenkreuzchen ausgewertet werden. Das Material (z.B. Interviewtranskript, Beobachtungsprotokoll) spricht nicht für sich selbst, es muss erst gedeutet, interpretiert werden. Vor allem bei sprachlichem Material wird der subjektiv gemeinte Sinn erschlossen. Aber auch hier wieder kann gesagt werden, dass ein solcher interpretativer Schritt nicht eine quantitative Weiterverarbeitung der Interpretationen ausschließen muss.

### *Einzelfallbezogenheit*

Die Fallanalyse hat in qualitativer Forschung immer eine besondere Rolle gespielt, weil am einzelnen Fall die Komplexität von Zusammenhängen direkter erforschbar ist (Stake 1995). Allerdings besteht die Schwierigkeit hier in der Verallgemeinerbarkeit der Ergebnisse. Mit Einzelfallbezogenheit ist dabei gemeint, dass auch im Rahmen großangelegter Studien Fallanalysen eingebaut werden können, die die Variablenzusammenhänge genauer erklären lassen. Der Rückgriff auf den Fall erlaubt dabei auch immer eine ganzheitlichere, subjektbezogener Sicht.

### *Alltagsnähe*

Qualitative Forschung hat sich oft durch eine Kritik am Laborexperiment begründet. Zwar seien hier Variablenzusammenhänge unter kontrollierten Standardbedingungen beobachtbar, allerdings stelle sich die Frage, ob die Versuchspersonen im Alltag ähnlich wie im Labor reagieren. Vor allem in der qualitativen Feldforschung ist es wichtig, in die alltägliche Lebenswelt der Beforschten einzudringen

(vgl. z. B. Girtler 1984). Auch hier wieder haben wir kein entweder – oder vor uns, sondern den Anspruch, mit qualitativ orientierter Forschung so nah wie möglich am Alltag der Beforschten anzuknüpfen.

### *Argumentative Verallgemeinerung der Ergebnisse*

Dies ist ein fundamentaler Unterschied zwischen quantitativer und qualitativer Forschung: Dort kann man oft automatisch von der Generalisierbarkeit der Ergebnisse ausgehen, da repräsentative Stichproben und standardisierte Instrumente zum Einsatz kamen. In qualitativer Forschung dagegen muss die Verallgemeinerbarkeit erst erwiesen und argumentativ abgesichert werden. Es muss begründet werden, welche Ergebnisaspekte unter welchen Bedingungen bis zu welchem Grade generalisiert werden können.

In diesen fünf Punkten zeigt sich die Spezifität eines qualitativ orientierten Herangehens, ohne damit Überschneidungen und Verbindungen zu quantitativer Analyse ausschließen zu wollen.

## **3. Beispiel eines qualitativ orientierten Ansatzes in der Lehr-Lern-Forschung**

An anderer Stelle habe ich bereits ausführliche Überblicke über qualitativ orientierte Ansätze in der Lehr-Lern-Forschung gegeben (Mayring 1996; 1999b). Wir finden solche Studien vor allem im Bereich pädagogischer Fallanalysen (z. B. Schülerportraits, Lehrerbiografien), ethnomethodologischer Schulfeldforschung, schulbezogener Praxisforschung, der qualitativen Unterrichtsbeobachtung, der Rekonstruktion schulbezogener Subjektiver Theorien und qualitativ orientierter Evaluationsstudien im Bereich Schule.

Hier soll nun ein Beispiel aus der eigenen aktuellen Forschungsarbeit kurz vorgestellt werden (vgl. Mayring, Rhöneck & Melenk 1997; Bleicher, Fix, Fuß et al. 2001; Laukenmann, Bleicher, Fuß, Gläser-Zikuda, Mayring & Rhöneck 2000; Gläser-Zikuda 2001). Dabei handelt es sich um ein Projekt mit einer quantitativen (652 SchülerInnen der Klassenstufe 8) und einer qualitativen Teilstudie (24 Achtklässler), über die nun berichtet werden soll.

### **3.1 Theorie und Fragestellung**

Neuere Untersuchungen der Lehr-Lern-Forschung betonen, dass Lernen nicht als rein kognitiver Prozess aufzufassen ist, sondern dass auch emotionalen Variablen wie Flow, Interesse, Lernfreude, schulischem Wohlbefinden, Langeweile und Angst eine entscheidende Rolle zukommt (vgl. Möller & Köller 1996; Jerusalem & Pekrun 1999; Hänze 1998). Die bisherigen Studien auf diesem Gebiet thematisieren allerdings wenig die fachspezifischen Aspekte des Einflusses von Emotio-

nen auf Lernen und sie analysieren wenig die Interaktion von Kognitionen und Emotionen, auch im Prozessverlauf. Deshalb hat sich die qualitative Teilstudie folgende Fragestellungen gesetzt:

- Lassen sich Lernemotionen mit qualitativen Erhebungsinstrumenten vergleichbar erheben?
- Lassen sich Zusammenhänge emotionaler Faktoren und Lernen spezifisch nachweisen?
- Sind dabei Fachunterschiede (Physik versus Deutsch) erkennbar?

### 3.2 Methode

Über Schulnoten und Lehrerbefragung wurden 12 leistungsstarke und 12 leistungsschwächere Schülerinnen und Schüler der achten Jahrgangsstufe ausgewählt, mit offenem Interview befragt und mit einem halb-strukturierten Lern- und Emotionstagebuch über sechs Wochen hinweg begleitet. Dieses Material wurde mittels Qualitativer Inhaltsanalyse (Mayring 2000) ausgewertet. Für die Kategorienentwicklung und Kodierung wurde eine formative und summative Reliabilitätsüberprüfung vorgenommen, die sehr zufriedenstellende Interkoderreliabilitäten zwischen .72 und .97 ergab (vgl. Gläser-Zikuda 2001).

### 3.3 Ergebnisse

Aus den Lern- und Emotionstagebüchern wurden induktiv Kategorien entwickelt und in einem zweiten Schritt zu Hauptkategorien zusammengefasst. Auf die zentrale Variable der Freude ergaben sich folgende Ergebnisse am Beispiel des Physikunterrichts (Elektrizitätslehre).

Wir sehen, dass die häufigste Kategorie die Freude über das Beherrschen des Lernstoffes ist. Freude über den Unterrichtsverlauf und Freude über die eigenen Leistungen tauchten ebenfalls relativ häufig in den Tagebüchern auf. Solche Kategorien wurden zur Hauptkategorie „Freude am Lernprozess“ zusammengefasst und der Hauptkategorie „Freude am Inhalt“ gegenübergestellt. Freude am Lernprozess taucht dabei in den Tagebüchern viel häufiger auf.

In einem zweiten Auswertungsgang wurden die Lernemotionen einer skalierenden Inhaltsanalyse (viel – mittel – wenig Freude) unterzogen und hier Gruppenunterschiede zwischen leistungsstarken und leistungsschwachen SchülerInnen untersucht.

Schule				Zu Hause		
Kat. Nr.	Keine Freude	Nen. insg.	N Schü	Keine Freude	Nen. insg.	N Schü
1	Nichts macht Freude	5	3	Nichts machte Freude	2	2
Kat. Nr.	Freude am Lernprozess	Nen. insg.	N Schü	Freude am Lernprozess	Nen. insg.	N Schü
1	Freude über das Beherrschen des Lernstoffs	14	6	Freude über das Beherrschen des Lernstoffs	19	8
2	Freude über den Unterrichtsverlauf	9	6	Freude über das Ende des Lernens	4	2
3	Freude über positive Leistungsrückmeldung	4	2	Freude über erledigte Hausaufgaben	3	2
4	Freude über Klassenarbeit/ Test	3	3	Freude über geringen Arbeitsaufwand	1	1
5	Freude über geringen Arbeitsaufwand	2	2		27	13
6	Freude über schnelles Arbeitsende	2	2			
		34	21			
Kat. Nr.	Freude am Physikinhalte	Nen. insg.	N Schü	Freude am Physikinhalte	Nen. insg.	N Schü
1	Freude an Elektrizitätslehre	3	2	Freude an Elektrizitätslehre	8	2

Tab. 2: Induktiv entwickelte Kategorien zu Freude aus Lern- und Emotionstagebüchern; 12 Emotionstagebücher, N = 79 Nennungen, Mehrfachnennungen möglich (Gläser-Zikuda 2000, S. 161; Laukenmann et al. 2000, S. 148)

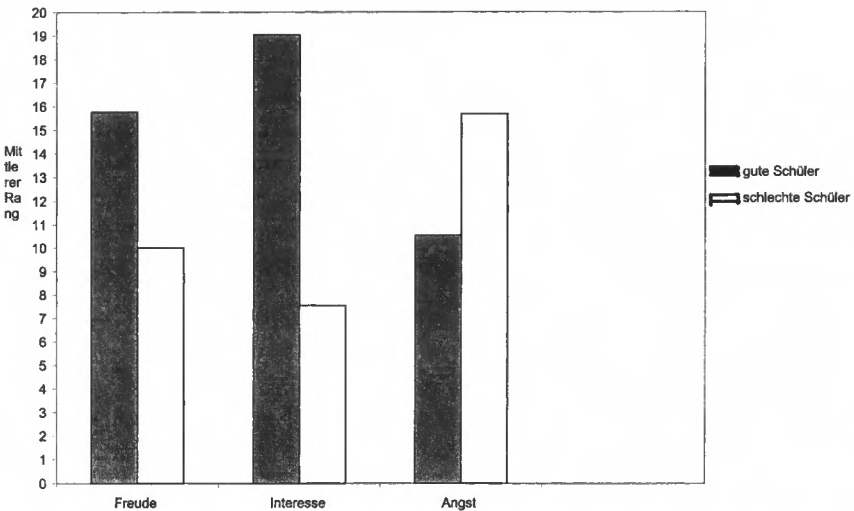


Abb. 1: Lernfreude, Interesse und Angst bei leistungsstarken vs. leistungsschwachen SchülerInnen (N = 25) (vgl. Gläser-Zikuda 2000; S. 210)

Hier zeigt sich ein eindeutiger Zusammenhang, der auch signifikant wird (Mann-Whitney-U-Test bei Freude  $p=.022^*$ , bei Interesse  $p=.000^{***}$ ). Bessere Lerner haben mehr Freude am Lernen als schlechtere.

### 3.4 Diskussion der Ergebnisse

Während in der quantitativen Teilstichprobe die Emotionsstärke über einen Skalenwert (Items wie „Die Stunde hat mir Freude gemacht“, „Der Unterricht hat mir Spaß gemacht“, gefolgt von einer vierstufigen Skala) bestimmt wird, kann die qualitative Teilstudie eruieren, woran die Schülerinnen und Schüler konkret Freude in der Schule haben. Überraschend war, dass nicht die Lerninhalte, sondern der Spaß am Lernprozess im Vordergrund standen. Wir schließen daraus, dass es in der Schule nicht darauf ankommt, nur gute Stimmung zu verbreiten, sondern die Freude an den Lernprozess zu koppeln.

Die Auswertung der deduktiven Kategorien (Grad der Lernfreude) ermöglichte eine Analyse der Zusammenhänge zur Leistung, die auch in der quantitativen Studie (durch Korrelations- und LISREL-Berechnungen) angestellt wurde. Die Zusammenhänge in der qualitativen Studie sind eindeutig stärker. In der quantitativen ergab sich ein signifikanter Zusammenhang nur im Fach Physik ( $r = .17^{**}$ ), nicht aber für das Fach Deutsch. Wir vermuten, dass die Zusammenhänge von Lernfreude zu Lernleistung stärker werden, wenn wir genauer (mittels qualitativer Analyse) auf die Lernprozesse sehen. Die Haupthypothese des Projekts (Positiver Effekt von Lernfreude auf Lernergebnisse) wird hier deutlich unterstützt.

## 4. Beispiele qualitativ orientierter Ansätze aus der Sachunterrichtsforschung

Auch im engeren Kreis der Forschungen zur Didaktik des Sachunterrichtes lassen sich mittlerweile eine Reihe von interessanten und vielversprechenden qualitativ orientierten Ansätzen auffinden. Möller (2000) führt hier bereits frühe deskriptiv orientierte Studien zu Schülervorstellungen naturwissenschaftlicher Phänomene (J. Piaget; M. Wagenschein) an. Ich möchte hier einige aktuelle Beispiele nennen:

- Gartz (1997) hat mit einem an der objektiven Hermeneutik (Oevermann) orientierten Vorgehen eine Unterrichtseröffnungsphase analysiert. Die Feinanalyse ergab, dass das qualitative Vorgehen verdeckte Muster, Elemente eines „heimlichen Lehrplans“ aufdecken kann.
- Hempel (1997) hat Lebensentwürfe (Berufswünsche, Lebensvorstellungen) von Grundschulkindern durch freie Beschreibung (Phantasiereisen, Wunschgeschichten) und teil-offene Leitfadeninterviews erhoben und gezeigt, dass eine Didaktik des Sachunterrichts an solchen Vorstellungen anknüpfen muss (Kind-



orientierung). Gerade bei Grundschulkindern kommt man hier mit standardisierten Fragebogentechniken nicht sehr weit.

- Hartinger (1997) hat in seiner Studie zur Interessenförderung von Grundschulkindern im Sachunterricht (Beispiel „Leben am Gewässer“) neben einem geschlossenen Fragebogen auch fokussierte Interviews mit 12 Schülern geführt und mittels qualitativer Inhaltsanalyse ausgewertet. Die Ergebnisse können die quantitativen Resultate erläutern, differenzieren und vertiefen.
- Möller (2000) hat eine Sachunterrichtseinheit („Wie kommt es, dass ein Eisen-schiff schwimmt?“) durch qualitative, halb-strukturierte Interviews mit elf SchülerInnen (jeweils vor und nach der Lerneinheit) und eine individuelle Lernweganalyse aufgrund der audiovisuellen Aufzeichnungen des Lernprozesses in Gruppen untersucht. Es konnten deutliche und bewusste Konzeptwechsel, z. B. von der Vorstellung „etwas Schweres geht unter“ zur Vorstellung „Holz schwimmt (fast) immer“ herausgearbeitet werden. Die Vorstellungen wurden differenzierter und hierarchischer. Solche qualitativen Veränderungen von impliziten und expliziten Vorstellungsmustern sind mittels rein quantitativer Untersuchungsanlagen nur schwer nachweisbar.
- Rheinoffer (2000) hat in einer explorativen Studie offene, halb-strukturierte Interviews mit Grundschullehrerinnen aus Großstadt-, Kleinstadt- und Land-schulen zum Stellenwert des Sachunterrichts durchgeführt. In der Auswertung des transkribierten Materials findet er mittels qualitativer Inhaltsanalyse unterschiedliche Bewertungsmuster und Begründungen, die dem Sachunterricht einen Vorrang, Gleichrang oder Nachrang gegenüber den anderen Grundschul-fächern geben.
- Besonders reichhaltig für qualitativ orientierte Sachunterrichtsforschung stellt sich der Sammelband von Richter (2000) dar. Hier wird eine transkribierte Un-terrichtsstunde („Ritter- und Bauernkinder im Vergleich mit heutigen Mädchen und Jungen“) aus den unterschiedlichsten qualitativ orientierten Analyseper-spektiven ausgewertet, wobei eher explorative Verfahren und mehr fragestel-lungsgeleitete Ansätze, auf Mikrostrukturen des Unterrichts wie auf Makro-strukturen der didaktischen Planung bezogene Auswertungen, historisch-in-terpretative, verstehende Analysen, Deutungsmusteranalysen und Diskursana-lysen eingesetzt werden.

Wir sehen also eine Vielzahl an qualitativ orientierten Ansätzen in moderner Sachunterrichtsforschung. Ein Problem bleibt dabei jedoch, dass eine explizite Einschätzung der methodischen Stringenz der Verfahrensweisen sowie eine ge-nauere Bewertung der erzielten Resultate, auch im Vergleich mit quantitativen Vorgehensweisen, oft schwer fällt. Angaben zu Objektivität, Reliabilität oder Vali-dität fehlen meistens. Deshalb soll abschließend auf Standards qualitativer For-schung eingegangen werden.

## 5. Methodische Standards qualitativ orientierter Forschung

Es ist oft betont worden, dass die klassischen Gütekriterien wie Objektivität, Reliabilität und Validität, aus der Testtheorie stammend, auf qualitative Forschung nicht völlig übertragbar sind (vgl. Mayring 1999a; Steinke 1999). Objektivität als Forderung nach völliger Unabhängigkeit der Untersuchung von individuellen Einflüssen der Untersucher wird von vielen Vertretern qualitativer Forschung nicht als Ziel formuliert. Im Gegenteil wird von einem dynamischen Verhältnis von Forschern und Beforschten gesprochen, von einer Interaktion, in der beide Beteiligten sich während des Forschungsprozesses verändern können. Der hermeneutische Zirkel, der von einer schrittweisen Veränderung des Vorverständnisses der Forscher ausgeht, wäre ein Beispiel für ein solches Konzept. Reliabilität als Gütekriterium wird dann fragwürdig, wenn Konzepte wie Retest-Reliabilität (erneute Durchführung der Studie mit der Hoffnung auf gleiche Ergebnisse) oder Paralleltest-Reliabilität oder interne Konsistenz zugrunde gelegt werden. Validität geht oft von einer deduktiven Methodologie aus oder sucht nach Außenkriterien, die bei qualitativer Forschung meist fehlen. Der allgemeine Anspruch auf eine genaue und auch gültige Forschung wird aber in der Regel auch von qualitativer Forschung aufrecht erhalten.

Als Konsequenz werden oft spezielle, für qualitativ orientierte Forschung geeignete Gütekriterien vorgeschlagen. Ich habe hier (Mayring 1999a) die Kriterien der Verfahrensdokumentation, der Argumentativen Interpretationsabsicherung, der Regelgeleitetheit, der Nähe zum Gegenstand, der kommunikativen Validierung und der Triangulation vorgeschlagen, wobei die ersten drei Kriterien eher in Richtung Reliabilität weisen, die letzten drei eher in Richtung Validität. Steinke (1999) hat in einer neueren Untersuchung zu Kriterien qualitativer Forschung folgende Vorschläge gemacht:

- Intersubjektive Nachvollziehbarkeit;
- Indikation des Forschungsprozesses und der Bewertungskriterien;
- Empirische Verankerung der Theoriebildung und -prüfung;
- Limitation (Bestimmung des Geltungsbereiches);
- Reflektierte Subjektivität;
- Kohärenz des theoretischen Hintergrunds;
- Relevanz von Fragestellung und Ergebnissen.

Diese Kriterien decken sich zum großen Teil mit den oben angeführten, sind aber allgemeiner auf den Forschungsablauf und seine Begründung bezogen. Sie verweisen damit darauf, dass qualitative Forschung sich im Wechselspiel zwischen Theorie und Empirie besonders ausweisen muss.

In jedem Falle sollten sich qualitativ orientierte Ansätze, die heute immer noch oft auf Skepsis treffen, hier besonders auszeichnen und explizit auf eine Diskussion der Forschungsabläufe und Ergebnisse anhand von klaren Gütekriterien eingehen.

## Literatur

- Asanger, R. & Weninger, G. (Hrsg.) (1999). *Handwörterbuch Psychologie*. Weinheim: Psychologie Verlags Union.
- Atteslander, P. (2000). *Methoden empirischer Sozialforschung*. 9. Aufl. Berlin: de Gruyter.
- Baacke, D. & Schulze, Th. (Hrsg.) (1993). *Aus Geschichten lernen. Zur Einübung pädagogischen Verstehens*. Weinheim: Juventa.
- Bleicher, M., Fix, M., Fuß, S., Gläser-Zikuda, M., Laukenmann, M., Mayring, Ph., Melenk, H. & v. Rhöneck, Ch. (2001). Einfluss emotionaler Faktoren auf das Lernen in den Fächern Physik und Deutsch – erste Ergebnisse aus dem Forschungsprojekt. In C. Finkbeiner & G.W. Schnaitmann (Hrsg.), *Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik* (S. 518-546). Donauwörth: Auer.
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation*. 2. Aufl. Berlin: Springer.
- Buer, J. v. (1984). ‚Quantitative‘ oder ‚qualitative‘ Unterrichtsbeobachtung? – Eine falsche Alternative. *Unterrichtswissenschaft*, 12, 252-267.
- Eisner, E.W. & Peshkin, A. (Eds.) (1990). *Qualitative inquiry in education*. New York: Teacher College Press.
- Gartz, D. (1997). Qualitative Forschungsmethoden für die Sachunterrichtsdidaktik. In B. Marquardt-Mau, W. Köhnlein & R. Lauterbach (Hrsg.), *Forschung zum Sachunterricht* (S. 43-60). Bad Heilbrunn: Klinkhardt.
- Girtler, R. (1984). *Methoden der qualitativen Sozialforschung. Anleitung zur Feldarbeit*. Wien: Böhlau.
- Gläser-Zikuda, M. (2001). *Emotionen und Lernstrategien in konkreten Lern- und Leistungssituationen: eine quantitativ – qualitative Studie mit Schülerinnen und Schülern*. Weinheim: Deutscher Studienverlag.
- Hänze, M. (1998). *Denken und Gefühl. Wechselwirkung von Emotion und Kognition im Unterricht*. Neuwied: Luchterhand.
- Haft, H. & Kordes, H. (Hrsg.) (1984). *Methoden der Erziehungs- und Bildungsforschung*. (= Enzyklopädie Erziehungswissenschaft, Bd 2) Stuttgart: Klett-Cotta.
- Hartertinger, A. (1997). *Interessenförderung. Eine Studie zum Sachunterricht*. Bad Heilbrunn: Klinkhardt.
- Hempel, M. (1997). Lebensentwürfe von Grundschulkindern – ein Forschungsthema für den Sachunterricht. In B. Marquardt-Mau, W. Köhnlein & R. Lauterbach (Hrsg.), *Forschung zum Sachunterricht* (S. 169-189). Bad Heilbrunn: Klinkhardt.
- Herrmann, T. & Tack, W. (Hrsg.) (1994). *Methodologische Grundlagen der Psychologie*. (= Enzyklopädie der Psychologie B, I, 1) Göttingen: Hogrefe.
- Jaeger, R. M. (Ed.) (1988). *Complementary methods for research in education*. Washington: AERA.
- Jerusalem, M. & Pekrun, R. (Hrsg.) (1999): *Emotion, Motivation und Leistung*. Göttingen: Hogrefe.
- König, E. & Zedler, P. (Hrsg.). (1995). *Bilanz qualitativer Forschung*. Weinheim: Deutscher Studien Verlag.
- Lamnek, S. (1988). *Qualitative Sozialforschung. Band 1: Methodologie*. München: Psychologie Verlags Union.
- Laukenmann, M., Bleicher, M., Fuß, S., Gläser-Zikuda, M., Mayring, Ph. & Rhöneck, C. v. (2000). Eine Untersuchung zum Einfluss emotionaler Faktoren auf das Lernen im Physikunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 6, 139-155.
- Mayring, Ph. (1996). Möglichkeiten qualitativer Ansätze in der Unterrichtsforschung. In G.W. Schnaitmann (Hrsg.), *Theorie und Praxis der Unterrichtsforschung* (S. 41-61). Donauwörth: Auer.

- Mayring, Ph. (1999a). *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. 4. Aufl. Weinheim: Psychologie Verlags Union.
- Mayring, Ph. (1999b) Qualitativ orientierte Forschungsmethoden in der Unterrichtswissenschaft – ein Anwendungsbeispiel aus der Lernstrategieforschung. *Unterrichtswissenschaft*, 27, 292-309.
- Mayring, Ph. (2000). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 7. Aufl. Weinheim: Deutscher Studien Verlag.
- Mayring, Ph. (2001). Kombination und Integration qualitativer und quantitativer Analyse (auch in Englisch und Spanisch). *Forum Qualitative Sozialforschung*, 2, H. 1) (<http://www.qualitative-research.net/fqs>).
- Mayring, Ph., Rhöneck, Ch. v. & Melenk, H. (1997). *Über den Einfluss emotionaler Faktoren auf das Lernen in den Fächern Physik und Deutsch*. Projektantrag. Ludwigsburg: Pädagogische Hochschule.
- Möller, J. & Köller, O. (Hrsg.) (1996). *Emotionen, Kognitionen und Schulleistung*. Weinheim: Psychologie-Verlags-Union.
- Möller, K. (2000). Lehr-Lernprozessforschung im naturwissenschaftlich-technischen Bereich des Sachunterrichts. In R. Duit & C. v. Rhöneck (Hrsg.), *Ergebnisse fachdidaktischer und psychologischer Lehr-Lern-Forschung. Beiträge zu einem Workshop an der Pädagogischen Hochschule Ludwigsburg* (S. 131-156). Kiel: IPN.
- Reinhoffer, B. (2000). *Heimatkunde und Sachunterricht im Anfangsunterricht. Entwicklungen, Stellenwert, Tendenzen*. Bad Heilbrunn: Klinkhardt.
- Richter, D. (Hrsg.) (2000). *Methoden der Unterrichtsinterpretation. Qualitative Analysen einer Sachunterrichtsstunde im Vergleich*. Weinheim: Juventa.
- Riedel, M. (1978). *Verstehen oder Erklären?* Stuttgart: Klett.
- Rost, D. (Hrsg.) (1998). *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz.
- Saldern, M. v. (1992). Qualitative Forschung - quantitative Forschung: Nekrolog auf einen Gegensatz. *Empirische Pädagogik*, 6, 377-399.
- Schmitz, B. (1996). Grundlagen der quantitativen Einzelfallanalyse. In Brähler, E. & Adler, C. (Hrsg.), *Quantitative Einzelfallanalysen und qualitative Verfahren* (S. 47-80). Gießen: Psycho-sozial-Verlag.
- Shulman, J. H. (Ed.) (1992). *Case methods in teacher education*. New York: Teachers College Press.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks: Sage.
- Steinke, I. (1999) *Kriterien qualitativer Forschung. Ansätze zur Bewertung qualitativ-empirischer Sozialforschung*. Weinheim: Juventa.
- Weitzman, E. B. & Miles, M. B. (1995). *Computer programs for qualitative data analysis*. Thousand Oaks: Sage.
- Wright, G. H. v. (1974). *Erklären und Verstehen*. Frankfurt: Athenäum.

## Qualitative und Quantitative Methoden in der fachdidaktischen Forschung

### 1. Die wissenschaftstheoretischen Grundlagen empirischer Forschung

Der Begriff der qualitativen Forschung stellt von seiner historischen Entwicklung her kein einheitliches Forschungsparadigma dar (vgl. Rost, 1998). Er ist vielmehr aus der Kritik an unterschiedlichen Aspekten der konventionellen empirischen Forschung hervorgegangen und definiert sich daher im wesentlichen *negativ*, d. h. dadurch, dass qualitative Forschung eben *nicht quantitativ* ist. Dies bringt uns einer Begriffsbestimmung aber noch nicht wesentlich näher, denn was sind die kritikwürdigen Merkmale herkömmlicher, als quantitativ bezeichneter Forschung und gibt es überhaupt überzeugende Alternativen hierzu, die sich als Merkmale qualitativer Forschung zusammenfassen ließen?

Die Kritik an herkömmlicher, empirischer Forschung, also an den derzeit anerkannten Forschungsparadigmen, bezieht sich so gut wie auf alle Stadien des Forschungsprozesses. In diesem Beitrag werden 6 Stadien oder Bestandteile des Forschungsprozesses unterschieden und jeweils diskutiert, inwiefern die herkömmlichen Konzeptionen als quantitativ bezeichnet und von alternativen, qualitativen Konzeptionen abgegrenzt werden können.

Beginnt man mit der wissenschaftstheoretischen Position, die der Forschung zugrunde liegt, so wendet sich die Kritik zunächst gegen das Popper'sche Falsifikationsprinzip. Nach diesem Paradigma besteht Forschung darin, aus Theorien abgeleitete Forschungshypothesen wiederholt empirischen Falsifikationsversuchen auszusetzen, bis die Empirie den Forscher dazu zwingt, seine Hypothese fallen zu lassen bzw. zu modifizieren und erneuten Falsifikationsversuchen auszusetzen (s. Kriz et al. 1987).

Dieses Forschungsprinzip kann in der empirischen Forschungspraxis tatsächlich zu kritikwürdigen Vorgehensweisen führen. So legt es den Hauptschwerpunkt der theoretischen Arbeit auf die Phase *vor* der Datenerhebung, in der es darum geht, den theoretischen Hintergrund einer empirischen Untersuchung auf wenige Hypothesen zuzuspitzen, über deren Wahrheitsgehalt dann anhand der Daten ent-

schieden werden kann. Die Auswertung der empirischen Daten besteht somit „lediglich“ darin, Auskunft zu geben, ob die jeweilige Hypothese abgelehnt werden muss oder nicht. Aus den empirischen Daten wird lediglich „1 Bit Information“ herausdestilliert (Falsifikation: ja oder nein), während die schwere Aufgabe der Theorienbildung *vor* der empirischen Datenerhebung (prä-empirisch) ausgeführt und abgeschlossen sein muss.

Der „laienhaften“ Vorstellung, dass man in der Forschung auf der Basis empirischer Beobachtungen Theorien *bildet*, steht im kritischen Rationalismus eine sehr reduzierte Funktion von Empirie gegenüber: Empirie gibt Auskunft darüber, ob eine erneute Theorienbildung oder -modifikation notwendig ist oder nicht, aber sie gibt keine Auskunft darüber, *wie* dies zu geschehen hat (Holzkamp, 1972).

Diese, aus der Sicht der qualitativen Forschung fragwürdige Reduktion empiristischer Erkenntnis, kann insofern mit dem Begriff „quantitativ“ charakterisiert werden, als es letztlich ein quantitatives Maß, d. h. eine Mittelwertsdifferenz, ein Korrelationskoeffizient oder ein Likelihoodquotient ist, das mittels einer zugeordneten Wahrscheinlichkeitsverteilung über die Beibehaltung oder Ablehnung der Forschungshypothese entscheidet. Die Forderung qualitativer Forschung ist es, aus empirischen Daten mehr qualitative Information zu ziehen, also Information darüber, *wie* die Theorie aussehen könnte, als lediglich die quantitative Information, inwieweit die Theorie falsifiziert werden muss (Flick et al. 1991; Lamnek, 1989; Mayring, 1996).

Wie ist diese Kritik zu bewerten? Es kann wohl nicht darum gehen, Popper in seinen grundlegenden erkenntnistheoretischen Aussagen und in der Stringenz seiner Argumentation zu kritisieren. Es ist zweifellos richtig, dass jede theoretische Erkenntnis aus dem Kopf heraus kommen muss (daher Rationalismus) und nicht aus den Daten abgelesen werden kann, wie der Empirismus dies so gerne hätte. Andererseits soll sich der Verstand (die Ratio) nicht irgendwelche Theorien ausdenken, sondern empirische Beobachtungen nutzen, um möglichst sinnvolle Theorien zu bilden. Genau dieser Prozess, wie man Theorien aufgrund von empirischen Daten *bildet*, um sie später mit Hilfe des Falsifikationsprinzips überprüfen zu können, wird von Popper nicht thematisiert, stellt aber einen wichtigen Schritt der Forschungspraxis dar (Popper, 1998).

Angesichts der vielen fragwürdigen Theorien, die in den empirischen Sozialwissenschaften einem Falsifikationsversuch ausgesetzt werden, kommt der Frage, *wie* der Forscher seine Theorie gebildet hat, eine immer größere Bedeutung zu. Beliebige, in ihrer Genese nicht begründete Theorien und Hypothesen immer erneuten Falsifikationsversuchen auszusetzen, kann zu einer absurden Landschaft von Forschungsergebnissen führen. Viele Kritiker sehen eine derartige Forschungslandschaft in der Psychologie oder Pädagogik bereits als gegeben an.

Erkenntnisfortschritt ist ein Wechselspiel von Theorie und Empirie. Eine Erkennt-

nislehre, die lediglich den deduktiven Prozess der Ableitung empirisch prüfbarer Aussagen aus Theorien beschreibt, aber nicht den induktiven Prozess der Entwicklung von Theorien auf der Grundlage empirischer Beobachtungen, ist für viele Bereiche sozialwissenschaftlicher Forschung tatsächlich defizitär. Erkenntnisfortschritt wird oft mit Hilfe eines Spiralmodells beschrieben, in dem sich Theoriebildung und empirische Beobachtung abwechseln (vgl. Abbildung 1, s. Sarris, 1990).

Aus einer Theorie wird mit Hilfe von deduktiven Prozessen eine Erwartung über empirische Sachverhalte abgeleitet, die einer empirischen Prüfung zugänglich ist. Aus empirischen Beobachtungen wird wiederum durch induktive Prozesse auf die Theorie zurückgeschlossen, diese modifiziert und erweitert. Erkenntnisfortschritt im Sinne einer spiralförmigen Weiterentwicklung kann nur stattfinden, wenn beide Richtungen einer Umdrehung in dieser Spirale funktionieren: die Überprüfung deduzierter Hypothesen muss ebenso erfolgen wie die Theoriemodifikation anhand von empirischen Daten. Qualitative Forschung betont den zweiten Halbkreis, also die Theorienbildung anhand von Daten und ergänzt sich somit mit dem deduktiven Halbkreis, der von Popper beschrieben wird. Die Forderung nach qualitativer Forschung bedeutet aus wissenschaftstheoretischer Perspektive, dass Theoriedefizite erkannt und benannt werden müssen, dass Theorien mit Hilfe von empirischen Daten weiterentwickelt werden müssen und Methoden zu entwickeln sind, aus empirischen Daten mittels induktiver Prozesse Theoriebildung zu betreiben.

Dass diesem Theorien-generierenden Halbkreis der Spirale des Erkenntnisfortschritts weniger wissenschaftliche Dignität zukommt liegt - historisch gesehen - am Scheitern des Positivismus. In der Positivismusdebatte zeigte sich, dass induktive Schlüsse von empirischen Beobachtungen auf theoretische Sätze nicht möglich sind, d. h. nicht die Qualität von *logisch wahren* Schlussfolgerungen besitzen (Kutschera, 1972). Insofern kann der induktive Weg von empirischen Daten zur Theorie niemals die Qualität logisch wahrer Ableitungen beanspruchen. Andererseits besteht der Wert und Nutzen des menschlichen Denkens vor allem in der Fähigkeit zum induktiven Denken, d. h. zum Ziehen von Schlussfolgerungen aus Wahrnehmungen und Beobachtungen. So gilt auch das induktive Denken (inductive reasoning) als zentrale Fähigkeit der menschlichen Intelligenz (Rost, 1997). Wir ziehen ständig (induktive) Schlüsse aus unseren alltäglichen Beobachtungen und tun gut daran so zu handeln, als ob diese Schlüsse wahr wären, obwohl sie niemals *logisch* wahr sein können.

Wir müssen uns in der empirischen Forschung wohl von der Vorstellung verabschieden, dass wir allein mit logisch korrekten Schlussfolgerungen arbeiten. Wenn wir den Kreis zwischen Empirie und Theorie vollständig durchlaufen wollen (und das müssen wir, um im Spiralmodell einen Erkenntnisfortschritt zu erzielen), so

müssen wir auch die Induktion von der Empirie auf die Theorie als Bestandteil des wissenschaftlichen Prozesses anerkennen.

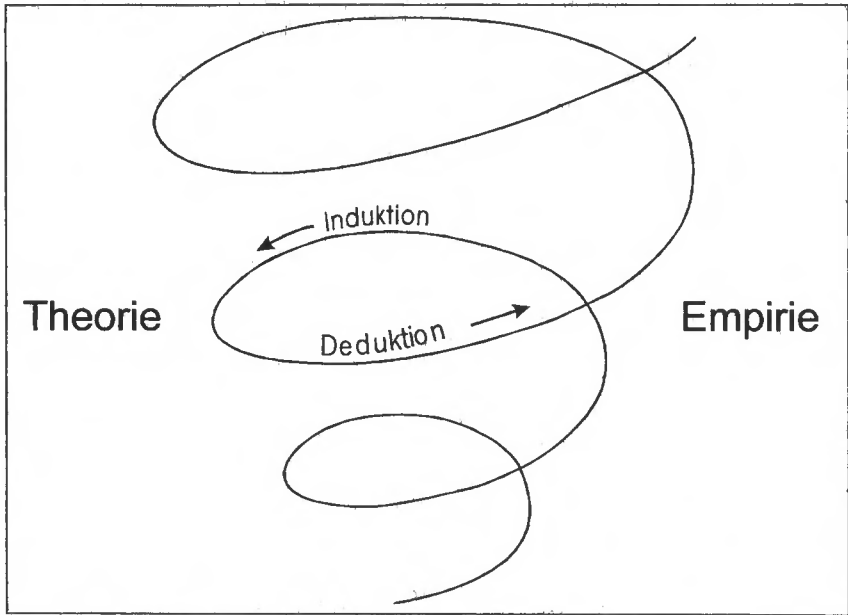


Abb. 1: Spiralmodell des Erkenntnisfortschritts

## 2. Hypothesen und Forschungsfragen

Am Anfang jeder Forschung steht eine Frage. Diese Feststellung ist ebenso allgemeingültig wie nichtssagend. Lässt sich doch jedem wissenschaftlichen oder nicht-wissenschaftlichen Erkenntnisvorgang nachträglich eine Fragestellung anheften, z. B. die allgemeine Frage „wie funktioniert das“ oder „wie lernen Schüler“.

Was macht eine Frage zum Ausgangspunkt empirischer Forschung oder wie muss eine Fragestellung formuliert sein, damit sie als Ausgangspunkt der methodischen Planung einer Forschungsarbeit dienen kann? In der herkömmlichen Forschung hat man weniger Wert darauf gelegt zu definieren was eine Fragestellung ist, sondern vielmehr was eine *Hypothese* ist. Hier wird auch nicht eine Fragestellung als Ausgangspunkt betrachtet, sondern vielmehr eine Hypothese. Eine Hypothese stellt eine vorweggenommene Antwort auf eine Frage dar und integriert somit die Fragestellung mit einer antizipierten Antwort. Dass nicht die Fragestellung sondern eine Hypothese als Ausgangspunkt von Forschung angesehen wird, ist aus der Sicht des kritischen Rationalismus nur konsequent, denn die Empirie kann unsere



Fragen nicht beantworten, sondern höchstens Auskunft darüber geben, ob unsere eigenen Antworten (Hypothesen) auf unsere Fragen richtig oder falsch sind.

Was macht man aber nun, wenn man zwar Fragen aber keine Hypothesen hat? Wie müssen Fragen beschaffen sein, um als Ausgangspunkt von Forschung dienen zu können, ohne dass man bereits Theorien hat, aus denen man erwartete Antworten (Hypothesen) auf diese Fragen ableiten kann? Die Lösung des Problems ist verblüffend einfach: Die Fragen müssen so beschaffen sein, dass sie potentiell durch Hypothesen beantwortbar sind. Unter einer Hypothese versteht man allgemein eine aus einer Theorie abgeleitete Aussage. Definierendes Merkmal einer Aussage ist es, dass sie wahr oder falsch sein kann. Aber was sind das für Aussagen, um die es in empirischer, psychologischer oder pädagogischer Forschung geht? Man spricht hier von Unterschiedshypothesen, also z. B. dass sich Versuchsgruppe und Kontrollgruppe in ihrem Mittelwert voneinander unterscheiden, oder von Zusammenhangshypothesen, z. B. dass das Merkmal  $xy$  eines Schülers mit seinem Lernerfolg zusammenhängt. Ebenso gibt es Entwicklungshypothesen, die zeitliche Verläufe beschreiben, oder prognostische Hypothesen, die das Eintreffen oder Ausbleiben bestimmter Ereignisse vorhersagen, z. B. Berufserfolg.

All diese Hypothesen beschreiben niemals singuläre Ereignisse, (z. B. Schüler  $xy$  hat das Lernziel der Klasse erreicht), sondern sie beziehen sich auf alle Mitglieder einer bestimmten Stichprobe oder Population. Wenn man einen Lernerfolg der Unterrichtsmethode  $xy$  erwartet, so besagt die Hypothese, dass alle Schüler, die diesen Unterricht genossen haben, einen höheren Lernerfolg haben als andere Schüler. Das „alle“ meint damit nicht, dass es für jeden einzelnen Schüler gelten muss, sondern es kann sich auch um eine statistische Aussage handeln, dass eben der Mittelwert dieser Gruppe höher ist. Dies ändert nichts an der Tatsache, dass solche Hypothesen von ihrer logischen Struktur her All-Aussagen sind, d. h. sich immer auf alle Mitglieder einer bestimmten Gruppe beziehen.

Betrachtet man die *Art* der Aussagen, die Hypothesen treffen, so kann man feststellen, dass Hypothesen im Bereich psychologisch-pädagogischer Forschung Aussagen über *Variablen* machen. Eine Variable ist eine Zuordnung von bestimmten Merkmalen zu allen Elementen einer Menge von Objekten, meist Personen. Wenn allen Mitgliedern einer Personenstichprobe das Merkmal „männlich“ oder „weiblich“ zugeordnet wird, so entsteht daraus eine Variable, nämlich die Variable „Geschlecht“. Ähnlich verhält es sich, wenn man den Personen Meßwerte in einem Test zuordnet oder ein qualitatives Urteil über eine erbrachte Leistung. Man kann also sagen, Hypothesen sind Aussagen über eine oder mehrere Variablen, wobei es sich in unserem Bereich meistens um Personenvariablen handelt.

Es läßt sich noch näher spezifizieren, *was* für Aussagen denn Hypothesen über Variablen machen. Die häufigsten Hypothesen sind, dass zwei Variablen zusammenhängen (korrelieren), oder dass eine Variable in unterschiedlichen Teilstich-

proben (Versuchsgruppen) unterschiedliche Werte oder Mittelwerte aufweist. Häufig sind auch Aussagen über den Zusammenhang von *drei* Variablen anzutreffen, z. B. dann wenn man Interaktionseffekte oder Moderatorwirkungen erwartet (s. u. Abschnitt 5). Generell gibt es keine Einschränkungen darüber, auf wie viele Variablen sich eine Hypothese bezieht. Gemeinsam haben jedoch all diese Hypothesen, dass sie Aussagen über bestimmte Verteilungseigenschaften von Variablen machen. Eine relativ präzise Definition dessen, was eine Hypothese ist, lautet daher: Eine Hypothese ist eine Aussage über Verteilungseigenschaften von Variablen (s. u. Abschnitt 6).

Nun zurück zur qualitativen Forschung. Wenn man in der qualitativen Forschung schon keine Hypothesen hat, so müssen doch die Fragestellungen so dargestellt sein, dass sie durch Hypothesen, also durch Aussagen über Verteilungseigenschaften von Variablen beantwortbar sind. Dies mag auf den ersten Blick sehr technisch und abstrakt klingen, ist jedoch äußerst hilfreich bei der Ausformulierung von handlungsleitenden Fragestellungen in der qualitativen Forschung. So macht z. B. die Fragestellung „Wie lernen Schüler im Unterricht xy?“ keine Aussage über irgendwelche Variablen und ihre Verteilungseigenschaften, und tatsächlich ist aus einer solchen Frage auch kein sinnvolles methodisches Vorgehen ableitbar. Demgegenüber stellt die Frage „Welche Verständnisschwierigkeiten haben Schüler, wenn sie mit Methode xy unterrichtet werden?“ eine Frage über die Verteilungseigenschaften von Variablen. Und zwar wird die Variable „Verständnisschwierigkeit“ angesprochen, die jedem Schüler als Merkmal ein oder mehrere Verständnisprobleme aus einer vielleicht noch unbekanntem Palette von solchen Problemen zuordnet. Außerdem gibt es die Variable der Unterrichtsmethode, die jedem Schüler als Merkmal zuweist, nach welcher Methode er unterrichtet wurde.

Die Formulierung von Forschungsfragen als Fragen über die Verteilungseigenschaften von Variablen ermöglicht es, die weitere Planung der Untersuchung mit Hilfe des bestehenden und bewährten Methodeninventars vorzunehmen. So gilt es, im nächsten Schritt bezüglich der in der Fragestellung angesprochenen Variablen zu klären, wie sie definiert sind, wie andere Autoren sie definiert haben, welche Forschungsergebnisse es bislang hierzu gibt, wie man sie operationalisieren kann, welche Probleme es bei ihrer Messung gibt usw. Eine Forschungsarbeit ist in dem Ausmaß als „qualitativ“ zu bezeichnen, je weniger gesichertes Wissen es über die beteiligten Variablen gibt. Wurden z. B. die vermuteten Verständnisschwierigkeiten bislang noch nie in dieser Form untersucht, so ist es ein erstes Anliegen der (qualitativen) Forschungsarbeit, erst einmal die Existenz der Variable „Verständnisschwierigkeiten“ nachzuweisen. Es ist nämlich durchaus nicht selbstverständlich, dass es gelingt, jedem Schüler genau ein Merkmal aus einer Menge von Verständnisschwierigkeiten zuzuordnen oder auch eine Kombination bestimmter Verständnisschwierigkeiten. Ergeben z. B. Voruntersuchungen, dass jeder Schüler

unter einem undefinierbaren Konglomerat verschiedener Verständnisprobleme leidet, so kann man mit dieser (Schein-)Variable keine Forschung betreiben. Aufgabe qualitativer Forschung wäre es daher, erst einmal unterscheidbare Verständnisschwierigkeiten zu identifizieren, um daraus dann eine Variable konstruieren zu können.

Die Funktion von Forschungsfragen ist es, das weitere methodische Vorgehen zu strukturieren. Forschungsfragen, die als Fragen über Verteilungseigenschaften von Variablen formuliert sind, erfüllen diese Funktion. Inwieweit die zur Beantwortung dieser Forschungsfragen durchgeführte Forschung qualitativ ist, drückt sich darin aus, in welchem Ausmaß neue Variablen erst operationalisiert und identifizierbar gemacht werden müssen, und ob über ihre Verteilungseigenschaften bereits Erwartungen vorliegen.

### 3. Die Erhebung von Daten

Zu den aus der Sicht der qualitativen Forschung kritisierten Verfahren der Datenerhebung gehören Fragebögen, standardisierte Testinstrumente, physiologische Messungen, standardisierte Beobachtung und standardisiertes Interview. Demgegenüber gelten narrative Interviews, teilnehmende Beobachtung, Videoaufzeichnungen und die Erhebung verbaler Daten, aber auch Inhaltsanalyse, biographische Daten und ethnographische Formen der Datenerhebung als qualitative Methoden (Mayring 1999). Versucht man hierin ein System zu entdecken, so ist es weniger die Unterscheidung von qualitativ und quantitativ, die hier weiterhelfen könnte. Ausschlaggebendes Moment ist vielmehr der *Informationsreichtum* der Datenquelle. In Multiple-Choice-Fragebögen wird das Verhalten der Menschen auf die Platzierung eines Kreuzes in das richtige Kästchen reduziert. Kennzeichnend für die Fragebogenmethode ist, dass der große Verhaltensspielraum der Menschen in einer Situation einschließlich verbaler, mimischer und gestischer Ausdrucksformen radikal eingegrenzt wird und man bei der Datenerhebung auf sehr viel Information verzichtet. Die Reduktion des Verhaltens auf die Platzierung eines Kreuzchens stellt zwar aus der Sicht der „quantitativen“ Forscher genau jene Reduktion dar, die für die Beantwortung einer Hypothese benötigt wird. Damit wird aber auf zusätzliche Informationen, wie z. B. die Sicherheit, mit der die Antwort gegeben wird, etwaige Randbedingungen, unter denen die jeweilige Antwort gilt, oder emotionale Begleitreaktionen verzichtet.

Ähnlich verhält es sich mit Leistungstests, bei denen nur das maximal erreichbare Leistungsniveau erhoben wird, nicht aber die subjektive Anstrengung, die Konzentrationsleistung, die Art der kognitiven Prozesse oder eventuelle Gründe für eine Leistungsminderung oder gar die subjektive Bedeutsamkeit der erbrachten Leistung. Auch hier ist es die Informationsarmut, also die Reduktion auf die An-

zahl richtig gelöster Testaufgaben oder die Bearbeitungsgeschwindigkeit, die Tests als quantitative Methoden auszeichnet und von qualitativen Methoden abgrenzt. Ob Interviews oder Beobachtungsmethoden eher den quantitativen oder den qualitativen Verfahren zugerechnet werden, hängt vom Grad ihrer Standardisierung ab. Ein hoher Grad an Standardisierung bedeutet eine Einschränkung des Spektrums beobachtbarer oder erfragbarer Verhaltensaspekte auf jene Aspekte, die zur Beantwortung der Untersuchungshypothese notwendig sind. Weniger gesteuerte Formen der Beobachtung oder des Interviews, die gar eine Interaktion mit den zu Untersuchenden zulassen, liefern dagegen eine größere Vielfalt an Informationen über die befragten oder beobachteten Personen.

Ganz eklatant wird die Informationsfülle als Definitionsmerkmal qualitativer Methoden dann, wenn man an Videoaufzeichnungen, Tonbandmitschnitte, umfangreiche verbale Daten, Lebensläufe oder Tagebuchaufzeichnungen denkt. Da der Informationsreichtum das ausschlaggebende Moment qualitativer Datenerhebungsmethoden ist, soll hier auch nicht von quantitativen und qualitativen Daten gesprochen werden (was höchst mißverständlich ist), sondern von informationsreichen und informationsarmen Daten(-quellen). Dabei kann ‚informationsarm‘ nicht per se abwertend gemeint sein, denn die Datenreduktion kann sehr sinnvoll auf diejenigen Verhaltensaspekte erfolgen, auf die es in der Untersuchung ankommt. Ebenso ist ‚informationsreich‘ per se noch kein Merkmal hoher Forschungsqualität, denn es kann passieren, dass man nicht weiß, wie man die enorme Informationsfülle reduzieren und für die Theorienbildung nutzen soll.

Dass der Informationsreichtum der Datenquelle dem Anliegen qualitativer Forschung und damit der Kritik an quantitativer Forschung entspricht, ist nicht zufällig. Ist es doch das im vorangehenden Kapitel beschriebene Defizit an einer wissenschaftlich fundierten Theorienbildung aufgrund empirischer Daten, die einen größeren Informationsreichtum empirischer Daten logisch notwendig werden läßt. Es läßt sich sogar die scheinbar paradoxe Beziehung aufstellen: je weiter fortgeschritten eine Theorienentwicklung ist, desto informationsärmer können die Daten sein, die man zur Überprüfung der Theorie benötigt. Je schwächer und unvollkommener eine Theorie ist, desto informationsreicher sollten die Daten sein, da es nicht nur gilt, die Theorie zu überprüfen, sondern sie zunächst zur Gänze zu entwickeln. Die Kritik an der quantitativen Forschung besteht natürlich nicht darin, dass ihre Theorien schon so weit entwickelt sind, dass man sie mit auf das Wesentliche reduzierten Daten testen kann. Der Kern der Kritik geht vielmehr dahin, dass Theorien, die sich mit informationsarmen Daten testen lassen, eher langweilig, trivial oder vom psychologischen Gehalt her reduktionistisch sind. Man mag der Kritik zustimmen, dass mit Test- und Fragebogendaten oft nur uninteressante Theorien untersucht werden. Jedoch ist der Beweis, dass man mit informationsreicheren Datenquellen zu wesentlich interessanteren und spannenderen Theorien gelangt, auch noch nicht erbracht.

#### 4. Kodierung und Skalierung

Die Kritik am konventionellen Paradigma der quantitativen Forschung geht dahin, dass die erhobenen Daten, wie informationsreich auch immer sie sein mögen, nach vorgefertigten Kodierungsregeln in Zahlen transformiert und weiterhin zu quantitativen Meßwerten verarbeitet werden. Im Rahmen der quantitativen Forschung müssen in der Tat die Kodierungsregeln prä-experimentell, d. h. vor der Datenerhebung festgelegt sein, da dies dem erkenntnistheoretischen Prinzip entspricht, dass die Datenauswertung allein der Bestätigung oder Ablehnung der Forschungshypothese dient. Würde die Datenkodierung und -auswertung selbst von den Daten abhängig gemacht, also das Kodierungsschema erst anhand der erhobenen Daten entwickelt werden, so wäre die Objektivität der Hypothesenprüfung eingeschränkt. Es wäre möglich, dass die Entwicklung eines Kodiersystems anhand der Daten so erfolgt, dass die Forschungshypothese damit eher bestätigt wird.

Andererseits hat es eine gewisse Absurdität, informationsreiche Daten zu erheben, wenn man sie dann doch auf Auswertungskategorien reduziert, die schon *vor* der Untersuchung feststanden (Strauß & Corbin, 1996). Die Erhebung informationsreicher Daten ist überflüssig, wenn man diesen keine Chance läßt, den Forscher von der Unzulänglichkeit seines Kategoriensystems zu überzeugen. Es gibt Möglichkeiten, ein Kategoriensystem anhand der Daten zu entwickeln und trotzdem der Forderung nach einer objektiven Datenauswertung gerecht zu werden. Es kann z. B. die nach einem Zufallsprinzip aufgeteilte erste Hälfte der Datenmenge zur Entwicklung eines Auswertungssystems benutzt werden, um dieses dann bei der zweiten Hälfte der Daten zur Hypothesentestung einzusetzen. Auch im Rahmen qualitativer Forschung kann und sollte die Intersubjektivität der Datenkodierung kontrolliert und überprüft werden. Dies ist z. B. dadurch realisierbar, dass alle Daten nach Entwicklung des Kategoriensystems erneut von unabhängigen Kodierern ausgewertet werden und anhand einer doppelt ausgewerteten Teilmenge der Daten die Übereinstimmung zwischen zwei unabhängigen Kodierern berechnet wird (s. Rost, 1996).

Sind die Daten kodiert, so sind in einem zweiten Schritt der Datenreduktion Meßwerte für die Variablen der Untersuchung abzuleiten. Die Kritik an quantitativer Forschung bedeutet nicht, gänzlich auf Meßwerte zu verzichten, denn Meßwerte benötigt man, um die Variablen der Untersuchung zu erfassen und die Forschungsfrage zu beantworten. Allerdings gilt auch hier, dass bei herkömmlicher „quantitativer Forschung“ die Meßvorschrift prä-experimentell festliegt oder festliegen sollte, während es bei qualitativer Forschung erst einen Akt der Datenauswertung darstellt, geeignete Meßwerte zu identifizieren und somit die Variablen zu operationalisieren.

Bei diesem Schritt der Bildung von Meßwerten ist erstmals der Unterschied zwischen quantitativer und qualitativer Forschung wörtlich zu nehmen: Üblich ist die

Bildung *quantitativer* Meßwerte, sei es durch die Bildung von Summenscores oder den Einsatz der Faktorenanalyse, die auf heuristische Weise quantitative latente Variablen ermittelt. Dass es auch möglich ist, *qualitative* also kategoriale Variablen zu erheben und mit entsprechenden Meßvorschriften meßbar zu machen, ist nicht nur im Rahmen quantitativer Forschung, sondern auch im Rahmen qualitativer Forschung weitgehend unbekannt.

„Messen“ wird in der Regel als synonym zu „quantifizieren“ verstanden, nicht nur in der naturwissenschaftlichen, sondern auch in der sozialwissenschaftlichen Forschung. Dabei wird übersehen, dass die zu messenden Konstrukte in der sozialwissenschaftlichen Forschung oft das Skalenniveau einer Nominalskala haben, also typologische Konstrukte sind. Beispiele sind etwa kognitive Stile, Attributionsstile, kognitive Strategien (Rost, 1995), Arten der Zielorientierung im Leistungsverhalten (Köller et al. 1998) oder die Stufen intrinsischer Motivation nach Deci und Ryan. Obwohl solche Konstrukte qualitative Unterschiede zwischen Personen beschreiben, werden sie in der empirischen Forschung in der Regel mit quantitativen Variablen erfaßt.

In den letzten Jahren sind Verfahren auf der Grundlage der Latent-Class-Analyse (Lazarsfeld und Henry, 1968) entwickelt worden, die es erlauben kategoriale, also qualitative Personenvariablen zu messen (Rost, 1996). Ausgangspunkt für diese Verfahren sind kategoriale oder ordinale Daten, also z. B. Antworten in einem Fragebogen, kategorisierte Beobachtungsdaten oder Interviewdaten. Es werden nicht Summenwerte über die Itemantworten gebildet, sondern es werden die Personen danach klassifiziert, wie ähnlich das Profil ihrer Itemantworten ist. Dieses qualitative Konzept von „Messen“, also die Klassifikation von Personen anhand ihres Antwortprofils, ist im Rahmen probabilistischer Testmodelle eine interessante Alternative zur Quantifizierung, wie sie im Rahmen der klassischen Testtheorie und der Item-Response-Theorie angestrebt wird. Es läßt sich bei den Latent-Class-analytischen Verfahren wie auch z. B. beim Rasch-Modell prüfen, ob die Items tatsächlich die gewünschte quantitative oder kategoriale Variable messen. Speziell ist auch der Vergleich möglich, ob die durch eine Menge von Items erfaßte Personenvariable eher quantitativ oder eher kategorial ist. Man braucht für diese Art der Testauswertung allerdings größere Personenstichproben (mindestens 60 - 70 Personen).

Hinsichtlich der Skalierung der Variablen sollte der Unterschied zwischen qualitativer und quantitativer Forschung nicht darin bestehen, ob eine meßtheoretisch fundierte Skalierung vorgenommen wird oder nicht, sondern darin ob qualitative Personenvariablen auch tatsächlich als solche skaliert werden. Darüber hinaus kommen Auswertungsverfahren wie die Latent-Class-Analyse auch für eine explorative Skalenanalyse in Frage, wie sie im Rahmen qualitativer Forschung meistens erforderlich ist. Es werden keine prä-experimentellen Hypothesen über die Art der

latentem Klassen vorausgesetzt, sondern die Latent-Class-Analyse ermittelt eine Klassifikation der Personen, die im Sinne einer Theorienbildung anhand von Daten interpretiert werden kann.

## 5. Das Forschungsdesign

Das in der klassischen, sprich quantitativen Forschung vorherrschende Design ist sicherlich das einer faktoriellen Versuchsplanung mit einer oder mehreren unabhängigen Variablen und der abhängigen Variablen, auf die eine Wirkung der unabhängigen Variablen erwartet wird. Als quantitative Forschung werden aber auch großflächige Fragebogenuntersuchungen bezeichnet, in denen eine Vielzahl von (quantitativen) Variablen erhoben und miteinander korreliert werden. Die beiden Prototypen experimenteller und nicht-experimenteller empirischer Forschung sind in ihrer Begründung und Aussagekraft derart unterschiedlich, dass man sich fragen kann, inwiefern sie beide in den gemeinsamen Topf der quantitativen Forschung getan werden können.

So benötigt man für experimentelle Studien im allgemeinen kleinere Stichproben, während für Fragebogenstudien sehr große Stichproben erforderlich sind. In experimentellen Studien ist man dazu gezwungen, prä-experimentelle Hypothesen festzulegen. Fragebogenstudien werden dagegen allzu oft ohne besonders begründete Theorien durchgeführt. In experimentellen Studien wird versucht, möglichst alle Drittvariablen zu kontrollieren, während Fragebogenstudien im allgemeinen Feldforschung darstellen und dem weißen Rauschen einer Vielzahl von unkontrollierbaren Drittvariablen ausgesetzt sind. Was haben diese Forschungsarten gemeinsam, was sie aus der Sicht der qualitativen Forschung disqualifiziert?

Gemeinsam ist ihnen z. B. das Mißverhältnis von organisatorischem Aufwand für die Datenerhebung (Planung eines Experiments oder Vorgabe eines komplexen Fragebogens an einer großen Stichprobe) und dem Ergebnis, also der Information, ob ein Mittelwertunterschied statistisch bedeutsam ist (im Fall experimenteller Studien) oder ein Korrelationskoeffizient signifikant von Null abweicht (im Fall von Fragebogenstudien). Für beide Arten von (quantitativer) Forschung haben sich im Lauf der Zeit gewisse Standards und dazugehöriges Handwerkszeug entwickelt, so dass sich die Bewertung von Forschung oft daran zu orientieren scheint, inwieweit das Handwerkszeug professionell eingesetzt wurde. Welche Erkenntnis für die Weiterentwicklung der Theorien mit dem hohen empirischen Aufwand gewonnen wurde, bleibt oft nebensächlich.

Gibt es denn typische qualitative Forschungsdesigns, die man den beiden klassischen quantitativen Forschungsdesigns gegenüberstellen könnte? Nun, ein qualitatives Forschungsdesign scheint sich durch kleine Stichproben, mögliche Interaktionen zwischen Versuchspersonen und Versuchsleiter, informationsreiche Daten-

aufnahme und minimale Restriktivität der Rahmenbedingungen auszuzeichnen. Mayring (1999) nennt als typische qualitativ-orientierte Forschungsdesigns die Einzelfallanalyse nach der biographischen Methode, die Dokumentenanalyse, Handlungsforschung, Feldforschung nach der ethnographischen Methode und formative Evaluationsforschung. Es stellt sich die Frage, ob im Rahmen derartiger Forschungsdesigns neue Regeln oder Prinzipien der Versuchsplanung gelten oder die klassischen bewährten Prinzipien sinngemäß anzuwenden sind.

An drei zentralen Beispielen soll im folgenden deutlich gemacht werden, dass die abstrakten Prinzipien der Versuchsplanung (Huber, 1987; Kerlinger, 1975) auch in der qualitativen Forschung - wenn auch in abgewandelter Ausgestaltung - angewendet werden können und müssen.

Das erste Beispiel ist die Forderung nach einer *Kontrollgruppe*. Es gibt viele Kritikpunkte am klassischen Kontrollgruppendesign: Die Realisierung der Kontrollbedingung erfordert gleichen Aufwand wie die Realisierung der Versuchsbedingung, ist aber von vornherein inhaltlich uninteressant, da die theoretisch interessante Maßnahme in ihr gar nicht realisiert ist. Sie muss „mitgeschleppt“ werden, nur um nachzuweisen, dass die Versuchsbedingung besser ist. *Warum* sie nun wirklich besser ist, geht aus dem Vergleich oft gar nicht hervor. Daraus den Schluß zu ziehen, dass auf jegliche Vergleichsgruppen verzichtet werden kann, wäre jedoch fatal. Auch in der qualitativen Forschung ist es sinnvoll, nicht nur *eine* homogene Gruppe zu untersuchen, da erst der Vergleich mit Gruppen, in denen etwas anderes passiert, Aufschluß darüber gibt, welches Ergebnis spezifisch für die untersuchte Population ist. Gerade der Prozess der *Generierung* von Erkenntnis erfordert den Kontrast. Erkenntnisse darüber, wie z. B. Grundschüler im Rahmen eines problemlösenden Unterrichts lernen, gewinnen erst dadurch an Bedeutung, dass man weiß, unter welchen anderen Bedingungen sie anders lernen würden.

Es gibt mehrere Alternativen zur klassischen Kontrollgruppe, die die Funktion erfüllen, einen Vergleichsmaßstab für die Größe und Bedeutsamkeit von empirischen Befunden zu liefern. Eine mögliche Alternative besteht darin, die Wirkung bestimmter Maßnahmen, z. B. zur Förderung einer autonomen Lernmotivation in verschiedenen Dosierungen zu untersuchen. Damit ist ebenfalls eine Abschätzung der Effektstärke des untersuchten Unterrichts möglich, ohne eine Kontrollgruppe mit sogenanntem Normalunterricht in die Untersuchung einbeziehen zu müssen. Eine andere Alternative ist es, die interessierende Personengruppe nicht mit einer „Normalpopulation“ als Kontrollgruppe zu vergleichen, sondern mit anderen Populationen, zu denen deutliche Unterschiede erwartet werden. Sollen z. B. bestimmte Leistungsmerkmale von Waldorfschülern untersucht werden, so können anstelle einer schwer definierbaren „Normalstichprobe“ auch Schüler von Schulen mit einem ausgeprägten musischen Zweig, Leistungskursschüler mit einem bestimmten Schwerpunkt oder Schüler der Montessoripädagogik herangezogen wer-



den. Solche wohldefinierten Vergleichsgruppen sind aussagekräftiger als eine schlecht definierte Kontrollgruppe, da in ihnen bestimmte Merkmale der zu untersuchenden Maßnahme (Waldorfpädagogik) in unterschiedlicher Kombination und Ausprägung realisiert sein können.

Die klassische Kontrollgruppe definiert sich durch die *Abwesenheit* einer bestimmten Maßnahme, was im pädagogischen Kontext oft nicht sinnvoll zu realisieren ist. Andersartig definierte Vergleichsgruppen können dieselbe Funktion erfüllen, nämlich einen Maßstab zur Beurteilung der Ergebnisse zu liefern, und geben darüber hinaus noch detailliertere Hinweise, worauf die beobachteten Gruppenunterschiede zurückzuführen sein könnten.

Das zweite zentrale Prinzip der Versuchsplanung betrifft die Kontrolle von Störvariablen. Ein Kritikpunkt an quantitativer Forschung besteht darin, dass mit relativ hohem Aufwand Laborbedingungen geschaffen werden, extrem homogene oder parallelisierte Stichproben rekrutiert werden, möglicherweise relevante Drittvariablen ausgeschaltet oder konstant gehalten werden, um in diesem Rahmen einzelne Variablen untersuchen zu können. Dieses Vorgehen schränkt die Aussagekraft der Untersuchung ein, beeinträchtigt ihre ökologische Validität und kann sogar dazu führen, dass das zu untersuchende komplexe Phänomen in dem Versuchsaufbau gar nicht in Erscheinung tritt. Aus dieser Kritik den Schluß zu ziehen, dass man im Rahmen qualitativer Forschung gänzlich auf die Kontrolle von Störvariablen verzichten sollte, wäre jedoch fatal. Sowohl in quantitativer wie in qualitativer Forschung geht es darum, auszuschließen, dass die empirisch ermittelten Zusammenhänge zwischen Variablen möglicherweise auf den Einfluß ganz anderer Variablen zurückzuführen sind.

Das zugrundeliegende abstrakte Prinzip heißt „Kontrolle von Störvariablen“ und man mag aus qualitativer Sicht einwenden, dass man noch gar nicht weiß, welche Variable eine Störvariable ist oder tatsächlich zum untersuchten Phänomen gehört. Dann gilt es, alle in Betracht kommenden Variablen zu benennen und jeweils zu entscheiden, ob sie mit untersucht werden sollen oder nicht. Auch in der qualitativen Forschung gibt es Störvariablen, die verhindern können, dass sich aus den empirischen Daten diejenigen Erkenntnisse ableiten lassen, die für die Theoriebildung wichtig wären.

Man unterscheidet verschiedene Formen der Kontrolle von Störvariablen, nämlich das Konstanthalten, die Randomisierung, Messwiederholungen, die Einbeziehung in das Forschungsdesign und die statistische Kontrolle. Keine dieser Kontrolltechniken ist von vornherein für qualitative Forschung untauglich, jedoch sind bestimmte Techniken, z. B. die statistische Kontrolle von Störvariablen eher dem Sachverhalt angemessen, dass in qualitativen Forschungsdesigns oft nur schwache Annahmen über potentielle Störvariablen existieren.

Das dritte zentrale Prinzip der Versuchsplanung besteht darin, mögliche *Mode-*

*ratorvariablen* zu entdecken und das Ausmaß ihres Einflusses abzuschätzen. Worin unterscheiden sich Störvariablen, die es zu kontrollieren gilt, von Moderatorvariablen? Störvariablen korrelieren mit Variablen, an deren Zusammenhang man interessiert ist. Korreliert z. B. eine Variable A mit der theoretisch interessierenden Variable B, aber auch mit der potentiellen Störvariable C, so besteht eine Unsicherheit, ob sich tatsächlich A und B gegenseitig beeinflussen oder ob die Störvariable C für den beobachteten Zusammenhang verantwortlich ist. Eine Moderatorvariable korreliert dagegen nicht notwendigerweise mit A oder B, sondern sie verändert (moderiert) den Zusammenhang zwischen A und B, d. h. dieser ist unterschiedlich für verschiedene Ausprägungen der Moderatorvariablen C.

Die Suche nach Moderatorvariablen dient der Absicherung von Kausalinterpretationen, denn ein empirisch gefundener Zusammenhang zwischen zwei Variablen oder ein Mittelwertunterschied zwischen zwei Gruppen sagt nur dann etwas aus, wenn dieser Zusammenhang oder Unterschied invariant ist für alle denkbaren Untergruppen. Hat z. B. ein bestimmter Unterricht eine nachweisbare Wirkung auf die Lernmotivation der Schüler, so kann die beobachtete Motivationsänderung nur dann auf den Unterricht zurückgeführt werden, wenn dieser Effekt bei Jungen und Mädchen, bei älteren und jüngeren Schülern, in verschiedenen Fächern etc. gleich ist. Ist dies nicht der Fall, so muss das Ergebnis: „Der Unterricht hat einen Effekt auf die Lernmotivation“ eingeschränkt oder modifiziert werden, etwa derart, dass der Unterricht besonders bei leistungsschwachen Jungen der 9. Jahrgangsstufe einen Effekt auf die Lernmotivation hat.

Auch kann sich bei der Suche nach Moderatoren ein bedeutsamer Effekt in bestimmten Teilgruppen zeigen, der in der Gesamtstichprobe nicht beobachtbar war. Z. B. kann sich ein positiver Effekt bei leistungsschwachen Schülern mit einem negativen Effekt bei leistungsstarken Schülern gegenseitig neutralisieren, so dass man bei einer Analyse der Gesamtpopulation zu dem Schluß kommt, der Unterricht habe keinen Effekt (s. u. Abschnitt 6).

Diese Grundprinzipien der Versuchsplanung, Bildung einer Vergleichsgruppe, Kontrolle von Störvariablen und Identifizierung von Moderatorvariablen, müssen ebenso für die Planung einer qualitativen Studie gelten wie für eine quantitative Studie. Es gibt keinen vernünftigen Grund, weswegen man im Rahmen qualitativer Forschungsdesigns darauf verzichten könnte. Ganz im Gegenteil: Wenn das Ziel qualitativer Forschung darin besteht, die Brücke von den empirischen Daten zurück zur Theorie zu schlagen (s. Abschnitt 1), so sind diese Prinzipien der Versuchsplanung umso bedeutsamer, denn sie tragen dazu bei, eine theoretische Interpretation empirischer Beobachtungen abzusichern.

## 6. Die Auswertung der Daten

Qualitative Forschung beantwortet Fragen, die sich auf die Verteilungseigenschaften von Variablen beziehen. Dementsprechend lassen sich Auswertungsverfahren danach einteilen, ob sie sich auf die Verteilungseigenschaften *einer* Variablen oder von zwei, drei oder mehreren Variablen beziehen. Oft geht es bei qualitativer Forschung um den Nachweis und die Exploration *einer einzelnen* Variable, z. B. die Unterscheidung verschiedener Lernmotivationen beim Schüler oder die Erhebung vorunterrichtlicher Schülervorstellungen zu einem bestimmten Sachgebiet. Egal, ob es sich hierbei um eine quantitative Variable handelt (Ausmaß der Lernmotivation) oder um eine qualitative Variable (Art der Lernmotivation), besteht der erste Auswertungsschritt darin, die Varianz der Variable zu berechnen und zu untersuchen. Nicht zufällig haben die Begriffe Variable und Varianz denselben Wortstamm, d. h. eine Variable ohne Varianz ist keine Variable, sondern eine Konstante. So banal diese Feststellung auch sein mag, sie betont die Notwendigkeit, das *jede* Variable im Untersuchungsdesign eine möglichst große Varianz haben muss. Nur was variiert, ist einer empirischen Analyse zugänglich, haben z. B. alle Schüler dieselbe Lernmotivation, so kann auch im Rahmen qualitativer Forschung nichts über die Zusammenhänge dieser Variable mit anderen Variablen herausgefunden werden.

Die Varianz ist neben dem Mittelwert die zweite zentrale Eigenschaft einer Verteilung. Beide Verteilungseigenschaften müssen gemeinsam interpretiert werden, denn ein Mittelwert mit geringer Varianz sagt etwas anderes aus als ein Mittelwert mit hoher Varianz. Beide Verteilungsmerkmale sind wiederum in Abhängigkeit von der *Form* der Verteilung zu interpretieren: Ist die Verteilung einer Variable zweigipflig (bimodal), so ist der Mittelwert eine mißverständliche Größe und die relativ große Varianz nur Ausdruck der Bimodalität.

Der nächste Schritt der Datenauswertung besteht darin abzuschätzen, welcher Anteil der beobachteten Varianz auf den (stets gegebenen) Meßfehler zurückzuführen ist und welcher Anteil sogenannte wahre Varianz darstellt. Der Anteil der wahren Varianz an der beobachteten Varianz einer Variable bezeichnet man als *Reliabilität*, die auf verschiedene Weise berechnet oder abgeschätzt werden kann (Lienert und Raatz 1994). Die Berechnung der Reliabilität einer Variable ist dann wichtig, wenn es um die Identifizierung und den Nachweis neuer, theoretisch interessanter Variablen geht.

Viele Fragestellungen qualitativer Forschung beziehen sich jedoch nicht auf die Verteilungseigenschaften einer Variable, sondern auf zwei Variablen. Man hat es mit einer sogenannten bivariaten Verteilung zu tun, zum Beispiel die gemeinsame Verteilung von Lernmotivation und Lernerfolg. Sind beide Variablen dieser bivariaten Verteilung quantitativ, so ist die interessanteste und am häufigsten be-

rechnete Verteilungseigenschaft die Korrelation. Der Korrelationskoeffizient drückt aus, inwieweit beide Variablen gleichsinnig variieren, d. h. *kovariieren*. Geht hohe Lernmotivation mit hohem Lernerfolg einher, so ist die Korrelation positiv, geht hohe Lernmotivation mit einem geringen Lernerfolg einher, so wäre die Korrelation negativ.

Ist eine Variable quantitativ und die andere Variable kategorial, z. B. bei einem Gruppenvergleich, so ist der Mittelwertunterschied die zentrale Eigenschaft dieser bivariaten Verteilung. Der Mittelwertunterschied der quantitativen Variable in den zwei oder mehr Versuchsgruppen drückt den Zusammenhang zwischen unabhängiger und abhängiger Variable aus, ist also ein Zusammenhangsmaß. Je größer der Mittelwertunterschied desto höher die Korrelation.

Sind beide Variablen kategorial, z. B. die Art der Lernmotivation und die Art der Aufgabenlösung, so repräsentiert eine sogenannte *Kreuztabelle* die bivariate Verteilung der beiden Variablen. Der Zusammenhang beider Variablen drückt sich in einer Kreuztabelle darin aus, dass einige Zellen häufiger besetzt sind, als es aufgrund der beiden (univariaten) Verteilungen und der Annahme, dass diese beiden Variablen unabhängig sind, zu erwarten wäre. Die Stärke des Zusammenhangs läßt sich in einem sogenannten Kontingenzmaß ausdrücken.

Der Zusammenhang zweier Variablen, sei es eine Korrelation, eine Kontingenz oder ein Mittelwertunterschied, ist sicherlich die interessanteste Verteilungseigenschaft einer bivariaten Verteilung. Jedoch repräsentieren diese Zusammenhangsparameter noch nicht die Art oder Qualität des Zusammenhangs. So kann der Zusammenhang zwischen zwei quantitativen Variablen nicht-linear sein oder die Kontingenz zwischen zwei kategorialen Variablen nur auf dem Unterschied zweier Valenzen der einen Variablen beruhen. Hierauf kann jedoch nicht weiter eingegangen werden (s. z. B. Bortz, 1984).

Viel interessanter wird die Angelegenheit, wenn drei Variablen an einer Fragestellung beteiligt sind. Hier gibt es sehr viele verschiedene Möglichkeiten die trivariate Verteilung zu beschreiben und verschiedene Begriffe kennzeichnen unterschiedliche Arten des Zusammenhangs von drei Variablen.

Die einfachste und am häufigsten angewendete Art, einen trivariaten Zusammenhang zu beschreiben, besteht in der Rückführung auf die drei bivariaten Zusammenhänge. D. h., die trivariate Verteilung von den drei Variablen A, B und C wird zu beschreiben versucht, indem man die drei Korrelationen AB, BC und AC betrachtet. Dies ist z. B. dann der Fall, wenn man die Korrelationsmatrix aller Variablen berechnet und mit dieser z. B. eine Faktorenanalyse oder ein lineares Strukturgleichungsmodell berechnet. Das Bemerkenswerte an diesem Vorgehen ist, dass sich „echte“ trivariate Zusammenhänge gar nicht in den bivariaten Zusammenhängen niederschlagen müssen bzw. durch sie sehr reduziert und verzerrt wiedergegeben werden können. Abbildung 2 verdeutlicht dies an einem Beispiel.

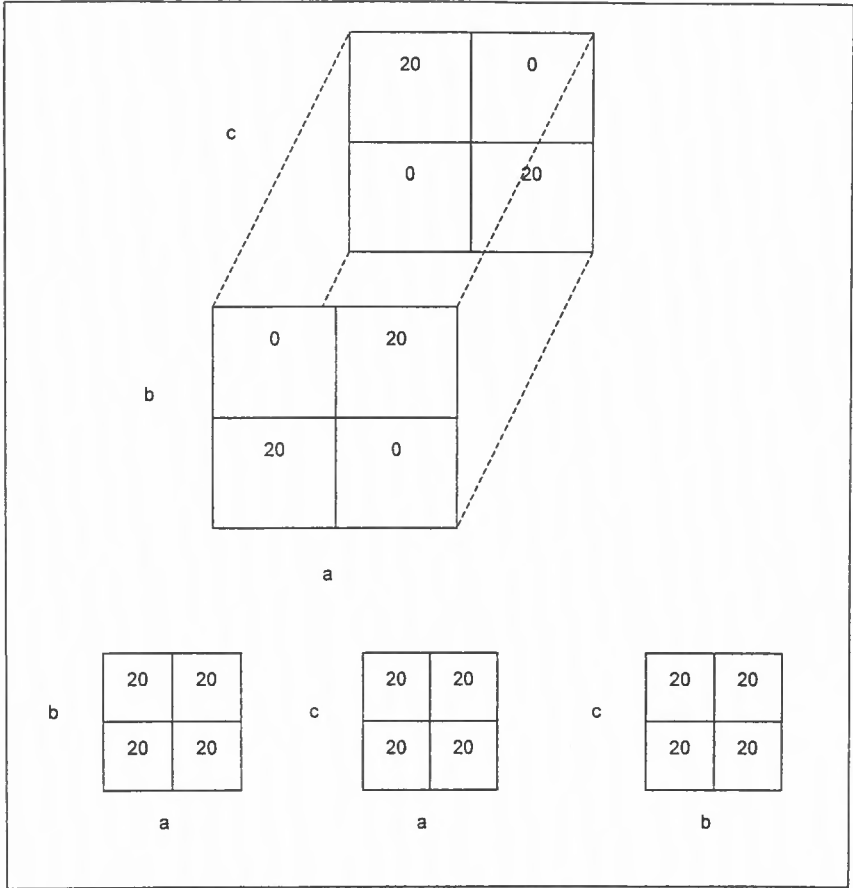


Abb. 2: Beispiel eines trivariaten Zusammenhangs

Der in Abbildung 2 dargestellte Würfel zeigt den Zusammenhang von drei dichotomen Variablen in Form einer dreidimensionalen Kreuztabelle sowie die drei bivariaten Kreuztabellen derselben Variablen. Wie man sieht, zeigt sich in keiner

der bivariaten Kreuztabellen ein Zusammenhang zwischen den beiden beteiligten Variablen, d. h. die Zellen der Kreuztabellen sind gleich häufig besetzt. Demgegenüber drückt der Würfel aus, dass ein sehr starker trivariater Zusammenhang besteht, dieser ist sogar maximal. Vier der acht Zellen des Würfels sind gleich häufig besetzt, die anderen vier weisen Nullhäufigkeiten auf.

Dieses Beispiel soll illustrieren, wie fahrlässig man handelt, wenn man multivariate Zusammenhänge lediglich über die bivariaten Zusammenhänge analysiert, wie dies in der gesamten Korrelationsstatistik passiert. Komplexe multivariate Zusammenhänge können durch Faktorenanalyse oder Strukturgleichungsmodelle gar nicht abgebildet werden, sondern nur ihre bivariaten Projektionen. Dabei bietet die Methodenlehre eine Reihe von Konzepten an, die echte trivariate Zusammenhänge beschreiben.

Der Begriff der *Interaktion* ist wohl der geläufigste. Eine Interaktion ist der Zusammenhang zwischen einer Variablen (der abhängigen Variablen) von der bivariaten Verteilung zweier anderer Variablen (der unabhängigen Variablen). Der Lernerfolg A kann z. B. von der Interaktion der Unterrichtsmethode B und des Geschlechts C abhängen, wenn Mädchen bei der Unterrichtsmethode besser lernen als Jungen.

Während der Begriff der Interaktion symmetrisch ist, beschreibt der Begriff der *Moderatorvariablen* dasselbe Phänomen aus einer asymmetrischen Perspektive. Das Geschlecht ist diejenige Variable, die den Einfluß der Unterrichtsmethode auf den Lernerfolg moderiert, d. h. verändert. Moderatorvariablen können nur identifiziert werden, indem man die Stichprobe nach den Werten der potentiellen Moderatorvariablen teilt und die Analysen in den Teilstichproben getrennt berechnet.

Auch der Begriff der *bedingten Effekte* beschreibt eine spezielle trivariate Interaktion. Im genannten Beispiel wäre z. B. der Effekt unter der Bedingung, dass es sich um Mädchen handelt, positiv, während der bedingte Effekt für die Jungen kleiner oder nicht vorhanden ist.

Man kann sich vorstellen, dass eine derartige intensive Analyse multivariater Verteilungen umso schwieriger wird, je mehr Variablen an ihr beteiligt sind. Bereits bei vier beteiligten Variablen ist die Anzahl möglicher Interaktionen, Moderatoreffekte oder bedingter Effekte beachtlich groß, so dass man sie in der Regel nur findet, wenn man sie theoriegeleitet sucht. Dies ist ein Punkt, wo das Vorliegen einer Theorie nicht nur ein erkenntnistheoretisches Gebot (s. oben), sondern eine ganz praktische Erfordernis darstellt.

Die Unterscheidung zwischen der Rückführung multivariater Zusammenhänge auf ihre bivariaten Zusammenhänge und der Berücksichtigung „echter“ multivariater Zusammenhänge spaltet auch die *multivariaten* Auswertungsmethoden in zwei Gruppen. Der Großteil der gängigen Statistik beruht auf der Korrelations-

statistik und fokussiert die bivariaten Zusammenhänge einer multivariaten Verteilung. Hierzu gehören wie bereits erwähnt, die Faktorenanalyse oder auch die Strukturgleichungsmodelle, welche die moderne Form von Pfadmodellen darstellen. Demgegenüber gibt es aber auch Auswertungsmethoden, die die gesamte Information über die multivariaten Zusammenhänge ausschöpfen. Diese Verfahren nennt man Full-Information-Methods. Hierzu gehört z. B. die Latent-Class-Analyse, mittels derer sich die multivariaten Zusammenhänge kategorialer Variablen beschreiben lassen. Die Latent-Class-Analyse sucht nach Mustern von Variablenausprägungen, die besonders häufig auftreten. Sie läßt sich daher gut zur Exploration multivariater Zusammenhänge kategorialer Daten nutzen, ohne dass sie eine Theorie voraussetzt, welche Muster zu erwarten sind.

Dieser Streifzug durch statistische Auswertungsmethoden sollte deutlich machen, dass die bewährten statistischen Verfahren nicht nur eingesetzt werden können, wenn man bestimmte Verteilungseigenschaften erwartet (d. h. Hypothesen hat), sondern auch, wenn man „nur“ die Verteilungseigenschaften verschiedener Variablen beschreiben will (d. h. qualitative Fragestellungen hat). Der Unterschied besteht darin, dass man im ersten Fall gezielt bestimmte Verteilungsparameter berechnet, um diese dann mit den Methoden der Inferenzstatistik mit den theoretisch erwarteten Parameterwerten zu vergleichen. Im Fall qualitativer Fragestellungen muss man dagegen sehr viele Verteilungseigenschaften unter die Lupe nehmen, um Phänomene zu entdecken, die zur Theoriebildung beitragen können. Die inferenzstatistische Absicherung der Ergebnisse erübrigt sich zwar, wenn man keine prä-experimentellen Hypothesen hat, jedoch läßt sich auch hier über die Berechnung von Konfidenzintervallen für die Verteilungsparameter beurteilen, inwieweit beobachtete Phänomene stark genug sind, um interpretationswürdig zu sein (s. z. B. Brandstätter, 1999).

## Literatur

- Bortz, J.: *Lehrbuch der empirischen Forschung für Sozialwissenschaftler*. Springer, Berlin 1984.
- Brandstätter, E.: *Confidence Intervals as an Alternative to Significance Testing*. In: *Methods of Psychological Research*. Pabst Science Publishers, www.mpr-online.de, Lengerich 1999.
- Flick, U.; Kardorff, E. v.; Keupp, H.; Rosenstil, L. v. & Wolff, S.: *Handbuch Qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen*. Psychologie Verlags Union, München 1991.
- Holzkamp, K.: *Kritische Psychologie*. Frankfurt 1972.
- Huber, O.: *Das psychologische Experiment: Eine Einführung*. Huber, Bern 1987.
- Kerlinger, F. N.: *Grundlagen der Sozialwissenschaft* (2 Bde.). Beltz, Weinheim 1975.
- Köller, O.; Baumert, J. & Rost, J.: *Zielorientierungen: Ihr typologischer Charakter und ihre Entwicklung im frühen Jugendalter*. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie. 1998, Jg. 30, Heft 3, S.128-138.
- Kriz, J.; Lück, H. E. & Heidbrink, H.: *Wissenschaftstheorie und Erkenntnistheorie. Eine Einführung für Psychologen und Humanwissenschaftler*. Leske & Buddrich. Opladen 1987.

- Kutschera, F. v.: *Wissenschaftstheorie I und II*. München 1972.
- Lamnek, S.: *Qualitative Sozialforschung. Bd. 2. Methoden und Techniken*. Psychologie Verlags Union. München 1989.
- Lazarsfeld, P. F. & Henry, N. W.: *Latent structure analysis*. Houghton Mifflin Co. Boston 1968.
- Lienert, G. A. & Raatz, U.: *Testaufbau und Testanalyse*. 5. Aufl. Beltz. Weinheim 1994.
- Mayring, Ph.: *Einführung in die Qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. 3. Auflage. Psychologie Verlags Union. Weinheim 1996.
- Mayring, Ph.: *Qualitativ orientierte Forschungsmethoden in der Unterrichtswissenschaft: Ein Anwendungsbeispiel aus der Lernstrategieforschung*. In: Unterrichtswissenschaft, 1999, S. 292 - 309.
- Popper, K. R.: *Logik der Forschung*. Wien 1935/1998. 2. Aufl. Tübingen 1966.
- Rost, J.: *Die testdiagnostische Erfassung von Typen*. In: Pawlik (Hrsg.): Bericht über den 39. Kongreß der DGPs in Hamburg. Hogrefe, Göttingen 1995, S. 392 - 398.
- Rost, J.: *Lehrbuch Testtheorie Testkonstruktion*. Huber. Bern 1996.
- Rost, J.: *Drei Thesen zum Konzept qualitativer Forschungsmethoden*. Zeitschrift für Didaktik der Naturwissenschaften, 1998, Jg. 4, Heft 3, S. 35 - 42.
- Rost, J.: *Intelligenz*. In Kempf, Straub & Werbik: Psychologie - Eine Einführung. Grundlagen, Methoden, Forschungsfelder. DTV 1997.
- Sarris, V.: *Methodologische Grundlagen der Experimentalpsychologie*. 1: Erkenntnisgewinnung und Methodik. PVU, München/Basel 1990.
- Strauss, A. & Corbin, J.: *Grounded Theory. Grundlagen Qualitativer Sozialforschung*. Psychologie Verlags Union (engl. Org. 1990). Weinheim 1996.



## **Assessment of Understanding of Science in Elementary School**

### **1. Assessment for Learning**

Three premises guide our thinking about assessment. They concern the purpose, enjoyment, and validity of assessment. Each of these premises is significant for assessment at any level of education, but there are particular aspects of each that are of greater impact with elementary school children. We begin by considering these aspects, and then give examples of methods for the assessment of science learning in elementary school.

Most people would say that the purpose of assessment is to find out what students know about something. That is correct, but limited. It also presumes that the purpose belongs to the assessor. We find that a different view of purpose changes one's ideas about the nature and methods of assessment. We see that the purpose of assessment is to aid the learning of the students, and so give them greater control over their lives. Assessment thus serves the purposes of the assessed. Assessment is for the learner, even more than for the teacher.

What this change in view means in practice is that assessment has to be much more than giving a test and summarising students' performances with number or letter grades. Assessment is more than tests. Assessment has to improve students' learning as well as reveal it, and has to reveal the student's learning to the student as well as to the teacher. This belief is embedded in our principle: Every good assessment procedure is a good teaching procedure and a good learning procedure. We shall try to show through the later examples how this principle is inherent in specific procedures.

### **2. Assessment for enjoyment**

Science should be fun to learn. That is especially important for young children. If early experiences turn them away from science, then they suffer a great loss. Their understanding of the world, and so their control of their lives, will remain weak. Fun in learning does not preclude assessment. Rather, assessment has to support enjoyment of science, not destroy it. This precept has implications for the forms of

assessment. We have to consider what sorts of things children like to do and what they dislike.

This line of thinking leads us to appreciate that assessment must be varied. Children, like adults, like routine but also enjoy variety within that routine. Hence there must be a wide range of methods in assessment. Also, a task that requires children to persist against their will can only create dislike for the topic. The central issue in consideration of how to have children wanting to undertake a task is not the length of time involved in the task but the extent to which they choose to engage intellectually with it. Thus, while it is often a reasonable working principle that assessment tasks with young children should be brief, this is not always appropriate.

Other characteristics that might help assessment to foster enjoyment are provision of a sense of accomplishment, room for self expression, partnership rather than isolation, collaboration rather than competition, and employment of colour. Each is important to the fostering of intellectual engagement with the task.

The specific methods that we describe below are intended to go some way towards meeting these requirements. We are conscious, however, that there is room for much more progress in designing assessment for enjoyment.

### **3. Valid assessment of understanding**

The stock definition in textbooks of validity is that it is the extent to which a test measures whatever it is supposed to measure. By itself, that does not take us far, but it does direct our attention to the outcomes of teaching. If we are going to have valid measures of children's learning of science, we shall have to answer two questions: Why is science taught in primary school? What do we intend the children to learn? These are not trivial questions. To answer them fully would require a book. All that we can state here is that we believe that we should teach science to give students a rich understanding of the natural world, an appreciation of science as a human achievement, and a wish to continue to learn more of it. Assessment should cover all of these outcomes.

Of course, the degrees to which we would aim to attain these objectives are affected by the ages of our students. For elementary school students understanding naturally has to be limited. We can, however, frame our teaching and assessment in ways that will promote understanding, appreciation, and interest to as full an extent as possible given the stage of development of our students.

Although we note the importance of appreciation of science as human achievement, and students' wishes to learn more of science, most of our attention in this chapter focusses on assessment of understanding. Valid assessment must include as complete and accurate a summary of children's understanding as possible. This is

not a simple requirement to meet, for understanding is a complex notion. Philosophers and educationists debate its nature at length. Our view of understanding is that it has numerous facets, and the more of these that are evaluated the more valid the assessment will be.

The most obvious facet of understanding is the extent of the individual's knowledge. By and large, the more that a person knows, the better his or her understanding. Tests should reflect that. This is not quite as simple or as obvious as it might appear. Most tests in common use do not reveal the extent of students' knowledge. Rather, they reveal the proportion that the student knows of the information that the test covers, but tell nothing about any other knowledge that the student might have about the topic. Most tests only sample knowledge. They are valid to the extent that the sampling is fair. Fairness involves an assumption that although students will differ in the amounts that they know, their knowledge will be similar in type. The one who knows more, according to the test, is assumed to know more about all other aspects of the topic that the test does not cover. People's knowledge can be uneven, however, with one person knowing much more about beetles, say, and maybe little about insects. A test on invertebrates that over-emphasises beetles will imply greater understanding than it should.

The characteristic of assessment that matters in regard to the assumption of fairness in sampling is the degree to which responses are open or closed. The extreme of closedness is a multiple choice test, in which the student is restricted to choosing between a small number of pre-determined responses. Such a format cannot reveal knowledge other than that included in the test. There is nothing wrong with a closed format, so long as we appreciate that used alone it will not provide valid assessment. It has to be complemented with more open forms, such as drawings, interviews, concept maps, and question production.

Extent is not the only property of knowledge that determines understanding. Another is the degree to which the learner perceives that the knowledge is interconnected, forming a meaningful whole. Two students might possess much the same amount of knowledge, but differ on this facet of understanding. One might perceive how each element of knowledge links with others, while for the second student the facts are isolated. The more extensive the pattern of links, the greater the understanding. The common forms of test, multiple choice and short answer items, are not as well suited to revealing interlinking as are other forms such as concept maps.

Understanding involves more than verbal knowledge. Though facts are important, the mental pictures we hold for objects, actions, and events are also fundamental to understanding. This is especially important for young children, who have not yet built up the mass of propositions that adults call knowledge. Images are part of knowledge, too. To be valid, assessment must tap students' stores of images. Draw-

ings and physical actions can do this, and so should be part of the teacher's repertoire of methods of assessment. Children also tend to enjoy drawing, modelling, and acting.

#### 4. Examples of innovative forms of assessment

Our consideration of purpose, enjoyment, and validity in assessment leads us to the principle that teachers should use diverse methods to find out how well their pupils understand any subject, science included. Our book, *Probing Understanding* (White & Gunstone, 1992), describes rather fully a range of methods, so we shall not go into full detail here of how to introduce each procedure, or its variations. We provide only a brief outline, with an example, for each method.

##### *POE (or, What happens next?)*

POE stands for Prediction, Observation, Explanation. Students are presented with a situation and are asked to predict what will happen. They must give reasons for their predictions, because the reasons tell more about their understanding than do the predictions themselves. They then watch what happens, and say or write what they see. Their observations also tell much about their understanding. Occasionally some will focus on irrelevant or minor matters, more commonly some will see only what they believe should happen. Finally, if an observation conflicts with a prediction, the student has to explain why this might be so. Again, the explanations reveal something of the student's understanding. Often a student provides reasons for the incorrect prediction that show a good understanding.

POE is a powerful probe of understanding that has been used widely and effectively in secondary and tertiary education. It is, of course, the basis of scientific method in which a hypothesis (prediction) is tested through experiment (observation) and the results are discussed (explanation).

The method has been used in primary teaching, although as yet much less than at other levels. The main thing for the teacher to be clear about is what principle students are expected to understand. Then a situation can be created that involves the principle.

We find it is often helpful to give students a limited range of possibilities for their predictions (e.g. increase, decrease, no change) so that the focus of the prediction can be on the principle underlying the POE and not on less relevant aspects such as the actual magnitude of any change or difference. This can also help focus the students' observations on things relevant to the essential principle.

##### *Example 1*

Principle: black materials absorb heat better than white.

Situation: two blocks of ice are covered with cloth, one cloth black and the other

white. Students are asked to predict what will happen to the ice. (A suitable limited range of predictions would be “ice melts faster under the white cloth”, “ice melts faster under the black cloth”, “ice melts the same under the two cloths”.)

*Example 2*

Principle: A difference in air pressure produces a force.

Situation: Coloured water is sucked up through a straw; the students are asked to predict what will happen when a pin is poked into the straw.

Here the range of predictions that are provided to the students could be “the water comes out of the pin hole,” “the water goes up to the mouth,” and “the water does not come up.”

*Example 3*

Principle: The amount of a material does not affect whether the material will float (this embodies a very common student view that flotation is related to size).

Situation: A ball of plasticine or clay is dropped into water and seen to sink. Then a very small amount of the plasticine is taken of the ball and formed into a ball shape; the students are asked to predict what will happen when the small ball is put on to the water surface.

Leach (1999) and Shepardson, Moje, and Kennard-McClelland (1994) have used POE to assess elementary school children’s understandings of science topics.

*Venn Diagrams*

Venn Diagrams are a sharp probe of understanding of classifications such as liquids, metals, living things, flowers. Here is a Venn diagram for the classes animals, dogs, cats.

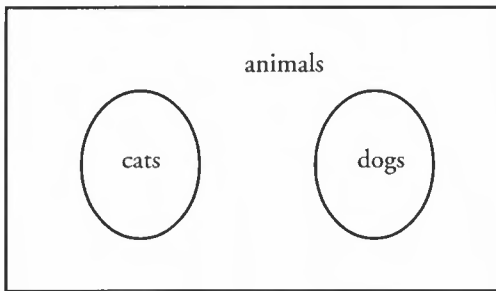


Fig. 1

The task can be extended by also asking children to mark on their diagram an appropriate spot for specific members of the classes, for example in this case “guide dog”, “my cat”.

While Venn diagrams are very useful as an assessment method, it is important to be clear about the specific aspects of understanding you seek to assess. This is well illustrated by the consequences of adding just one class to the above example: pets. The diagram now is much more complex in its demands on children.

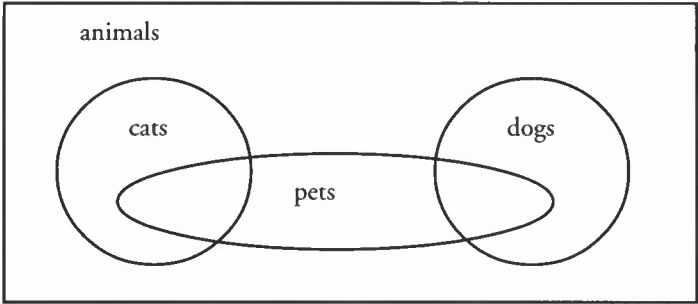


Fig. 2

If the concept of pet is taken to include other than animals, such as plants or inanimate things, then the diagram becomes even more complex.

Once children have been taught about Venn diagrams, the diagrams can be used in several ways to test their understanding of classes. Older children can be asked to draw the diagram, while younger ones might be shown a diagram and asked to name an example from a given section of it, or asked in which section a specific case should go.

Note that since Venn diagrams are about the relation between classes of things, the regions are for *sets* of objects, not individual items. So they always involve plurals.

**Example 1**

Draw a Venn diagram for *plants, trees*.

**Example 2**

Here is a Venn diagram for *things that never were alive, living things, animals and plants*.

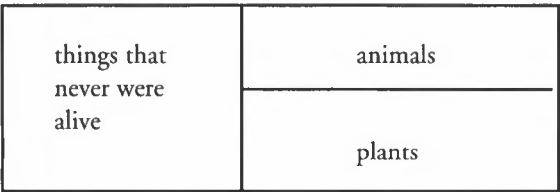


Fig. 3

Write H on the diagram for where *Horse* fits, S where *Seaweed* fits, and W where *Wind* fits.

Other examples of classes of objects suitable for primary school science are animals, mammals, animals with backbones; plants that we eat, plants that have flowers; objects in the solar system, planets, stars.

Unfortunately, researchers have so far missed the opportunity to use Venn diagrams in probing understanding, not just in science and at elementary school level but generally. We can find no relevant examples other than our own early articles: Gunstone and White (1983, 1986).

### *Putting Things Together*

Putting Things Together is a partner of Venn Diagrams. Like Venn Diagrams, it tests children's understanding of the meaning of terms for sets of objects. The procedure is simple: the children are given a collection of objects, which the teacher has chosen to represent things that they have recently been taught about, and they have to place them in groups. When they have finished they can be asked why they chose the groupings.

It is best to use actual objects, especially with younger children, but cards bearing pictures or names of things can be used.

#### *Example 1*

After a lesson on types of animals, the students are given cards with pictures of mammals, reptiles, and birds, and asked to place them in three groups. About five of each type of animal is appropriate.

#### *Example 2*

A class of young children undertaking an investigation of floating is to find out which of a number of given objects floats in water. Drawings of each object are given to the children to be placed in groups according to the results of the investigation.

### *Drawings*

Most children like to draw, and are less inhibited than adults by concerns about the quality of their productions. Drawings are an efficient and revealing method of probing understanding. Efficient, because they do not take children long to produce, and the teacher can inspect a whole class's drawings in a short time.

#### *Example 1*

This drawing was made by a first grade child after a lesson on the water cycle. This involved a demonstration of water coming from a kettle and condensing on a cool plate and then falling as drops. The child has formed the belief that kettles dot the fields.

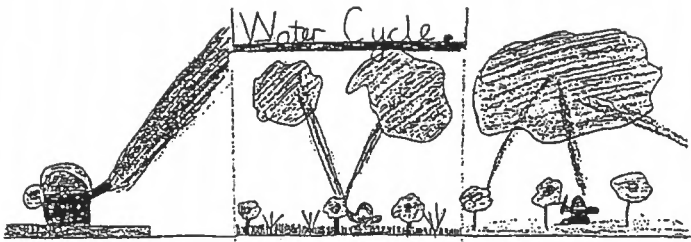


Fig. 4

*Example 2*

Upper primary children who have been taught about body parts can trace the outline of one of their class lying on a large sheet of paper, then sketch in the various organs. Here is a drawing by a fifth grade student, who shows some understanding of the placement of several organs in the body.

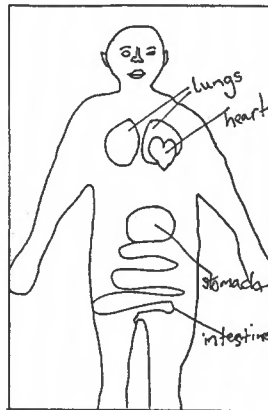


Fig. 5

*Example 3*

Students' understanding of scientists can be tested by asking them to draw a scientist at work. Unfortunately, most students draw a middle-aged to elderly male, often of eccentric appearance. It is, however, a good test of their understanding of scientists.

Research articles by Johnson (1998b), Kibble (1999), and Tunnicliffe and Reiss (1999) demonstrate the power that drawings have to reveal understanding.

*Question Production*

Older students find it difficult to frame questions, probably because most school systems do not encourage them to ask. Question production is, however, an im-



portant skill. Also, the questions that students ask reveal much about their understanding. This procedure should be much more widely used in education at all levels.

The procedure is to show students a phenomenon connected with the topic that they have recently learned about, and to ask them to write a question about it. We recommend write, rather than say, because many students are afraid to ask questions out loud in class in case they look silly. We also have found that training in asking questions is necessary. Some possible approaches to this training and more general structuring of student question asking are in our book *Probing Understanding*. At first students tend to frame questions about trivial or irrelevant features of the situation, but with guidance and experience they become capable of questions that often reveal depths of understanding that other forms of test miss. Examples of appropriate phenomena are evaporation of a liquid, burning of a candle, bouncing of a ball, flowers on a bush or tree.

Here are examples of questions that 10 year old children who had experience in question production asked about some flowers in a vase.

- How do the flowers stay alive?
- Are the flowers still alive?
- Why do you put water in the vase?
- Will the flowers lose their colour?

### *Interviews*

One of the best of all ways of assessing someone's understanding is to interview them. The Socratic method of teaching is a form of whole-class interview. Teachers might also find it useful to interview individual students about their understanding of a topic. Although it may not be practical to interview individually every member of the class, a talk with one or two students will give the teacher an impression of how well the class has grasped the essential ideas and knowledge.

### *Example*

A valuable approach to interviewing, developed and widely used by researchers at the University of Waikato in New Zealand, involves a set of cards containing drawings of situations/events. Each card is shown to the interviewees and they are asked if, in the way they think about it, this is an example of the concept on which the interview is focussed. Some examples of the cards are in Osborne and Freyberg (1985). An interview focussed on the concept of living, that is exploring the ways in which the interviewee understands whether a range of object is living or not, will reveal thinking like "yes, it can move" (9 year old, responding to card showing a drawing of a spider when asked if this is alive), "no, it doesn't drink" (11 year old, cloud), "yes, because it grows taller" (9 year old, tree).

Interviews are now common in research reports for science education. For elementary school examples, see Fleer and Hardy (1993), Johnson (1998a 1998c), McGinn and Roth (1998), Nakhleh and Samarapungavan (1999), and Palmer (1999).

### *Telling Links*

In Telling Links, students are asked to write about the connection between two items from the topic that they have learned about. Like all methods, this one needs explanation to students and some training before they appreciate the sort of link that they should write.

### *Example*

After exploring a number of phenomena involving water and air pressure, such as using a drinking straw, students were asked to link *suction* and *air pressure*. Responses showed a wide range of understanding: “the water goes up the straw because of air pressure not suction”, “suction keeps the card on the glass and the air pressure tries to push it off”.

Other terms that students might be asked to link are:

- rocks, soil
- trees, air
- birds, seeds
- light, sound
- food, energy

We have found no published reports on the use of this technique in elementary science.

## 5. Learning to use new forms of assessment

Most of the methods we described above involve a rather different approach to assessment than will be the norm for many students. This has important implications for the ways such methods are introduced to students. In most of these methods there is a procedure to be learned before the method can create the representation of learning. If students do not understand the procedure, then the representation they create will not indicate their learning of science; it will only indicate that they have not learned how to respond to the task. Therefore it is central to the use of these methods that students are helped to learn the procedure, and to gain some sense of its value. This is no different from other procedures that are more widely used in schooling. Examples are the writing of essays, for which learning about procedure is a gradual and continuous experience over many years of school, and answering multiple choice questions for the first time.

Teachers have to learn as well as students. We have to learn when to use each method, how to introduce it to students, and how to interpret the results. Most of

us are so accustomed to essay tasks and multiple choice questions we have forgotten that we had to learn how to write the task or question and how to score the students' answers. We have probably forgotten that students have to be shown how to write an essay, or how to select an alternative in multiple choice. These things did, however have to be learned and done. In the same way as they once did for familiar tasks, teachers and students have to learn about these innovative methods.

Learning of the innovative methods causes no special difficulties. The procedure in each is clear, and easy to learn. Scoring is more straightforward than it is for essays. Some, such as Venn Diagrams, have a defensible single answer, which can be marked right or wrong as easily as can a multiple choice question. With others, such as Drawings and Question Production, judgements must be made, but this is at least no more difficult than the judgements that have to be made in scoring essays or paragraph-length answers.

The methods should be used in both formative and summative evaluation. Diversity of methods produces more valid assessment than will a single procedure. Of even more importance than validity is the influence assessment has on learning. Students learn from tests, especially those used in summative evaluation, what the teacher values. If their grades derive entirely from multiple choice tests, then they will learn to select the authoritative alternatives from fields of distractors. They will not learn to explain, nor to link topics nor to represent knowledge visually. There is nothing wrong with good multiple choice tests in themselves; it is their over-use that is damaging.

Whether we like it or not, tests influence learning, and therefore the tests that we use must have as full a spread of forms as possible. The innovative methods described here empower teachers to reveal the quality of students' understandings; they also empower students through making their learning comprehensive.

## References

- Fleer, M. & Hardy, T. (1993). *How can we find out what 3 and 4 year olds think? New approaches to eliciting very young children's understandings in science.*
- Gunstone, R. F. & White, R. T. (1983), Testing and teaching with Venn diagrams. *Australian Science Teachers Journal*, 29, No.3, 63-64.
- Gunstone, R. F. & White, R. T. (1986). Assessing understanding by means of Venn diagrams. *Science Education*, 70, 151-158.
- Johnson, P. (1998a). Progression in children's understanding of a 'basic' particle theory: a longitudinal study. *International Journal of Science Education*, 20, 393-412.
- Johnson, P. (1998b). Children's understanding of changes of state involving the gas state, Part 1: Boiling water and the particle theory. *International Journal of Science Education*, 20, 567-583.
- Johnson, P. (1998c). Children's understanding of changes of state involving the gas state, Part 2: Evaporation and condensation below boiling point. *International Journal of Science Education*, 20, 695-709.

- Kibble, B. (1999). How do you picture electricity? *Physics Education*, 34, 226-229.
- Leach, J. (1999). Students' understanding of the co-ordination of theory and evidence in science. *International Journal of Science Education*, 21, 789-806
- McGinn, M. K., & Roth, W-M. (1998). Assessing students' understanding about levers: better test instruments are not enough. *International Journal of Science Education*, 20, 813-832.
- Nakhleh, M. B., & Samarapungavan, A. (1999). Elementary school children's beliefs about matter. *Journal of Research in Science Teaching*, 36, 777-805.
- Osborne, R. & Freyberg, P. (Eds.) (1985). *Learning in science: The implications of children's science*. Auckland, NZ: Heinemann.
- Palmer, D. H. (1999). Exploring the link between students' scientific and non-scientific conceptions. *Science Education*, 83, 639-653.
- Shepardson, D. P., Moje, E. B., & Kennard-McClelland, A. M. (1994). The impact of a science demonstration on children's understandings of air pressure. *Journal of Research in Science Teaching*, 31, 243-258.
- Tunncliffe, S. D. & Reiss, M. J. (1999). Students' understandings about animal skeletons. *International Journal of Science Education*, 21, 1187-1200.
- White, R. & Gunstone, R. (1992). *Probing understanding*. London: Falmer.

**Empirische Befunde  
zum Lehren und Lernen  
im Sachunterricht**



## Naturwissenschaftliches Denken im Grundschulalter Die Koordination von Theorie und Evidenz

### 1. Die Koordination von Theorie und Evidenz

In der traditionellen, an Piaget orientierten Entwicklungspsychologie wurde Kindern im Grundschulalter die Fähigkeit zum naturwissenschaftlichen Denken im Sinne der systematischen Bildung, Prüfung und Revision von Theorien und Hypothesen abgesprochen. Die Fähigkeit, hypothetisch (besser: theoretisch) zu denken, entwickle sich 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).

Die Auffassung, dass die Fähigkeit zum wissenschaftlichen Denken bei präadoleszenten Kindern in gravierender Weise eingeschränkt sei, wurde jedoch bis vor kurzem nicht hinterfragt. Obwohl seit langem bekannt ist, dass Kinder systematische Strategien der Hypothesenprüfung, die sie spontan nicht anwenden (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), wird erst seit kurzem die differenzierte Charakterisierung der Ursprünge des naturwissenschaftlichen Denkens im Kindesalter als Desiderat erkannt (Kuhn & Pearsall 2000).

Da die entwicklungspsychologische Forschung der letzten 20 Jahre grundlegende *Ähnlichkeiten* zwischen dem Denken und Lernen von Kindern und Erwachsenen demonstriert hat, besteht über eine Reihe von Voraussetzungen für wissenschaftliches Denken bei Kindern im Grundschulalter kein Zweifel mehr: Schon lange vor Beginn der Grundschulzeit können Kinder kausale Schlüsse nach den gleichen Prinzipien ziehen wie Erwachsene (Bullock, Gelman & Baillargeon 1982); insbesondere können Grundschüler Kovariationsinformationen<sup>1</sup> zur Bildung bzw. Revi-

---

<sup>1</sup> d. h. Information über die gemeinsame Veränderung zweier oder mehrerer Variablen

sion von Kausalhypothesen auswerten (Siegler 1976). Ihr Defizit scheint also nicht in der Fähigkeit zur *Nutzung* von empirischer Evidenz zur Erklärung von Phänomenen der Welt zu liegen. Wissenschaftliches Denken (im Sinne einer systematischen, reflektierten Bildung und Prüfung von Hypothesen und Theorien über reale Phänomene) erfordert jedoch die *metakognitive* Fähigkeit, die eigenen Hypothesen und Theorien selbst zum Gegenstand der Reflexion zu machen.

Eine in der neueren entwicklungspsychologischen Literatur einflussreiche Hypothese besagt daher, dass die Defizite von Kindern im wissenschaftlichen Denken nicht auf der Ebene fundamentaler Konzepte (Kausalität) bzw. auf der Ebene des logischen Schließens, sondern auf der *metakognitiven* Ebene zu suchen seien (Carey, Evans, Honda, Jay & Unger 1989; Kuhn 1989; Kuhn, Garcia-Milar, Zohar & Andersen 1995; Sodian, Zaitchik & Carey 1991): Kinder können, so diese Position, empirische Evidenz *nutzen*, um Theorien und Hypothesen zu bilden, haben jedoch Schwierigkeiten, den Theorie-Evidenz-Bezug zu *verstehen*. Das Verständnis des Theorie-Evidenz-Bezugs ist Teil eines umfassenderen epistemologischen Verständnisses, d. h. des Verständnisses des Zustandekommens und der Begründbarkeit (wissenschaftlichen) Wissens. Daher wird das intuitive epistemologische Verständnis des Kindes in der neueren Literatur als Grundlage des naturwissenschaftlichen Denkens betrachtet (Kuhn & Pearsall 2000).

Wichtigstes Merkmal des Verständnisses von Wissenschaft ist nach Kuhn (1989) die Fähigkeit zur Koordination von Theorie und Evidenz, die zunächst eine Differenzierung zwischen beiden voraussetzt. Nach Deanna Kuhns Analyse liegt in dieser Differenzierung die Hauptschwierigkeit von Kindern – und manchen Erwachsenen – beim Verständnis wissenschaftlicher Argumentation: Kinder tendieren dazu, so Kuhn (1989), Theorie und Evidenz zu einer skriptartigen Repräsentation „der Dinge, wie sie nun eben sind“, zu vermischen.

Was bedeutet es, zwischen Theorie und Evidenz zu differenzieren? Erste Voraussetzung für die Differenzierung zwischen Theorien/ Hypothesen einerseits und empirischer Evidenz andererseits ist die Unterscheidung zwischen Überzeugungen (*beliefs*) und der Realität, die Kinder im Alter von etwa vier Jahren meistern (Wimmer & Perner 1983; Perner 1991). Im Altersbereich zwischen etwa vier und sechs Jahren entwickelt sich auch bereits ein explizites Verständnis von Informationsquellen und ihrer Rolle beim Erwerb von Wissen: Während 4-jährige nur direkten Informationszugang (visuelle Wahrnehmung, verbale Mitteilung) kennen, verstehen 6-jährige, dass Wissen auch durch schlussfolgerndes Denken zustande kommen kann (Sodian & Wimmer 1987).

Im gleichen Altersbereich beginnen Kinder auch, korrekte Antworten auf die Frage „Woher weißt du das?“ zu geben, bzw. selbst nach den Quellen des Wissens anderer zu fragen (Gopnik & Graf 1988). Jedoch unterscheiden Vorschulkinder noch nicht systematisch zwischen *Evidenz* für eine Behauptung und *Erklärungen* für das



Zustandekommen eines Handlungsergebnisses: Nach der Evidenz für eine (korrekte) Folgerung gefragt („Woher weißt du, dass x beim Wetlauf gewonnen hat?“), antworten Vorschulkinder häufig mit einer Hypothese über die Ursache des Erfolgs („weil er gute Laufschuhe hat“) nicht mit der Evidenz („weil er den Pokal hält“) (Kuhn & Pearsall 1998). Ob es sich bei dieser Konfundierung um ein robustes begriffliches Problem oder um ein semantisches Missverständnis handelt, ist bisher nicht geklärt.

Belege für die Fähigkeit zur Hypothese/ Evidenz-Differenzierung bei jüngeren Kindern kommen aus einer Reihe neuerer Studien zu Vorläufern des wissenschaftlichen Denkens im Grundschulalter. Ruffman, Perner, Olson & Doherty (1993) zeigten, dass 5- bis 7-jährige Kinder unabhängig von eigenen Überzeugungen aufgrund von aktuell präsentierter Evidenz (einfache Muster von Kovariationsdaten) valide theoriebezogene Folgerungen ziehen können. Sodian, Zaitchik und Carey (1991) fanden, dass Erst- und Zweitklässler zwischen konklusiven (schlüssigen) und nicht konklusiven (nicht schlüssigen) empirischen Tests für eine Hypothese über einen einfachen Sachverhalt („das Erdschwein hat einen guten Geruchssinn“) unterscheiden können und die Wahl eines konklusiven Tests (schwach riechendes Futter in die Erde eingraben und prüfen, ob das Tier es findet oder nicht) korrekt begründen können.

Weitere Studien demonstrieren ein Grundverständnis der Logik der Prüfung von *Kausalhypothesen* (Hypothesen über Ursache-Wirkungs-Zusammenhänge zwischen Variablen) bei älteren Grundschulern: Die Mehrheit der Viertklässler wählte in einer Studie von Bullock und Ziegler (1999) ein kontrolliertes Experiment und konnte ihre Wahl korrekt begründen. Viertklässler können außerdem Befundmuster, die unter einer bestimmten Hypothese zu erwarten und solche, die *nicht* zu erwarten sind, spezifizieren, und geben valide evidenzbezogene Begründungen für kausale Schlüsse aus Kovariationsdaten. Dies wurde in inhaltsarmen Aufgabekontexten (z. B. Hypothesen über die Variablen, die für die Qualität von Tennisschlägern ausschlaggebend sind), wie auch in einer vorkenntnisreichen Wissensdomäne (Attribution von Leistungsergebnissen) gezeigt (Sodian 1992; Schrempf & Sodian 1999).

Diese Befunde deuten auf die Entwicklung wichtiger metabegrifflicher Kompetenzen im Verlauf des Grundschulalters hin. Grundschulkindern können nicht nur (kausale) Inferenzen aus empirischen Daten ziehen, sondern sie besitzen ein Grundverständnis der Relation von Kausalität und Kovariation. Daher scheint die Annahme einer mangelhaften begrifflichen Trennung zwischen Hypothese und Evidenz für das Grundschulalter nicht angemessen zu sein.

Die wenigen bisher vorliegenden Studien zum intuitiven Wissenschaftsverständnis bei Grundschulkindern sind typisch für die erste Phase der kritischen Prüfung einer Defizitbehauptung: Möglichst einfache Aufgaben wurden konstruiert, um zu

prüfen, ob eine globale Defizitannahme (wie die der Vermischung von Hypothese und Evidenz) einer kritischen Prüfung standhält. Die gefundenen Kompetenzen dürfen jedoch nicht im Sinne einer ebenso globalen Kompetenzbehauptung fehlinterpretiert werden. Bisher wurden nur wenige Teilkomponenten des Verständnisses von Wissenschaft an Kindern im Grundschulalter untersucht (vgl. Samarapungavan, 1992).

Umfassendere Analysen liegen für das Sekundarschulalter vor, da bei älteren Kindern und Jugendlichen längere halbstandardisierte Interviews eingesetzt werden können (Carey et al. 1989; Sodian 1992; Sodian, Thoermer, Schrempf & Bullock 1999). In diesen Interviewstudien wurde sowohl das abstrakt-definitorische Verständnis des wissenschaftlichen Erkenntnisprozesses exploriert (Carey et al. 1989; Carey & Smith, 1993) als auch das Verständnis konkreter Beispiele wissenschaftlicher Argumentation (eine Debatte zwischen zwei fiktiven Wissenschaftlern über die Ursachen von Nervosität bei Hunden; eine Hexerei-Theorie vs. eine moderne Theorie der Genese einer Infektionskrankheit). Bei den jüngeren Probanden (12-jährige) überwog ein aktivitätsorientiertes Wissenschaftsverständnis (Wissenschaftler probieren etwas aus, schauen, ob es funktioniert) bzw. ein faktenorientiertes Verständnis von Erkenntnis (Wissenschaftler beobachten, sammeln Fakten); selbst bei Jugendlichen und Erwachsenen wurde nur selten der Theorie-Evidenz-Bezug explizit hergestellt, wenn die meisten Probanden auch theoriebezogene Schlussfolgerungen ableiten und konkurrierende theoretische Perspektiven nachvollziehen konnten.

In der vorliegenden Studie wird erstmals versucht, ein Interview zum Wissenschaftsverständnis mit Grundschulern (der vierten Klassenstufe) durchzuführen, um zu einer differenzierten Charakterisierung des Theorie-Evidenz-Verständnisses (bzw. möglicher Misskonzepte) zu kommen. Die Kinder erhielten die gekürzte und vereinfachte Version eines der Interviews, die bereits an älteren Schülern und Erwachsenen eingesetzt wurden (Sodian et al. 1999) (siehe Anhang). Es handelt sich um ein fiktives Beispiel einer wissenschaftlichen Kontroverse, einer Auseinandersetzung zwischen einem Anlage- und einem Umwelttheoretiker über Nervosität bei Hunden. Beide Wissenschaftler führen ein Experiment durch, um aus ihrer Theorie abgeleitete Hypothesen zu testen – der Anlagetheoretiker versucht, das vermutete somatische Defizit durch Medikation zu kurieren, der Umwelttheoretiker erteilt den Hundebesitzern Unterricht im Umgang mit ihrem Hund. Geprüft wird, (a) ob die Probanden die Logik der Konstruktion der Experimente und der Interpretation der Daten verstehen, (b) ob sie zwischen Interpretationen, die durch Evidenz gestützt sind, und reinen Plausibilitätsargumenten unterscheiden können und (c) wie sie ein fiktives Streitgespräch zwischen den Wissenschaftlern verstehen, das im Kern in der Reinterpretation der Befunde des anderen im Sinne der eigenen Theorie besteht. Im Folgenden wird über Teilergebnisse dieser Studie berichtet.

## 2. Eine Pilotstudie zum intuitiven Wissenschaftsverständnis von Grundschulern der vierten Klassenstufe

*Versuchspersonen* waren 27 Kinder, 15 Mädchen und 12 Jungen der vierten Klasse einer Grundschule im Raum Würzburg. Die Kinder wurden von trainierten Interviewern einzeln in einem ruhigen Raum in der Schule getestet. Die Interviews wurden auf Tonband aufgezeichnet und anschließend transkribiert.

Die Kodierung der Antworten erfolgte nach einem *Kodiersystem*, das für die Vorläuferstudien entwickelt worden war und für die Grundschulstudie angepasst wurde. Das Kodiersystem erlaubt eine differenzierte qualitative Analyse verschiedener Typen von Antworten und ist gleichzeitig so konstruiert, dass die Antworten verschiedenen übergeordneten Ebenen des Wissenschaftsverständnisses zugeordnet werden können: dem *Aktivitätsniveau* (Ebene 1a), gekennzeichnet dadurch, dass konkrete Aktivitäten genannt werden („Blutdruck prüfen“), die nicht auf eine Hypothese bezogen sind; die Probanden auf diesem Niveau zeigen kein Verständnis des Testens von Hypothesen und Theorien; dem Niveau der *einfachen Überzeugungen* (Ebene 1b), gekennzeichnet durch ein Verständnis des Sammelns von Informationen und der Interpretation der Theorien der Wissenschaftler als *Meinungen*; auf diesem Niveau wird kein explizites Verständnis der kritischen Prüfung einer Hypothese gezeigt, vielmehr wird auf Fakten bzw. Ereignisse in der Realität verwiesen, die der Meinung des Wissenschaftlers entsprechen. Auf dem Niveau der *Hypothesenprüfung und der Suche nach Erklärungen* (Ebene 2) werden kontrastive<sup>2</sup> Tests zur Prüfung von Hypothesen vorgeschlagen und die Logik der Falsifikationsstrategie verstanden; die Interpretationen der Ergebnisse werden bezogen auf die Theorien der Wissenschaftler, die als konkurrierende Erklärungsversuche für ein Phänomen (bzw. einen Phänomenbereich) verstanden werden. Da viele Antworten nicht eindeutig Ebene 2 zugeordnet werden konnten, wurde ein Niveau des *impliziten* Verständnisses der Hypothesenprüfung und der Suche nach Erklärungen identifiziert (Ebene 1,5), auf dem kontrastive Tests vorgeschlagen und theoretische Interpretationen von Befunden rekonstruiert werden, jedoch der Theorie-Evidenz-Bezug nicht expliziert wird.

## 3. Ergebnisse

- Logik der Hypothesenprüfung (F 1.1; siehe Anhang):

Die Mehrheit der Probanden (69%) konnte Möglichkeiten zur Gewinnung relevanter empirischer Daten zur Hypothesenprüfung vorschlagen, jedoch nannte nur

---

<sup>2</sup> Ein *kontrastiver Test* liegt dann vor, wenn zwei Ausprägungen der in Frage stehenden unabhängigen Variable hinsichtlich ihrer Kovariation mit der abhängigen Variable verglichen werden (z. B. die Nervosität von Hunden (abhängige Variable) mit und ohne Medikation (unabhängige Variable)). Das *kontrollierte Experiment* setzt zusätzlich die Konstanzhaltung aller anderen Variablen voraus.

eine Minderheit (14%) spontan einen kontrastiven Test. Antworten auf dem Aktivitätsniveau waren selten (12%). Ebenfalls konnte eine knappe Mehrheit der Kinder (57%) sowohl zu erwartende als auch nicht zu erwartende Ergebnisse eines Experiments spezifizieren (F 1.3; siehe Anhang). Auf die Frage nach der Notwendigkeit einer Kontrollgruppe (F 1.2a/b; siehe Anhang) gaben 60 % der Kinder richtig an, dass zwei Gruppen untersucht werden müssten. Die Begründungen für diese Antwort offenbarten allerdings bei 60 % dieser Kinder ein Missverständnis: Sie glaubten, die Kontrollgruppe diene als eine Art „Ersatztruppe“, um mehr Information zu bekommen. Nur ein Kind gab die (nicht weiter spezifizierte) Begründung, die Kontrollgruppe diene „zum Vergleich“. Auch auf die Frage nach der Konfundierung von Variablen (Hunderasse und Treatment, F 1.4; siehe Anhang) antwortete die Mehrheit der Probanden (65 %), bei dem konfundierten Experiment handle es sich „nicht um ein gutes Experiment“.

Als Begründung nannten sie jedoch, dass man die Ergebnisse nicht generalisieren könne, weil nur eine Hunderasse untersucht worden sei, bzw. weil es keine objektiven Bedingungen biete. Nur ein Kind zeigte ein explizites Verständnis der Notwendigkeit der Variablenkontrolle. Insgesamt zeigen die Antworten der Kinder in diesem Interviewteil ein oberflächliches Verständnis des Experimentierens: Die überwiegende Mehrheit der Antworten konnte Niveau 1b zugeordnet werden, d. h., Experimente werden als Mittel der Informationsgewinnung verstanden, die Logik des kontrastiven Tests und des kontrollierten Experiments wird jedoch, wenn überhaupt, nur sehr oberflächlich im Sinne eines „Vergleichs“ verstanden.

- Evidenzbezogene Argumentation vs. Plausibilität:

Die Unterscheidung zwischen zulässigen (evidenzbasierten) und nicht zulässigen (nicht evidenzbasierten) Schlussfolgerungen (F 2.1/2/3; siehe Anhang) gelang zwei Dritteln der Kinder (66 %) mindestens in einem Fall: In der Regel wiesen sie eine nicht zulässige Schlussfolgerung („die Besitzer behandeln die Hunde schlecht, weil sie selbst als Kinder schlecht behandelt wurden“) zurück, mit der Begründung, der Wissenschaftler könne das nicht wissen, da er nichts über die Kindheit der Besitzer wisse. Immerhin ein Drittel der Probanden akzeptierte jedoch alle drei Schlussfolgerungen und argumentierte trotz mehrerer leitender Nachfragen durchgehend auf der Plausibilitätsebene. Diese Kinder zeigten eine substantielle Konfundierung von Theorie und Evidenz im Sinne von Kuhn (1989).

- Theoriekonflikt:

Die überwiegende Mehrheit der Kinder konnte die beiden Theorien (Umwelt- vs. Anlagetheorie) unterscheiden und im gegebenen Theorierahmen konforme Schlussfolgerungen ableiten. Dies zeigte sich besonders deutlich auf die Fragen nach den Erklärungen der beiden Wissenschaftler für „neue Befunde“ (F 3.2a/b/c;

siehe Anhang): Nur 16 % der Kinder konnten auf keine der drei Fragen zwei unterschiedliche theoriekonforme Erklärungen ableiten. Allen anderen gelang dies in mindestens einem Fall (z. B.: Wie würden die beiden Wissenschaftler erklären, warum nervöse Hunde auch häufiger nervöse Welpen haben als normale Hunde? – Wissenschaftler „Freundlich“ würde sagen, die haben das nachgemacht, weil die Eltern ja auch nervös sind. Wissenschaftler „Blutrot“ würde sagen, die haben das geerbt, weil ihnen auch ein Stoff im Blut fehlt.“)

Dieses Ergebnis zeigt, dass die beiden Theorien inhaltlich von den Kindern überwiegend verstanden wurden, dass sie ein Grundverständnis theoriekonformer Mechanismen (Lernen, Vererbung) hatten und dass sie zwischen den Theorien differenzieren konnten. Dies ist eine notwendige Voraussetzung für metatheoretisches Verständnis (da fehlendes inhaltliches Verständnis der Theorien bzw. eine globale Konfundierung der Theorien trivialerweise ein Verständnis des Theoriekonflikts verhindern würde).

Die Antworten auf die Frage „Könnten beide Wissenschaftler recht haben? Warum bzw. Warum nicht?“ (F 3.1; siehe Anhang) zeigen, dass es sich um keine hinreichende Bedingung für metatheoretisches Verständnis handelt: Die Mehrheit der Kinder (61 %) argumentierte entweder, es könnten beide Recht haben, da ja beide Experimente funktioniert hatten (beide einen Effekt gefunden hatten) (22 %) oder es könnten beide Recht haben, da es unterschiedliche Arten von nervösen Hunden gebe (39 %), wobei unterstellt wurde, dass unterschiedliche Behandlungsmethoden bei verschiedenen Hunden zum Erfolg führen könnten, da „Hunde eben verschieden seien“ bzw., dass es Hunde geben könne, die wegen falscher Behandlung nervös seien und solche, die wegen somatischer Defizite nervös seien. Nur zwei Versuchspersonen gaben auf diese Frage eine Antwort auf Niveau 1.5, eine argumentierte, beide Wissenschaftler könnten Recht haben, da beide Ursachen interagieren könnten und eine argumentierte, es könne nicht sein, dass beide Recht haben, da sie die Daten völlig unterschiedlich interpretierten. (ca. 30 % der Versuchspersonen gaben keine oder keine interpretierbaren Begründungen für ihre Antworten auf diese Frage).

Diese Befunde sind ein deutlicher Hinweis auf fehlendes metatheoretisches Verständnis bei Grundschulkindern: Trotz guten inhaltlichen Verständnisses scheinen die Kinder die Theorien nicht als konfligierende Wahrheitsbehauptungen zu verstehen, sondern entweder auf Aktivitätsniveau (1a) im Sinne unterschiedlicher Handlungsanweisungen zu interpretieren (wenn beide Handlungsanweisungen erfolgreich waren, haben beide Recht) oder als Aussagen mit eingeschränktem Geltungsbereich (Aussagen über eine Teilgruppe nervöser Hunde, die für diese „wahr“ sein können), nicht als generalisierende Behauptungen.

#### 4. Diskussion und Schlussfolgerungen

Die vorliegende Pilotstudie zeigt zunächst, dass es möglich ist, ein ausführliches halbstandardisiertes Interview zum Wissenschaftsverständnis schon bei älteren Grundschulern einzusetzen. Obwohl die vorliegende Fassung des Interviews noch kürzbar ist, ist eine Mindestlänge von ca. 30 Minuten wohl notwendig, wenn konkurrierende *Theorien* minimaler Komplexität (und nicht nur arbiträre Hypothesen) eingeführt und altersgerecht erläutert werden sollen.

Die Erfahrungen mit Viertklässlern stimmen im Hinblick auf die *inhaltlichen* Verständnisvoraussetzungen der Kinder optimistisch: Die weit überwiegende Mehrheit war imstande, eine simplifizierte Populärversion einer Umwelt- und einer Anlagentheorie in Grundzügen zu verstehen und theoriekonforme Inferenzen zu ziehen. Der Gewinn der vorliegenden Methodik – gegenüber den traditionellen, sehr inhaltsarmen Beispielen – liegt darin, dass das *Theorie-Evidenz-Verständnis* im eigentlichen Sinne untersucht werden kann und nicht lediglich das Verständnis des *Hypothese-Evidenz-Bezugs*.

Die vorliegenden Befunde weisen auf gravierende Defizite im metatheoretischen Verständnis von Grundschulern hin, die weiter exploriert werden sollten: Die Kinder tendierten dazu, Theorien als praktische Handlungsanweisungen oder „persönliche Meinungen“ zu trivialisieren. Trotz ausreichenden inhaltlichen Verständnisses der konkurrierenden Theorien verstanden sie diese nicht als konkurrierende Behauptungen über die Wahrheit. Vielmehr scheinen sie davon auszugehen, dass jeweils „passende“ Evidenz zur Stützung der einen wie der anderen Theorie herangezogen werden kann, ohne dass ein Konflikt auftritt.

Ähnliche Misskonzepte wurden auch bei älteren Schülern und sogar bei manchen Erwachsenen gefunden (Sodian et al. 1999), jedoch zeigte die Mehrheit der Jugendlichen und Erwachsenen ein sophisticatederes metatheoretisches Verständnis: Sie gingen entweder von der Kombination bzw. Interaktion verschiedener Ursachenfaktoren aus und postulierten eine integrative Theorie oder sie zeigten Möglichkeiten zur Gewinnung neuer Evidenz auf. Jedoch zeigten auch ältere Probanden häufig die Tendenz, den Theoriekonflikt zu negieren, indem sie annahmen, dass unterschiedliche Typen von Nervosität bei Hunden existieren und unterschiedliche Probandengruppen untersucht wurden. Die Einsicht, dass ein- und derselbe Datensatz (bzw. ein- und dasselbe Phänomen) aus unterschiedlichen theoretischen Perspektiven unterschiedlich interpretiert werden kann, scheint nicht Teil unserer Common-Sense-Epistemologie zu sein.

Dies ist über den naturwissenschaftlichen Kontext hinaus von Relevanz, da nicht nur für den Bereich (natur-) wissenschaftlicher Theorien, sondern auch in Bezug auf das Verständnis sozial- bzw. gesellschaftlicher Konflikte und „Meinungsverschiedenheiten“, wie z. B. unterschiedliche Geschichtsschreibung über einen Krieg

aus der Sicht der beteiligten Staaten bzw. Kulturen (Kuhn, Amsel & O'Loughlin 1988; Chandler, Boyes & Ball 1990) oder den angenommenen Gründen für Straffälligkeit (Kuhn et al. 1995) ähnliche Defizite im Verständnis interpretationsleitender Kausalsysteme (Theorien) selbst bei älteren Jugendlichen und Erwachsenen gefunden wurden.

Was kann die Schule – schon die Grundschule – tun, um das Verständnis von Wissenschaft zu fördern? Smith, Maclin, Houghton & Hennessey (2000) konnten zeigen, dass die epistemologische Sicht des Lehrers und die damit einhergehende Unterrichtsgestaltung (exploratives, diskursorientiertes Lernen) langfristig selbst bei jungen Sekundarschülern zu einem Verständnis des wissenschaftlichen Erkenntnisprozesses führten, das sonst erst bei deutlich älteren Sekundarschülern gefunden wird (Leach 1999). Andererseits fanden Bullock und Ziegler (1997) in einer umfangreichen Längsschnittstudie kaum interpretierbare Zusammenhänge zwischen Merkmalen der Unterrichtsqualität und dem Abschneiden der Schüler in Aufgaben zum naturwissenschaftlichen Denken. Jedoch handelte es sich um hochinferente Einschätzungen von Variablen der Unterrichtsqualität (z.B. Klassenmanagement), nicht um spezifische Formen der Vermittlung von Wissen im Sachunterricht.

Die Entwicklung von Curricula für die Grundschule, die speziell der Vermittlung von Wissenschaftsverständnis dienen, steht noch am Anfang (Grygier & Kircher 1999). Aus entwicklungspsychologischer Sicht sollte ein wesentliches Ziel solcher Curricula in der Vermittlung der Fähigkeit zur Koordination von Theorie und Evidenz liegen. Curricula, die allzu einseitig Strategien (z. B. der Variablenkontrolle) trainieren, werden diesem Ziel nicht gerecht.

## Literatur

- Bullock, M., Gelman, R. & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedmann (Ed.), *The Development of Psychology of Time* (S. 209-254). New York: Academic Press.
- Bullock, M. & Ziegler, A. (1997). Entwicklung der Intelligenz und des Denkens: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 27-36). Weinheim: Psychologie Verlags Union.
- Bullock, M. & Ziegler, A. (1999). Scientific reasoning: Developmental and individual differences. In F. E. Weinert & W. Schneider (Eds.), *Individual development from 3 to 12. Findings from the Munich Longitudinal Study*. Cambridge: Cambridge University Press.
- 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. *International Journal of Science Education*, 11, 514-529.
- Carey, S. & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, 28, 235-251.
- Case, R. (1974). Structures and strictures: Some functional limitations on the course of cognitive growth. *Cognitive Psychology*, 6, 544-574.

- Chandler, M., Boyes, M. & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology*, 50, 370-395.
- Chen, Z. & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the Control of Variables Strategy. *Child Development*, 70, 1098-1120.
- Chinn, C. A. & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research*, 63, 1-49.
- Gopnik, A. & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development*, 59, 1366-1371.
- Grygier, P. & Kircher, E. (1999). Wie zuverlässig ist unsere Wahrnehmung? In H. Schreier (Hg.), *Nachdenken mit Kindern*. Klinkhardt.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689.
- Kuhn, D., Amsel, E. & O'Loughlin (1988). *The Development of Scientific Thinking Skills*. London: Academic Press.
- Kuhn, D., Garcia-Mila, M., Zohar, A. & Andersen, C. (1995). Strategies of knowledge acquisition. *Society for Research in Child Development Monographs*, 60 (4, serial no. 245).
- Kuhn, D. & Pearsall, S. (1998). *Metacognitive dimensions of knowledge acquisition: Awareness of theory and evidence as sources of one's knowledge*. Unpublished manuscript, Teacher College, Columbia University, New York.
- Kuhn, D. & Pearsall, S. (2000). Developmental origins of scientific thinking. *Journal of Cognition and Development*, 1, 113-129.
- Leach, J. (1999). Students' understanding of the co-ordination of theory and evidence in science. *International Journal of Science Education*, 21, 789-806.
- Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: MIT Press.
- Ruffman, T., Perner, J., Olson, D. & Doherty, M. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis-evidence relation. *Child Development*, 64, 1617-1636.
- Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45, 1-32.
- Schrempf, I. & Sodian, B. (1999). Wissenschaftliches Denken im Grundschulalter. Die Fähigkeit zur Hypothesenprüfung und Evidenzevaluation im Kontext der Attribution von Leistungsergebnissen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31, 67-77.
- Siegler, R. S. (1976). The effects of simple necessity and sufficiency relationships in children's causal inference. *Child Development*, 47, 1058-1063.
- 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. *Cognition and Instruction*, 18, 349-422.
- Sodian, B. (1992). *Bereichsübergreifende Veränderungen in der kognitiven Entwicklung. Das kindliche Verständnis des wissenschaftlichen Denkens*. Universität München: Habilitationsschrift.
- Sodian, B. & Schrempf, I. (1997). Metaconceptual knowledge and the development of scientific reasoning skills. *Paper presented at the Annual Meeting American Educational Research Association, Chicago, IL*.
- Sodian, B., Thoermer, C., Schrempf, I. & Bullock, M. (1999). Understanding the nature of scientific knowledge – what develops from early adolescence to adulthood? *Poster presented at the Biennial Meeting of the European Developmental Society, Spetses, Greece*.
- Sodian, B. & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 58, 424-433.
- Sodian, B., Zaitchik, D. & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, 62, 753-766.
- Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.



# Anhang

Verlauf des Interviews zum wissenschaftlichen Denken am Beispiel des Theoriekonflikts eines Anlage- und eines Umwelttheoretikers zu Gründen für Nervosität bei Hunden

Teil 1: Logik der Hypothesenprüfung

**Phänomen:** Bei allen Hunderassen gibt es „nervöse“ Hunde, die viel bellen und beißen und gegenüber „normalen“ Hunden erhöhten Blutdruck und Pulsfrequenz haben.

**Umwelt-Theoretiker<sup>3</sup> (Herr Freundlich):** Eine falsche Behandlung durch die Besitzer ist Ursache für auffälliges Verhalten und für einen hohen Blutdruck der Hunde.

*F 1.1: Wenn Herr Freundlich prüfen will, ob seine Theorie stimmt, was könnte er tun?*

*F 1.2a: Müsste er eine Gruppe von Hunden untersuchen oder zwei Gruppen? Warum?*

**Hypothese:** Wenn eine falsche Behandlung der Hunde durch die Besitzer die Ursache für die Nervosität ist, sollten sich das Verhalten und Blutdruck/Pulsfrequenz der Hunde bessern, wenn die Besitzer der Hunde trainiert werden.

**Experiment:** 2 Gruppen, in jeder Gruppe gleich viele Hunde unterschiedlicher Rassen; Experimentalgruppe: Besitzertraining; Kontrollgruppe: kein Training

*F 1.3: Was für Ergebnisse hat Herr Freundlich bei seinem Experiment wohl erwartet? Was wäre für ihn ein unerwartetes Ergebnis?*

*F 1.2b: Warum hat er denn in dem Experiment eine Gruppe gehabt, bei der er den Hunden und Besitzern gar nichts beigebracht hat?*

Ein anderer Wissenschaftler führt das gleiche Experiment durch, jedoch sind in der einen Gruppe (Experimentalgruppe) nur Pudeln, in der anderen Gruppe (Kontrollgruppe) nur Dackel.

*F 1.4: Ist das eine gute Idee, das so zu machen? Warum/nicht? Warum muss man die Gruppen vergleichen?*

## Teil 2: Schlussfolgerungen

**Ergebnis:** Verhalten und Blutdruck/Pulsfrequenz haben sich bei der Experimentalgruppe normalisiert, bei der Kontrollgruppe blieben sie unverändert.

Herr Freundlich hat aufgeschrieben, was er jetzt über das Problem denkt:

---

<sup>3</sup> Im Folgenden wird der besseren Lesbarkeit halber immer die männliche Form für beide Wissenschaftler verwendet. Von dem Interview liegen jedoch Parallelversionen mit männlichem und weiblichem Text- und Bildmaterial vor. Die befragten Schülerinnen erhielten die weibliche, die Schüler die männliche Form.

- Ich glaube, dass Hunde ruhiger und freundlicher werden, wenn man den *Besitzern* beibringt, wie man seinen Hund richtig behandelt.
- Ich glaube, dass der Grund dafür, dass die Hunde so nervös sind, ist, dass die Besitzer sie *lange Zeit* nicht richtig behandelt haben.
- Ich glaube, dass der Grund, warum die Besitzer die Hunde nicht richtig behandeln, ist, dass sie *selber als Kinder* nicht richtig behandelt wurden.

*F 2.1/2/3: Zeigt das Experiment von Herrn Freundlich das oder nicht? Warum nicht?*

### Teil 3: Theoriekonflikt

#### Anlagetheorie

**Experiment:** Ein anderer Wissenschaftler, Herr Blutrot, beschäftigt sich ebenfalls mit nervösen Hunden. Er hat ein anderes Experiment gemacht: Gruppe 1 (Experimentalgruppe) erhält Medizin, Gruppe 2 (Kontrollgruppe) erhält keine Medizin. Auch hier normalisieren sich Verhalten und Blutdruck/Pulsfrequenz in der Experimentalgruppe, während die Kontrollgruppe keine Veränderungen aufweist.

**Theorie:** Ein angeborener Hormonmangel bedingt hohen Puls und Blutdruck, was zu auffälligem Verhalten führt. Daher erfolgt eine Behandlung durch ein Medikament, das ein fehlendes Hormon enthält.

Bericht Blutrot:

Das Besitzertraining von Wissenschaftler Freundlich bewirkt ein besseres Fütterungsverhalten und dadurch einen teilweisen Ausgleich des fehlenden Hormons. Die Ursache für Nervosität ist aber das fehlende Hormon, und das kann nur durch Medikamente ausgeglichen werden.

Bericht Freundlich:

Die Besitzer der Hunde in Wissenschaftler Blutrots Experiment mussten zur Verabreichung der Medizin zum Tierarzt, wo sie ein besseres Verhalten gegenüber Hunden „abgeguckt“ haben. Die Ursache für Nervosität ist aber ein Fehlverhalten der Besitzer. Dieses muss verändert werden, um das Verhalten der Hunde zu normalisieren.

*F 3.1: Könnten beide Wissenschaftler gleichzeitig Recht haben? Wie kann es sein, dass sie so unterschiedlicher Meinung sind, beide wissen doch, was bei beiden Experimenten herausgekommen ist?*

#### Neue Befunde zu nervösen Hunden:

- Man hat herausgefunden, dass nervöse Hunde häufiger erkältet sind als normale Hunde.
- Man hat herausgefunden, dass nervöse Familien von ihren Hunden weniger gern gemocht werden als normale Familien.
- Man hat herausgefunden, dass die Jungen (Welpen/Babys) von nervösen Hunden oft auch nervös sind.

*F 3.2a/b/c: Wenn nun die Wissenschaftler Freundlich und Blutrot davon hören, was sagen sie wohl dazu? Würden sie das auf gleiche Weise erklären oder unterschiedlich?*



## **Die Nutzung grafisch-visueller Repräsentationsformen im Sachunterricht**

### **1. Die Bedeutung von Repräsentationsformen für das naturwissenschaftliche Verständnis**

Verstehendes Lernen ist zentraler Bestandteil des mathematischen und naturwissenschaftlichen Unterrichts. Ein Fremdsprachenunterricht, in dem die Schüler fehlerfrei neue Sätze produzieren können, obwohl sie die zugrunde liegenden grammatischen Regeln nicht bewusst verfügbar haben, kann als erfolgreich bezeichnet werden. Ganz anders sieht dies für mathematischen und naturwissenschaftlichen Unterricht aus: Dieser gilt als gescheitert, wenn Schüler lernen, Aufgaben eines bestimmten Formats korrekt zu lösen, ohne die zugrunde liegenden Konzepte verstanden zu haben. Letzteres ist jedoch häufig der Fall. Von der Primarstufe bis zur gymnasialen Oberstufe sind sogenannte Überlebensstrategien von Schülern gut dokumentiert, mit deren Hilfe sie fehlendes Verständnis kompensieren können und damit Leistungen zeigen, denen keine tiefergehenden Kompetenzen zugrunde liegen. Derartige Strategien zu erkennen und ihnen konzeptuelles Verständnis entgegenzusetzen, gehört zu den großen Herausforderungen des erfolgreichen mathematischen und naturwissenschaftlichen Unterrichts. Die kognitive Psychologie kann hierzu u.a. beitragen, indem sie mehr Klarheit in den etwas mysteriösen Begriff des Verstehens bringt und zudem Gesetzmäßigkeiten des verstehenden Lernens aufklärt. Wir möchten im Folgenden einige Ansätze hierzu vorstellen und darüber hinaus Möglichkeiten aufzeigen, wie bereits im Sachunterricht der Grundschule die Voraussetzungen für ein besseres naturwissenschaftliches Verständnis geschaffen werden können.

### **2. Verstehen: Ein Alltagsbegriff in der Wissenschaft**

Jeder kennt das introspektiv zugängliche Gefühl, etwas verstanden zu haben. Ein solcher intuitiver Zugang zu einem Begriff ist für dessen wissenschaftliche Nutzung häufig jedoch eher hinderlich. Dennoch ist es der kognitiven Psychologie

gelingen, durch Explikationen wie „understanding comes from having a good representation“ (Novick & Hmelo 1994, S. 1299) das Konstrukt des Verstehens für den wissenschaftlichen Diskurs brauchbar zu machen. Natürlich wirft dieser Satz streng genommen nur ein neues Problem auf, weil der Begriff der Repräsentation noch mysteriöser ist als der des Verstehens. Ein ganzer Zweig der modernen analytischen Philosophie deckt Widersprüche und Ungereimtheiten seines Gebrauchs auf (vgl. Schumacher 1997).

Dessen ungeachtet hat sich der Begriff der Repräsentation als Arbeitsbegriff für die kognitive Lernpsychologie durchaus bewährt. In einem pragmatischen Sinne kann man von Repräsentation sprechen, wenn über Wahrnehmungseindrücke in unterschiedlichen Sinnesmodalitäten hinaus ein Sachverhalt intern oder extern symbolisch repräsentiert werden kann. Diese symbolische Repräsentation kann mehr oder weniger explizit verfügbar sein. Deshalb ist Verstehen nicht ein Alles-oder-Nichts-Vorgang. Der höchste Grad an Explizitheit ist erreicht, wenn ein Symbolsystem aktiv zur Produktion von kommunizierbarem Wissen genutzt werden kann. Es gibt aber auch implizite Formen des Verstehens, bei denen symbolisch vermitteltes Wissen nachvollzogen, aber noch nicht produziert werden kann. An anderer Stelle (vgl. Stern 1998) wird hervorgehoben, dass es insbesondere für Lehrer wichtig ist, erste Zeichen von implizitem Verstehen zu erkennen.

Naturwissenschaftliches Wissen wird vorwiegend in drei Symbolsystemen repräsentiert: in der mündlichen und schriftlichen Sprache, in formal-mathematischen Darstellungsformen und in räumlich-visuellen Veranschaulichungen. Letztere umfassen analoge Bilder ebenso wie Diagramme und grafische Abbildungen, in denen der Raum zur Abbildung nicht-räumlicher Information genutzt werden kann. Einen Sachverhalt oder einen Begriff zu verstehen heißt, ihn in möglichst vielfältiger Weise in unterschiedlichen Symbolsystemen repräsentieren zu können und zudem in der Lage zu sein, zwischen diesen Repräsentationsformen in flexibler Weise zu wechseln. Nach Larkin (1985) zeigt sich naturwissenschaftliches Verständnis, etwa das Verständnis des Kraftbegriffs, auf drei Ebenen der Wissensrepräsentation: 1) Die alltagsnahe Vorstellung einer konkreten Situation und deren Beschreibung: „Der Baum knickte um, als das Auto dagegen fuhr.“ 2) Die formal-mathematische Beschreibung eines Begriffs: „Kraft = Masse  $\times$  Beschleunigung“. Dieses Wissen resultiert oft aus dem naturwissenschaftlichen Unterricht. 3) Erklärende räumlich-visuelle Veranschaulichungsformen, aus denen das Zusammenwirken unterschiedlicher Größen hervorgeht, wie z.B. aus einem Kräfteparallelogramm.

Da sich Wissen auf Ebene 2) oft nur auf die Kompetenz bezieht, Variablen einer Formel zu manipulieren, ohne diese mit zugrundeliegenden Konzepten verbunden zu haben, können grafisch-visuelle Veranschaulichungen als Brücke zwischen sprachlichen Situationsrepräsentationen und formalem Wissen dienen. Sie füllen die Formeln mit Inhalt und reduzieren die Situationsrepräsentation auf wesentli-

che Strukturen. Experten auf naturwissenschaftlichen Gebieten können Wissen auf Ebene 3) als Denkwerkzeug nutzen, um potenzielle Situationen zu konstruieren. Auf der Grundlage grafisch-visueller Veranschaulichungen kann auch Transfer auf neue Inhaltsgebiete ermöglicht werden, wenn sich Interpretationsprinzipien übertragen lassen. So kann der Graph einer Funktion in einem kartesischen Koordinatensystem sowohl Verhältnisse aus extensiven Größen wie Saftmischungen abbilden, die sich auf dieselbe Maßeinheit beziehen (z.B. Anzahl an Gläsern), wie auch Verhältnisse aus intensiven Größen wie Geschwindigkeit, Dichte oder Stückpreis, die sich aus unterschiedlichen Maßeinheiten zusammensetzen (z.B. Kilometer pro Stunde).

### **3. Warum ermöglichen und fördern grafisch-visuelle Veranschaulichungen das naturwissenschaftliche Verständnis?**

Seit mehreren Jahrtausenden nutzen Menschen die Schriftsprache sowie formal-mathematische Symbole. Die räumlich-visuelle Veranschaulichung nicht-räumlicher Information hingegen wird noch nicht lange praktiziert, nimmt aber dennoch eine zentrale Rolle insbesondere in den Naturwissenschaften ein, was beispielsweise aus der häufigen Verwendung von Graphen und Diagrammen in naturwissenschaftlichen Lehrbüchern ersichtlich ist. Grafisch-visuelle Veranschaulichungen sind hier das zentrale Werkzeug zur Darstellung von Sachverhalten, da Sprache und formale mathematische Symbolsysteme nicht ausreichen, um wichtige naturwissenschaftliche Konzepte zu vermitteln. Der Grund hierfür ist, dass zwar die Auswirkungen naturwissenschaftlicher Konzepte sinnlich zugänglich sind, die Konzepte selbst jedoch nicht. So können wir durchaus die Auswirkungen wahrnehmen, die sich ergeben, wenn ein Objekt eine Kraft auf ein anderes Objekt ausübt. Der Kraftbegriff selbst ist jedoch ein reines Konstrukt, das erst im Rahmen einer größeren physikalischen Theorie bedeutungsvoll wird. Kraft und Gegenkraft, zentrale Begriffe der Newtonschen Mechanik, sind unserer Wahrnehmung in vielen Alltagssituationen unzugänglich, können aber durch Pfeile so veranschaulicht werden, dass sich nicht nur ein Zugang zur Mechanik ergibt, sondern sich infolgedessen auch unser Alltagsverständnis des Kraftbegriffes ändern kann. Die visuelle Veranschaulichung führt also zur Ausdifferenzierung des Schwerkraftkonzeptes und kann so Kinder dazu bringen, naive Vorstellungen von physikalischen Zusammenhängen aufzugeben (vgl. Hardy 1999).

Andere zentrale physikalische Konzepte wie z. B. Dichte oder Geschwindigkeit können auch von Laien rein mathematisch verstanden werden, sobald sie die Bruchrechnung beherrschen. Zieht man aber beispielsweise das Konzept der Dichte zur Erklärung komplexer situativer Zusammenhänge heran, etwa im Kontext von Schwimmen und Sinken, ist eine rein mathematische Verfügbarekeit des Dicht-

tebegriffs ohne Einbindung in ein konzeptuelles Situationsverständnis nicht ausreichend. Wird die unterschiedliche Dichte von Körpern und Flüssigkeiten hingegen mit Hilfe einer Punktematrix veranschaulicht, kann man damit verdeutlichen, warum manche Gegenstände schwimmen und andere sinken.

Teil des Repertoires von Naturwissenschaftlern ist es also, den Raum zur Abbildung nicht-räumlicher Information zu nutzen und Visualisierungen als Problemlösungs- und Kommunikationsmedien einzusetzen. Dieser für Experten der Naturwissenschaften selbstverständliche, flexible Umgang mit räumlich-visuellen Darstellungsformen kann und sollte schon im Grundschulalter gefördert werden. Hier setzt das am Max-Planck-Institut für Bildungsforschung durchgeführte ENTERPRISE-Projekt an (Enhancing KNowledge Transfer and Efficient Reasoning by Practicing Representation In Science Education), das im Folgenden näher beschrieben wird.

#### **4. Kompetenzen und Defizite im Umgang mit grafisch-visuellen Veranschaulichungen**

Im ENTERPRISE-Projekt gehen wir in Anlehnung an den Entwicklungspsychologen Vygotsky davon aus, dass sich geistiger Fortschritt in einer veränderten Nutzung von Symbolsystemen zeigt: Symbole werden nicht ausschließlich zur Bezeichnung der sinnlich wahrnehmbaren Welt genutzt, sondern können nach bestimmten Regeln unabhängig von der konkreten Welt modifiziert werden. Bestimmte Begriffe lassen sich ausschließlich durch reine Manipulation auf der Ebene der Symbole erklären. Der Begriff der Endlichkeit lässt sich konkret erlebten Situationen zuordnen, der Begriff der Unendlichkeit hingegen entstand aus der Möglichkeit, das Präfix „un“ einem Begriff voranzustellen und damit einen Gegenbegriff zu entwickeln, der nicht zwangsläufig eine Referenz zur externen Welt hat. Ganz generell lassen sich mit der Möglichkeit, auf der Ebene der Symbole Manipulationen vorzunehmen, die in der realen Welt nicht möglich sind, neue, insbesondere wissenschaftliche, Bedeutungen und Konstrukte schaffen. Aus dieser Annahme ergibt sich, dass anspruchsvolles wissenschaftliches Denken vor allen Dingen an Kompetenzen im Umgang mit Symbolsystemen gebunden ist und dass diese in der Schule in verschiedensten Inhaltsgebieten erworben werden müssen. Für sprachliche und mathematische Zeichensysteme wird dies bekanntermaßen ausgiebig praktiziert. Die Konstruktion und Nutzung grafisch-visueller Veranschaulichungen, die zentrale Repräsentationsform naturwissenschaftlicher Inhalte also, wird hingegen in der Schule nur selten aktiv gefördert. Selbst mathematische Textaufgaben, die sich hervorragend grafisch-visuell veranschaulichen lassen, werden in der Schule selten mit Hilfe solcher Repräsentationsformen erläutert (vgl. dazu: Stern & Staub, 2000). Herauszufinden, wie der aktive und passive Umgang mit grafisch-visuellen



Veranschaulichungen in den Schulunterricht integriert werden kann, gehört zu den Herausforderungen der gegenwärtigen Lehr-Lern-Forschung.

Generell gilt, dass mit Hilfe von Symbolsystemen neue Bedeutungen konstruiert werden können, wenn Regeln (Möglichkeiten und Einschränkungen) für den Umgang mit ihnen bekannt sind. Je mehr Einschränkungen man bei der Nutzung von Symbolsystemen in Kauf nimmt, um so mehr Möglichkeiten hat man, komplexe Bedeutungen darzustellen. Dieser etwas paradox klingende Satz soll am Beispiel des Pfeils als grafisch-visuelles Symbol erläutert werden. Werden Pfeile als Hinweiszeichen z. B. auf einer Landkarte genutzt, unterliegt man keinerlei Einschränkungen hinsichtlich Ausrichtung und Armlänge des Pfeils. Es muss lediglich sichergestellt sein, dass die Pfeilspitze auf den gewünschten Punkt zeigt. Der Preis für die Freiheiten bei der Gestaltung des Pfeils ist der eher geringe Informationsgehalt. Wird ein Pfeil in einem komplexeren inhaltlichen Kontext genutzt, um eine Richtung zu markieren, wie z. B. beim Verkehrsschild „Einbahnstraße“, hat man bei der Ausrichtung des Pfeilarms keine Freiheiten, bei der Gestaltung der Länge des Pfeilarms hingegen schon. Wird der Pfeil als Vektor z. B. im Kräfteparallelogramm genutzt, ist man bei Ausrichtung und Länge des Pfeilarms an inhaltlich sinnvolle Vorgaben gebunden und hat deshalb keine Freiheiten bei der Gestaltung dieser Elemente. Belohnt wird man mit der Möglichkeit, Situationen zu beschreiben, die strukturell komplex sind.

Aus dem bisher Gesagten lässt sich die Forderung nach einer verstärkten Förderung grafisch-visueller Kompetenzen im naturwissenschaftlichen Unterricht ableiten, um so bei Schülern ein besseres Verständnis naturwissenschaftlicher Konzepte zu erreichen. In der Grundschule kommen dafür u.a. der Mathematikunterricht und der Sachunterricht in Frage. Im ENTERPRISE-Projekt steht letzterer im Mittelpunkt des Interesses. Wir untersuchen, wie Inhalte des Sachunterrichts so vermittelt werden können, dass gleichzeitig grafisch-visuelle Kompetenzen erworben werden, die im naturwissenschaftlichen Unterricht der späteren Sekundarstufe einsetzbar sind.

## 5. Studien aus dem ENTERPRISE-Projekt

Das ENTERPRISE-Projekt hat jeweils einen Schwerpunkt in der Sekundarstufe und in der Grundschule. Studien mit Gymnasiastinnen und Gymnasiasten der Sekundarstufe machen deutlich, dass ein Nachholbedarf im aktiven und passiven Umgang mit grafisch-visuellen Veranschaulichungen sogar für die Darstellung inhaltlich wenig komplexer Sachverhalte besteht. Beispielsweise zeigten Untersuchungen von Catrin Rode mit Schülern der gymnasialen Klassenstufen 7 bis 10, dass selbst bei sehr einfach zu verstehenden Inhalten Schüler und insbesondere auch Schülerinnen nur sehr selten spontan die Möglichkeit nutzten, durch lineare Anordnungen, Baumdiagramme, Matrizen oder andere Repräsentationsformen die in einem

kurzen Text gegebene Information auf das Wesentliche zu reduzieren. Dass ein grundlegendes Verständnis der Eigenschaften grafisch-visueller Repräsentationen eine wichtige Voraussetzung zur Förderung tiefergehenden naturwissenschaftlichen Verständnisses ist, zeigte auch Elwin Savelsbergh. Für Schüler der Sekundarstufe entwickelte er grafisch-visuelle Veranschaulichungen zur Vermittlung von Grundkonzepten der Mechanik durch Vektoren. Es stellte sich jedoch heraus, dass erst Schüler der neunten Klasse, die schon prinzipiell mit der Funktion von symbolischen Darstellungen vertraut waren, diese Hilfen nutzen konnten, während die jüngeren Gymnasiasten nicht die hierzu nötigen grafisch-visuellen Kompetenzen mitbrachten.

Es stellt sich also die Frage, ob und wie grafisch-visuelle Fähigkeiten schon im Grundschulalter so gefördert werden können, dass sie einerseits das Verständnis naturwissenschaftlicher Grundkonzepte unterstützen und andererseits als grundlegende Werkzeuge zum Problemlösen in unterschiedlichen wissenschaftlichen Inhaltsgebieten herangezogen werden können. Trainingsmöglichkeiten für Grundschul Kinder werden im Folgenden am Beispiel von zwei Untersuchungen dargestellt. Beide Studien befassen sich mit der Förderung zweidimensionalen Verständnisses von Viertklässlern durch den Einsatz unterschiedlicher Darstellungsformen. Uns interessiert dabei insbesondere, wie schon im Grundschulalter ein Verständnis kartesischer Koordinatensysteme systematisch entwickelt werden kann, da diese durch ihre vielseitige Einsetzbarkeit ein hohes Transferpotential für die Vermittlung anspruchsvollerer naturwissenschaftlicher Konzepte in der Sekundarstufe aufweisen.

### **5.1 Studie 1: Effekte unterschiedlicher Repräsentationsformen auf das proportionale Verständnis**

Susanne Koerber ging im Rahmen ihres Dissertationsprojekts (vgl. Koerber 2000) der Frage nach, wie sich drei unterschiedlich kontextualisierte Repräsentationsformen auf den Erwerb und den Transfer eines proportionalen Verständnisses auswirken. Ausgangspunkt dieser Untersuchung ist die Annahme, dass sich durch den Umgang mit geeigneten Repräsentationsformen bereits bestehende Misskonzepte überwinden lassen. Im Bereich des proportionalen Denkens gilt es, das sogenannte additive Misskonzept aufzugeben, demzufolge proportionale Beziehungen zwischen zwei Verhältnissen auf der Addition der gleichen Zahl beruhen. Beispielsweise wird oft fälschlich angenommen, dass die Verhältnisse 1:2 und 2:3 proportional zueinander seien, weil die Differenz zwischen 1 und 2 sowie zwischen 2 und 3 gleich ist.

Um die multiplikative Struktur proportionaler Größen zu erkennen, die die Herstellung äquivalenter Verhältnisse (1:2 ist proportional zu 2:4) ermöglicht, muss

deshalb die Wahrnehmung von Verhältnissen zwischen zwei Größen umstrukturiert werden. Repräsentationsformen, welche die Beziehung zwischen zwei Dimensionen explizit visualisieren, sollten also besonders geeignet sein, die Aufmerksamkeit von Kindern in proportionalen Problemsituationen (in dieser Studie: Saftmischungen, bestehend aus einer bestimmten Anzahl Orangen- und Zitronensaftbecher) auf multiplikative Strukturen zu lenken. Kartesische Koordinatensysteme bieten diese Möglichkeit der visuellen Darstellung von proportionalen Verhältnissen in der Steigung einer linearen Funktion. Da die Funktionsweise dieser Repräsentationsform für Grundschul Kinder jedoch nicht spontan einsehbar ist, muss diese zunächst in einem Training erarbeitet werden.

In diesem Trainingsexperiment wurden drei Repräsentationsformen variiert, darunter zwei Formen eines Koordinatensystems mit unterschiedlichem Kontextbezug (siehe Abbildung 1). So wurden im *kontextualisierten Koordinatensystem* Orangen- und Zitronensaftbecher an den entsprechenden Koordinaten abgetragen, und das Mischverhältnis wurde durch einen verschieden kolorierten Hintergrund visualisiert. Dem *konventionellen Koordinatensystem* hingegen fehlten diese Elemente. Als dritte experimentelle Bedingung wurde die *Balkenwaage* eingesetzt, die ein manipulierbares Medium der Darstellung des Verhältnisses zweier Größen ist. Auf jeder Seite des Balkens können die beiden Größen abgetragen werden (in unserem Falle orangegelbe und farblose Schraubenmutter, die auf am jeweiligen Balkenende befindliche Stäbe aufgesteckt werden). Bei zueinander proportionalen Verhältnissen ändert sich die Balance nicht, während sich der Balken bei nicht proportionalen Verhältnissen zu einer Seite neigt. Wenn also bei einem Verhältnis von 1 Mutter (orange): 2 Muttern (gelb) nur jeweils eine Mutter auf jeder Seite hinzugegeben wurde, ist die Balkenwaage nicht mehr im Gleichgewicht.

Erst wenn die Seite mit gelben Muttern auf vier Muttern erweitert wird, bestehen proportionale Verhältnisse, was durch die wiedergewonnene Balance des Balkens sichtbar ist. Es ist zu erwarten, dass die Funktionsweise der Balkenwaage durch ihre physische Manipulierbarkeit und die direkte Erfahrung von Wirkzusammenhängen den Kindern intuitiv zugänglich ist. Außerdem kann, wie beim kontextualisierten Koordinatensystem, ein Bezug zur Problemsituation hergestellt werden, da Mengen durch farbige Schrauben und nicht auf symbolische Weise durch Zahlen und Linien dargestellt werden.

Am Trainingsexperiment nahmen 67 Viertklässler teil, die trainiert wurden, mit einer der drei Repräsentationsformen proportionale Verhältnisse aus Saftmischungen herzustellen. Die Studie umfasste sieben Phasen (2 x 2 Stunden), bestehend aus Vortest und zwei Nachtests, sowie zwei Tests zum Nah- und Ferntransfer (Sandmischungen und Geschwindigkeiten). In den Trainingsphasen lernten die Kinder zunächst die Funktionsweise ihrer Repräsentationsform anhand leichter 1:1-Mischungen kennen. In dieser Phase des Experimentes sollten sie durch das

Herstellen von Mischungen gleichen Geschmacks mit wichtigen Prinzipien der jeweiligen Repräsentationsform vertraut werden, etwa dass Mischungen gleichen Geschmacks auf einer Linie liegen. Die Kinder sollten dann ihre bisherigen Erkenntnisse über Darstellungs- und Funktionsweise der Repräsentationsform dazu nutzen, die multiplikative Struktur von Proportionen zu entdecken.

Die Tests zum proportionalen Denken beinhalteten einerseits Aufgaben, in denen die Verwendung multiplikativer und additiver Strategien getestet wurde (beispielsweise sollte bestimmt werden, welches von vier Mischverhältnissen genauso schmeckt wie eine gegebene Mischung), und andererseits Aufgaben, in denen Kinder schwierige Verhältnisse miteinander vergleichen sollten und die Möglichkeit bestand, die jeweilige Repräsentationsform zur Lösungsfindung einzusetzen (es sollte bestimmt werden, welche von zwei Mischungen mehr nach Orange schmeckt oder ob sie gleich schmecken).

Es zeigte sich, dass dieses nur kurze Training mit Repräsentationsformen das proportionale Verständnis der Viertklässler signifikant verbessern konnte. Der größte Trainingseffekt stellte sich bei der Gruppe ein, die mit der Balkenwaage gearbeitet hatte. Die Kinder lösten einerseits die Aufgaben zum proportionalen Verständnis am erfolgreichsten und setzten zudem ihre Repräsentationsform am häufigsten zur Lösung schwieriger Vergleichsaufgaben ein. Zur Überwindung des sogenannten additiven Misskonzeptes scheinen also sowohl die leichte Interpretierbarkeit als auch der Kontextbezug dieser Repräsentationsform beigetragen zu haben. Dennoch ist es bemerkenswert, dass auch die Gruppen, die mit den abstrakteren Koordinatensystemen trainiert wurden, erfolgreich lernten, ihre Repräsentationsform sowohl bei Aufgaben der Trainingsdomäne als auch in Transferkontexten (wie Geschwindigkeit) einzusetzen. Als besonders beachtenswertes Ergebnis bleibt festzuhalten, dass der Umgang mit Koordinatensystemen, d.h. die sinnvolle und kontextbezogene Interpretation von Steigungen, schon im Grundschulalter erlernt werden kann.

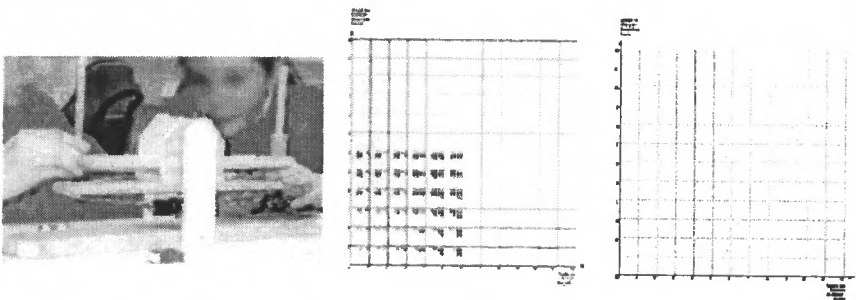


Abb. 1: Balkenwaage, kontextualisiertes Koordinatensystem, konventionelles Koordinatensystem

## 5.2 Studie 2: Effekte unterschiedlicher Repräsentationsaktivitäten auf das Verständnis linearer Zusammenhänge

Auch in Studie 2 von Ilonca Hardy steht das Verständnis von Koordinatensystemen zur Abbildung physikalischer Größen bei Grundschulkindern im Mittelpunkt (vgl. Hardy im Druck). Untersucht wurden hier die Effekte unterschiedlicher Repräsentationsaktivitäten auf 1) das Verständnis linearer Zusammenhänge in Problemlösesituationen und 2) die Interpretation dieser linearen Zusammenhänge in ihrer symbolischen Repräsentation im kartesischen Koordinatensystem. Unter der generellen Annahme, dass Repräsentationsformen zweidimensionales Verständnis fördern, indem sie die mathematische Struktur einer Situation offen legen, kann man davon ausgehen, dass je nach den Eigenschaften einer Repräsentationsform die Wahrnehmung unterschiedlicher mathematischer Strukturen einer Problemsituation fokussiert wird. So können sich Situationen, die lineare Prinzipien beinhalten, dem Lernenden entweder additiv (z. B.  $3+3+3+3=12$ ) oder multiplikativ (z. B.  $3 \times 4=12$ ) darbieten, je nach dem, welches Prinzip im Umgang mit der Repräsentationsform gefördert wurde. Wenn zudem Prinzipien der Interpretation einer Repräsentationsform transferiert werden, können Lernaktivitäten über das mathematische Situationsverständnis hinaus auch den Zugang zu anderen, komplexeren Repräsentationsformen wie Koordinatensystemen erleichtern.

Die in dieser Studie variierten Repräsentationsaktivitäten beinhalteten einerseits den Umgang mit vorgegebenen *Tabellen* und andererseits die Anfertigung von *selbst konstruierten Repräsentationsformen*.

Tabellarische Aktivitäten beruhen einerseits auf dem sukzessiven Addieren von Zahlenpaaren und andererseits auf der expliziten Visualisierung von zwei Dimensionen durch die Aufteilung in Spalten. In Bezug auf die dargestellte lineare Problemsituation sollten sie also die Aufmerksamkeit auf additive Zusammenhänge lenken und wenig fortgeschrittenes proportionales Verständnis aktivieren, wobei jeweils innerhalb eines Verhältnisses Einheiten addiert werden. Ein wirklicher Vergleich zwischen Verhältnissen ist jedoch erst durch multiplikative Strategien möglich. Von Schülern selbst konstruierte Repräsentationen könnten diese Art des proportionalen Verständnisses fördern. Werden Grundschulkindern ohne Vorgaben aufgefordert, lineare Zusammenhänge darzustellen, kann man davon ausgehen, dass alle wahrgenommenen Zusammenhänge bildlich (d. h. wenig abstrahiert) repräsentiert werden. Durch diese bildliche Darstellung von Größen wird jedoch gleichzeitig ein grundlegendes qualitatives Proportionsverständnis aktiviert, das ein Entwicklungspotenzial zur Erkenntnis fortgeschrittener, multiplikativer Zusammenhänge in einer Problemsituation beinhaltet. Allerdings ist nicht zu erwarten, dass die selbst konstruierten Abbildungen auf das Koordinatensystem hinfüh-

ren, während Tabellen durch Transfer ihrer Interpretationsprinzipien den Zugang zum Koordinatensystem erleichtern sollten. Effekte der variierten Repräsentationsaktivitäten wurden also einerseits in der Ausprägung additiver und multiplikativer Problemlösestrategien und andererseits beim Verständnis kartesischer Koordinatensysteme erwartet.

In dieser Studie mit 48 Viertklässlern waren sowohl die tabellarischen als auch die selbst konstruierten Repräsentationen in den experimentellen Umgang mit einem Tanksystem eingebettet. Dieses System besteht aus zwei skalierten Wassertanks (ca. 50 cm hohe Zylinder), die mit unterschiedlich großen Messbechern (Tassen) gefüllt werden können, d. h. der Zusammenhang zwischen den mathematischen Dimensionen Tassengröße, Einschüttungen und Wasserstand kann aktiv erfahren werden, indem gefüllte Tassen in die Tanks geschüttet werden und der jeweils resultierende Wasserstand beobachtet wird. Die zwei Dimensionen, die die Kinder aufgrund des Umgangs mit dem Tanksystem entweder tabellarisch oder in Eigenkonstruktion repräsentieren sollten, waren also die Anzahl an Einschüttungen mit einer bestimmten Tasse und der jeweils resultierende Wasserstand (siehe Abbildung 2 für ein Beispiel einer selbst konstruierten Repräsentation, die das dreimalige Einschütten mit einer „Zweiertasse“ zeigt).

Es zeigte sich, dass die Tabellengruppe zur Lösung linearer Aufgaben häufiger additive Strategien heranzog, während die Gruppe, die selbst Repräsentationen produzierte, multiplikative Strategien nutzte, was auf ein fortgeschritteneres proportionales Verständnis schließen ließ. Bei der Interpretation kartesischer Koordinatensysteme jedoch zeigte sich der umgekehrte Trend. Hier transferierten die Kinder der Tabellengruppe erfolgreich ihr Verständnis der Relation zweier Dimensionen auf die Steigung des Graphen als der Tassengröße, während dies Kindern mit selbst konstruierten Repräsentationen weniger gut gelang. Während 50% der Kinder der Tabellengruppe die Beziehung zwischen Einschüttungen und Wasserstand bei den vorgegebenen Aufgaben korrekt im Graphen repräsentieren und damit beide Dimensionen integrieren konnten (zweidimensionales Verständnis), gelang dies der Mehrheit der Kinder mit selbst konstruierten Repräsentationen nur manchmal (inkonsistentes Verständnis). In beiden Gruppen war der Anteil der Kinder, die sich ausschließlich auf eine Dimension konzentrierten (eindimensionales Verständnis), nur gering (siehe Abbildung 3).

Zusammenfassend zeigen diese Ergebnisse, dass der Zugang zu Koordinatensystemen zwar mit tabellarischen Aktivitäten gefördert werden kann, dass aber der erfolgreichen Interpretation von Funktionsgraphen nicht grundsätzlich ein tiefergehendes mathematisches Verständnis von linearen Beziehungen zugrunde liegt. Gleichzeitig weisen unsere Ergebnisse auf das Potenzial von selbst konstruierten Repräsentationsformen für die Förderung eines generalisierten zweidimensionalen Verständnisses bereits bei Grundschulkindern hin.

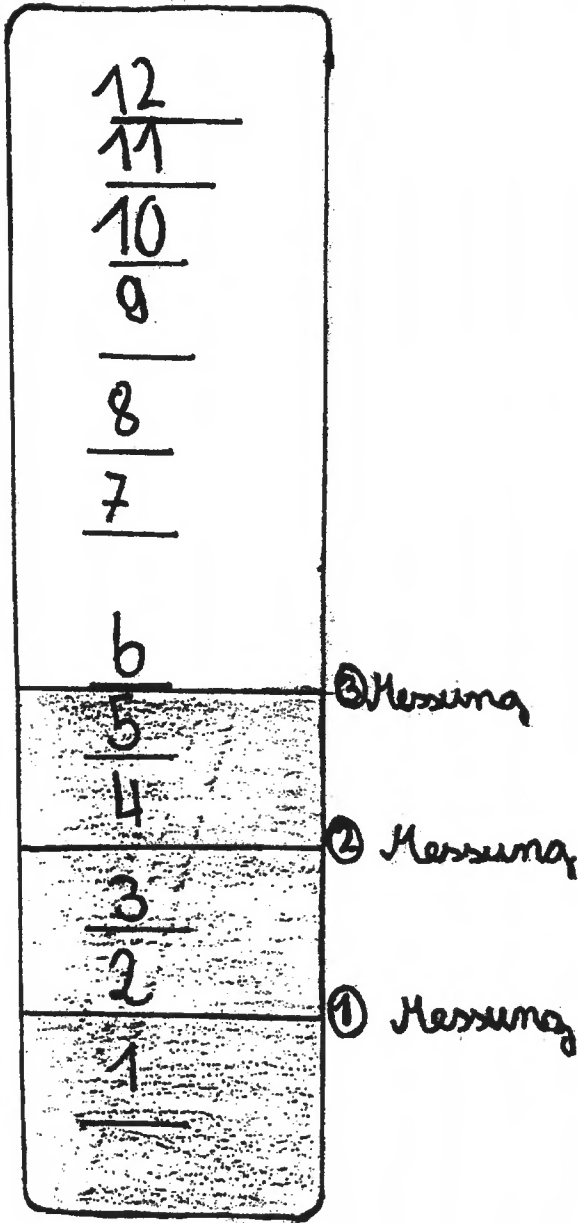


Abb. 2: Selbst konstruierte Repräsentation

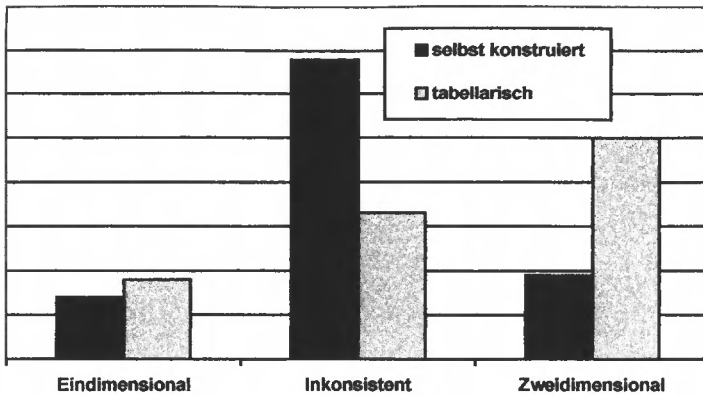


Abb. 3: Verständnis grafischer Darstellungsformen

## 6. Relevanz für den Sachunterricht der Grundschule

Die Ergebnisse beider Studien zeigen, dass mit Hilfe von Repräsentationsformen zweidimensionales Denken bereits in der Grundschule erworben werden kann. Zwar handelt es sich bei unseren Trainingsprogrammen nicht um Unterrichtseinheiten, doch könnten alle von uns verwendeten Materialien in den Sachunterricht eingebunden werden. Unsere Ergebnisse legen nahe, dass Grundschul Kinder zu einem flexiblen Umgang mit der Balkenwaage und sogar mit dem kartesischen Koordinatensystem fähig sind, wenn durch die Unterrichtsgestaltung eine sinnvolle Einbindung der repräsentierten Konzepte in das Situationsverständnis der Kinder gelingt.

Hierbei muss beachtet werden, dass Grundschüler zunächst die nötigen Kompetenzen für den Umgang mit Repräsentationswerkzeugen und ihre Interpretation erwerben müssen, bevor sie diese auf neue inhaltliche Konzepte anwenden können. Dabei ist die Einbettung der strukturellen Eigenschaften einer Darstellungsform in einen den Kindern bereits bekannten Kontext entscheidend. Beim Training mit Saftmischungen wurde beispielsweise damit begonnen, die Eigenschaften der unterschiedlichen Repräsentationsformen mit für die Kinder leicht verständlichen 1:1-Mischungen zu entdecken, bevor dieses Verständnis des Mediums dann wiederum dazu diente, das additive Misskonzept zu überwinden. Obwohl das Koordinatensystem Grundschulkindern kaum intuitiv zugänglich ist, konnte durch eine Integration dieser Darstellungsform in ein kindliches Situationsverständnis (sei es durch vorheriges Experimentieren am Tanksystem oder durch einfaches Proportionalitätsverständnis bei 1:1-Mischungen) ein grundlegendes Verständnis seiner Funktionsweise erreicht werden.



Sowohl die Balkenwaage als auch das Koordinatensystem sind geeignete Mittel zum Experimentieren mit Verhältnissen zwischen Größen. Sie können neben den hier beschriebenen Aufgaben auch zur Veranschaulichung von zentralen physikalischen Konzepten wie Dichte, Druck und Geschwindigkeit herangezogen werden, so dass das in einem Inhaltsgebiet erworbene Verständnis als Grundlage zum Erschließen weiterer Konzepte dienen kann. Mit der Einbindung von Repräsentationsformen in den Sachunterricht kann so auch eine Verbindung zur Mathematik hergestellt werden, die für das naturwissenschaftliche Verständnis in der Sekundarstufe unabdingbar ist.

## Literatur

- Hardy, I. (in Druck). Repräsentationsaktivitäten zur Förderung proportionalen und graphisch-visuellen Verständnisses im Grundschulalter. Erscheint in: H. Petillon (Hrsg.), Jahrbuch Grundschulforschung, 5.
- Hardy, I. (1999). Wie Kinder lernen: Von der Bedeutung sozio-kultureller Lernerfahrungen für die Denkentwicklung. Grundschulmagazin, 14, H.11, 4-7.
- Koerber, S. (2000). Der Einfluß externer Repräsentationsformen auf proportionales Denken im Grundschulalter. Dissertation, Technische Universität Berlin.
- Larkin, J. H. (1985). Understanding, problem representation and skill in physics. In S. F. Chipman, J. W. Segal & R. Glaser (Eds.), *Thinking and Learning Skills, Volume 2: Research on open questions* (S. 141-159). Hillsdale, NJ: Erlbaum.
- Novick, L. R. & Hmelo, C.E. (1994). Transferring symbolic representations across nonisomorphic problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, H. 6, 1296-1321.
- Schumacher, R. (1997). Philosophische Theorien mentaler Repräsentation. *Deutsche Zeitschrift für Philosophie*, 45, H.5, 785-815.
- Stern, E. (1998). Die Entwicklung des mathematischen Verständnisses im Kindesalter. Lengerich: Pabst Publisher.
- Stern, E. & Staub, F. (2000). Mathematik lernen und verstehen. Anforderungen an die Gestaltung des Mathematikunterrichts. In E. Inckemann, J. Kahlert & A. Speck-Hamdan (Hrsg.), *Sich Lernen leisten. Grundschule vor den Herausforderungen der Wissenschaft* (S.9 0-100). Neuwied: Luchterhand Verlag.



## Nachdenken über physikalische Probleme Eine Untersuchung zum A/K-Wert in der Begegnung von Grundschulern mit physikalischen Phänomenen

### 1. Einführung

Seit einer ganzen Reihe von Jahren beschäftigen wir uns in Kassel mit der Frage, wie Grundschüler physikalischen Phänomenen begegnen. Dabei wurde ein spezielles Versuchsdesign entwickelt, in dem eine Reihe von Grundschulkindern (meist zwei Jungen und zwei Mädchen gleicher Klassenstufe) mit einfachen physikalischen Versuchssphänomenen konfrontiert und unter größtmöglicher Zurückhaltung des Versuchsleiters zu Eingriffen in den Versuchsablauf sowie zu Gesprächen über das Geschehen angeregt wurden (Hagstedt & Spreckelsen 1986). Bei den ausgewählten Phänomenen griffen wir dabei in den meisten Fällen auf Vorschläge aus Kinder-Sachbüchern zurück wie z. B. auf „Spiel – das Wissen schafft“ (Press 1964). Meist wurden etwa fünf Versuche präsentiert und diskutiert, die Sitzungen dauerten zwischen 30 und 60 Minuten. Dabei gehörten vier der Versuche dem gleichen physikalischen Funktionsprinzip (wie etwa Balancieren/Gleichgewicht) an; einer der Versuche war davon unterschieden, so dass den Kindern strukturelle Ähnlichkeiten bzw. Unterschiede zwischen den Versuchen aus eigener Anschauung deutlich werden konnten. Die Sitzungen wurden video-dokumentiert und die Verbalprotokolle unter Zuhilfenahme der Video-Aufzeichnungen ausgewertet. Die Auswertungen erfolgen unter je spezifischen Fragestellungen im Rahmen von wissenschaftlichen Hausarbeiten für die erste Lehramtsprüfung. Eine Übersicht über die bearbeiteten Themen der Hausarbeiten sowie über die Versuche ist an anderer Stelle (Spreckelsen 1998, S. 207 ff.) gegeben.

Im folgenden wollen wir vor allem das von uns verwendete Verfahren der A/K-Wert-Bestimmung erläutern und dabei weniger auf die damit erzielten Untersuchungsergebnisse abheben, die allein schon wegen der geringen Probandenzahlen statistisch nicht repräsentativ sind.

## 2. A/K-Wert

Bei den Auswertungen der Verbalprotokolle stießen wir vor etwa sieben Jahren (Krüger 1993) auf die Möglichkeit, die Kinderäußerungen in einem nicht unbedeutlichen Maße in zwei Kategorien einzuteilen:

- Äußerungen, die sich unmittelbar beschreibend auf das zu beobachtende Versuchsgeschehen beziehen, nennen wir „*konkret*“ (Beispiel: „Die Kerze geht jetzt hoch.“ – Versuch Kerzenschaukel). Die Anzahl solcher konkreten Schüleräußerungen bezeichnen wir mit K.
- Äußerungen, die einen gedanklichen Prozess beim Schüler erkennen lassen, nennen wir „*abstrakt*“ (Beispiel: „Die Kerze geht hoch, weil da Wachs abtropft.“ – Versuch Kerzenschaukel). Die Anzahl solcher abstrakten Schüleräußerungen bezeichnen wir mit A.

Unter dem A/K-Wert verstehen wir nun den Quotienten der Anzahlen der als abstrakt bzw. als konkret einzuschätzenden Schüleräußerungen innerhalb einer bestimmten Video-Text-Passage. Diesen A/K-Wert kann man beispielsweise schülerbezogen ermitteln, d. h. es werden innerhalb einer Videosequenz alle entsprechenden Äußerungen eines bestimmten Probanden zur Quotientenbildung herangezogen. Der Wert lässt sich aber auch versuchsbezogen ermitteln, indem die entsprechenden Äußerungen aller beteiligten Probanden innerhalb eines bestimmten Versuches der Sequenz zur Quotientenbildung herangezogen werden. Er kann beispielsweise auch altersbezogen oder geschlechtsspezifisch berechnet werden.

Nicht alle Schüleräußerungen können den beiden genannten Klassen zugeordnet werden. Sie bleiben dann als „nicht zuordenbar“ für die Bestimmung des A/K-Wertes unberücksichtigt. Dies möge an einem Beispiel aus dem Versuchsgeschehen „Gleichgewichtsakrobat“ erläutert werden, in dessen Verlauf eine der Schülerinnen ausführt: „Und in der anderen (Hand) ist auch was drin. Ja. Da ist in jeder Hand was drin. Auf der Seite ist hier was drinnen, und hier ist auf der Seite was drinnen.“ Diese Äußerung erscheint zunächst als eine rein konkret-anschauliche Schilderung des Gleichgewichtsakrobaten, in dessen beiden Händen je ein Geldstück zur Beschwerung kaum sichtbar eingeklebt ist. Andererseits könnte sich in der angeführten Äußerung durchaus auch die Erkenntnis ausdrücken, dass es die beiden Gewichtsstücke sind, die dafür sorgen, dass der Akrobat auf seiner Nase stehend das Gleichgewicht zu halten in der Lage ist. Dies ist schwierig zu entscheiden, und so haben wir in Zweifelsfällen derartige Äußerungen der Kategorie „nicht zuordenbar“ zugewiesen.

Natürlich kann es von allgemeinem Interesse sein, zu betrachten, welcher Anteil der Gesamt-Äußerungen berücksichtigt werden kann und welcher nicht. Dieser letztere Anteil bleibt in aller Regel unter 25% (vgl. dazu beispielsweise Tab. 1). Damit kommen wir zugleich zu einer grundsätzlichen Problematik bei der Bildung

des A/K-Wertes. Auch wenn die Einteilung der Schüleräußerungen zumeist recht eindeutig ist, so ist die Zuordnung der Schüleräußerungen schwer zu objektivieren, und es bleibt immer wenigstens zum Teil ein Rest an subjektiver Einschätzung durch den jeweiligen Auswertenden. Bei Vergleichsuntersuchungen sollte dabei nach Möglichkeit stets dieselbe Person die Auswertung vornehmen. Im Zweifelsfall wurde dann eine Äußerung auch eher der Kategorie „nicht zuordenbar“ zugewiesen. Wegen der Schwierigkeiten bei der Zuordnung enthält auch die Kategorie der „nicht zuordenbaren“ Äußerungen meist immer noch einen beträchtlichen Anteil von Äußerungen, die sich auf das Versuchsgeschehen beziehen. Man vergleiche dazu beispielsweise die oben mitgeteilte Äußerung zum Gleichgewichtsakrobaten.

### 3. A/K-Werte in den Nordhäuser Untersuchungen von 1993

Im folgenden sollen in einem ersten Beispiel die A/K-Werte aus der Videodokumentation einer Begegnungssituation von vier Grundschulern eines vierten Schuljahres mit physikalischen Phänomenen angeführt werden. Die Untersuchung wurde im Jahre 1993 in einer Grundschule in Nordhausen am Harz durchgeführt, bei der die dortigen Lehrkräfte als Versuchsleiter fungierten.

Zunächst ist die Anzahl der Schüleräußerungen geordnet nach abstrakt-kognitiven (A) und konkret-anschaulichen (K) Äußerungen [N = nicht zuordenbar]; daneben ist der A/K-Wert aufgeführt und zwar zuerst bezogen auf die einzelnen Schüler, sodann bezogen auf die einzelnen Phänomene:

Schüler:	insgesamt	A	K	N	A/K
Markus	45	13	29	3	0,45
Benjamin	69	17	49	3	0,35
Katja	74	29	42	3	0,69
Steffi	110	48	48	14	1,00
Versuche:					
Verhexter Karton	78	36	36	6	1,00
Akrobat	55	23	26	6	0,88
Balancierender Knopf	69	22	42	5	0,52
Kerzenschaukel	51	14	34	3	0,41
Flaschengeist	45	12	30	3	0,40
Insgesamt:	298	107	168	23	0,64

Tab. 1: Nordhausen (Gruppe K)

An dieser Aufstellung werden bereits einige typische Merkmale der Nordhäuser Untersuchung sichtbar:

1. Die A/K-Werte für die einzelnen Schüler zeigen deutlich höhere Werte für die beiden Mädchen im Vergleich zu den beiden Jungen.

Offensichtlich scheinen die Mädchen in ihren Äußerungen eine stärker „nachdenkliche“ Komponente aufzuweisen als die Jungen, was aber möglicherweise auch auf unterschiedliche Formulierungskompetenzen hinweisen könnte.

2. Die A/K-Werte für die einzelnen Versuche sind zwei unterschiedlichen Gruppen zuzuordnen: Die Werte für den „verhexten Karton“ sowie für den „Akrobaten“ liegen deutlich über denjenigen für die anderen drei Versuche.

Auch scheinen die einzelnen Versuche – möglicherweise auch aufgrund unterschiedlicher Präsentationsweisen durch die Versuchsleiterin – die Probanden in unterschiedlichem Maße zu Denkleistungen anzuregen.

In der Nordhäuser Untersuchung wurden nun aber insgesamt vier Schülergruppen beobachtet, drei Gruppen zu je vier Schülerinnen und Schülern aus dem vierten Schuljahr sowie eine Gruppe zu sechs Schülerinnen und Schülern aus dem dritten Schuljahr. Als Versuchsleiterinnen fungierten die jeweiligen Klassenlehrerinnen. So ergab sich die Möglichkeit für zusätzliche Vergleiche, die in einer weiteren wissenschaftlichen Hausarbeit vorgenommen wurden (Günther 1994).

	Schüler	A/K Wert	Schülerin	A/K Wert
3. Schuljahr	Frank	1,34	Anika	1,67
	Thomas	0,53	Antje	0,78
	Tobias	0,70	Nicole	0,83
4. Schuljahr	Markus	0,45	Katja	0,69
	Benjamin	0,35	Steffi	1,00
	Daniel (1)	0,79	Nadine	0,63
	Falk	0,86	Susanne	0,74
	Daniel (2)	1,11	Anne	0,96
	Sebastian	0,91	Patricia	2,08

Tab. 2: Nordhausen (alle Gruppen, schülerbezogen)

Ermittelt man aus den Rohdaten Mittelwerte für alle Jungen und alle Mädchen, so erhält man für den A/K-Wert der Jungen den Wert 0,78, für denjenigen der Mädchen den Wert 1,04. Dabei ist übrigens die Anzahl der berücksichtigten Äußerungen der Jungen mit 464 derjenigen der Mädchen mit 457 nahezu identisch. Für den A/K-Wert aller Probanden der Nordhäuser Untersuchung ergibt sich aus insgesamt 414 als „abstrakt“ und 507 als „konkret“ klassifizierten Äußerungen ein A/K-Wert von 0,82.

Der oben für eine einzelne Gruppe bereits beschriebene Befund, dass die Mädchen einen merklich höheren A/K-Wert aufzuweisen haben, reproduziert sich also bei Ansehung aller vier Gruppen. Um nun eine Abschätzung dafür zu erhalten, wie aussagekräftig die erhaltenen Mittelwerte angesichts der vorhandenen Streuungen sind, betrachten wir den sog. quadratischen Fehler des Mittelwertes (= Streuung, dividiert durch  $\sqrt{n}$ ). Dieser einfach zu berechnende Wert ist – unter Normalverteilungsannahme – ein Maß für den Vertrauensbereich des Testmittelwertes (vgl. Lienert & von Eye 1994, S. 135 ff.). Für die Jungen liefert die Rechnung den Wert 0,11, für die Mädchen 0,17, in der Summe beider Werte also 0,28. Da die beiden oben genannten Mittelwerte um 0,26 auseinanderliegen, ihr Abstand also – wenn auch nur geringfügig – kleiner ist als die Summe der beiden Fehler, so ist die Aussagekraft des Befundes insoweit nur eingeschränkt zu sehen.

Ein entsprechender Befund ergibt sich, wenn man die Ergebnisse bezüglich der einzelnen Versuche betrachtet:

Versuch	A/K-Wert Jungen	A/K-Wert Mädchen	A/K-Wert insgesamt
FG	1,12	1,12	1,12
GA	0,85	1,15	0,95
VK	0,59	1,13	0,85
BK	0,67	0,81	0,72
KS	0,60	0,85	0,71

(FG = Flaschengeist, GA = Gleichgewichtsakrobat, VK = Verhexter Karton, BK = Balancierender Knopf, KS = Kerzenschaukel)

Tab. 3: Nordhausen (alle Gruppen, versuchsbezogen)

Zunächst ersieht man, dass mit einer Ausnahme die A/K-Werte für die Mädchen durchweg über denen für die Jungen liegen. Lediglich der Flaschengeist-Versuch führt zum gleichen A/K-Wert, was darin begründet liegen könnte, dass dieser Ver-

such, der ja als einziger nicht zum sonst gewählten Funktionsprinzip „Balancieren/ Gleichgewicht“ gehört, sondern zum Phänomen der Druckerhöhung von Luft bei Erwärmung, nur wenig Eingriffsmöglichkeiten in das Versuchsgeschehen bietet und von seinem Funktionsprinzip her für die Schüler weniger offensichtlich ist. Nimmt man diesen Versuch aus der Vergleichsbilanz zwischen Jungen und Mädchen heraus, erhält man folgende A/K-Werte für Jungen und Mädchen:

	abstrakt	konkret	A/K-Wert
Jungen	173	259	0,67
Mädchen	174	186	0,94

Tab. 4: Nordhausen (alle Gruppen, ohne Flaschengeist)

Die ermittelten A/K-Werte liegen (wie oben) um 0,27 auseinander. Eine detailliertere Analyse zeigt darüber hinaus einen bedeutsamen Unterschied zwischen Mädchen und Jungen beim Versuch „Verzauberter Karton“. Eine schlüssige Interpretation fehlt hier noch. Möglicherweise „schwätzen“ die Jungen einfach mehr als die Mädchen. Jedenfalls liegt dieser Schluss bei Einsicht in die Tabelle 3 insofern nahe, als die abstrakt zu wertenden Äußerungen hinsichtlich der Geschlechter kaum auseinanderfallen und die Jungen lediglich im konkret zu wertenden Anteil stark dominieren.

Im Übrigen liegen die A/K-Werte jeweils in einem begrenzten Bereich, beginnend mit dem Versuch „Gleichgewichtsakrobat“, fallend bis zum Versuch „Kerzenschaukel“. Dieser Versuch scheint die stärkste Anregungsqualität zu rein phänomenbezogenen Äußerungen zu besitzen, während der Versuch „Flaschengeist“ offensichtlich am stärksten zum „Spekulieren“ Anlass gibt.

#### 4. A/K-Werte in den Bergshäuser Untersuchungen von 1998

Im Sommer 1998 hatten wir die Gelegenheit, die Sequenz unserer Phänomenbegegnungssituationen in Vierergruppen an einer ganzen Schulklasse (2. Schuljahr, 16 Schülerinnen und Schüler) aufzuzeichnen, wobei diesmal die Versuchsleiterinnen bei allen vier Gruppen dieselbe war. So war hinsichtlich der Präsentation der Phänomene davon auszugehen, dass diese in den einzelnen Gruppendurchläufen weitgehend identisch vorgenommen wurde. Dabei hatten wir eine reine Mädchen- und eine reine Jungen-Gruppe sowie zwei gemischte Gruppen für unsere Untersuchung zur Verfügung. Die rechnerische Auswertung führte Claudia Kurz in ihrer Wissenschaftlichen Hausarbeit durch (Kurz 2000).



	A	K	A/K
Lisa	18	27	0,67
Rochelle	30	47	0,64
Isabelle	32	50	0,64
Paulina	4	11	0,36
<b>Gruppe 1</b>	<b>84</b>	<b>135</b>	<b>0,62</b>
Marius	25	22	1,14
Jonas	12	22	0,55
Julien	24	17	1,41
Nico	24	30	0,80
<b>Gruppe 2</b>	<b>85</b>	<b>91</b>	<b>0,93</b>
Fabrizio	37	16	2,31
Florian	32	24	1,33
Nadine	38	38	1,00
Sonja	51	33	1,55
<b>Gruppe 3</b>	<b>158</b>	<b>111</b>	<b>1,42</b>
Alexandra	5	3	1,67
Francesco	19	20	0,95
Patrick	34	33	1,03
John	16	11	1,45
<b>Gruppe 4</b>	<b>74</b>	<b>67</b>	<b>1,10</b>
<b>insgesamt</b>	<b>401</b>	<b>404</b>	<b>0,99</b>

Tab. 5: Gesamtübersicht über die vier Probandengruppen

Im Gesamtergebnis zeigt die Auswertung (siehe Tab. 5) von 805 verwendbaren Schüleraussagen ein ausgeglichenes Verhältnis von konkreten und abstrakten Äußerungen (A/K-Wert 0,99). Betrachtet man dagegen die einzelnen Gruppen, so fällt zunächst das deutlich geringere Abschneiden der reinen Mädchen-Gruppe (A/K-Wert 0,62) und der reinen Jungen-Gruppe (A/K-Wert 0,93) im Vergleich zu den beiden gemischten Gruppen (A/K-Werte 1,42 und 1,10) ins Auge. Die geschlechtsspezifischen Mittelwerte aus den gemischten Gruppen sind erstaunlicherweise bis auf die zweite Nachkommastelle identisch (1,41), woraus man den Schluss ziehen könnte, dass in den gemischten Gruppen erstens sämtliche Kinder profitieren, allerdings die Mädchen (Anwachsen im Mittel von 0,62 auf 1,41) möglicherweise sehr viel stärker als die Jungen (Anwachsen im Mittel von 0,93 auf 1,41). Hier wären weitere gezielte Untersuchungen unbedingt wünschenswert, da die hier dargestellten Ergebnisse keine Repräsentativität beanspruchen können. Grundsätzlich ist aber auch der Unterschied bezüglich der geschlechtsspezifischen Differenz in unserer Nordhäuser Auswertung bemerkenswert. In den gemischten Gruppen der Bergshäuser Untersuchung liegen Mädchen und Jungen (im Mittel) gleichauf, in den getrennten Gruppen haben die Mädchen hier sogar die niedrigeren A/K-Werte aufzuweisen als die Jungen. Auch hier wären angesichts der geringen Probandenzahlen weiterführende Untersuchungen erforderlich.

Insbesondere ist anzumerken, dass sich das Gesamtergebnis der Bergshäuser Auswertung (2. Schuljahr) auf einen A/K-Wert von 0,99 im Vergleich zur Nordhäuser Untersuchung (3./4. Schuljahr) von 0,82 beläuft. Dieser unerwartete Trend zeigt sich auch bei der im folgenden Abschnitt dargestellten Kasseler Vergleichsuntersuchung zwischen dem 2. und 4. Schuljahr.

Auch bei dem Vergleich der erzielten A/K-Werte hinsichtlich der einzelnen Phänomene ergeben sich deutliche Unterschiede zur Nordhäuser Untersuchung, wie die folgende Tabelle (Tab. 6) ausweist. Der Vergleich wird allerdings insofern etwas erschwert, da wir in Bergshausen den Versuch „Flaschengeist“ gegen den ähnlich strukturierten Versuch „Buddelthermometer“ (BT) austauschen mussten, da der Flaschengeist-Versuch der Klasse bereits bekannt war.

Man bemerkt beispielsweise, dass der Kerzenschaukelversuch, der in der Nordhäuser Auswertung an letzter Stelle stand, nunmehr an die Spitze gerückt ist. Als Erklärung wäre denkbar, dass dieser Versuch Kinder des zweiten Schuljahres mehr zum schweifenden Überlegen anregt, während er für Kinder zum Ende der Grundschulzeit direkter verstehbar und damit eher „abgehakt“ erscheint. Hierzu müsste man das Verbalprotokoll noch einmal heranziehen. Diese Vermutung erhält jedoch durch eine neuere Vergleichsuntersuchung zwischen dem zweiten und dem vierten Schuljahr eine gewisse Bestätigung (Heitmann 2001).

Versuch	A	K	A/K
KS	163	117	1,39
VK	86	88	1.12
BK	62	56	1.11
GA	32	35	0,91
BT	58	119	0,49

Tab. 6: Berghausen (versuchsbezogen)

## 5. Präzisierungen des Verfahrens

Das vorgestellte Verfahren der Quotientenbildung von Anzahlen von Schüleräußerungen besitzt einen grundlegenden Mangel. Je nach Anzahl der sog. „konkreten“ Schüleräußerungen kann es vorkommen, dass der Quotient sehr hohe Werte annimmt, falls nur wenige derartige Äußerungen vorliegen. Lassen sich in speziellen Fällen sogar überhaupt keine konkreten Äußerungen feststellen, so ist der Quotient nicht definiert (bzw. „wird unendlich groß“). Diesem Problem lässt sich abhelfen, indem man den A/K-Wert modifiziert und stattdessen den Wert  $A/(A+K)$  bildet. Dieser Wert existiert immer, er liegt zwischen null und eins (vgl. Tab. 7).

Bewertung der Einzelaussagen	Punkte
Vorschlag zu Versuchsvarianten	10
Beschreibung und Begründung einer Beobachtung	9
Handlungsunterstützte Begründung	9
Erklärung des Phänomens/Begründung einer Beobachtung	8
Analogiebildung/Versuch einer Analogiebildung	8
Handlungsunterstützte Vermutungsäußerung	7
Vermutungsäußerung	6
Wiederholung einer Begründung/Vermutung (abstrakt-kognitiver Art); Zustimmung	4
Handlungsunterstützte Beschreibung einer Beobachtung	4
Beschreibung einer Beobachtung	3
Wiederholung einer Beschreibung einer Beobachtung	2

Tab. 7: Bewertungsraster für Einzelaussagen nach Schorch (2001) und Heitmann

Aber auch bei diesem Vorgehen bleibt ein grundlegender Einwand noch bestehen: Die einfache Dichotomie der verwertbaren Äußerungen in „abstrakt“ und „konkret“ führe zu einer zu pauschalen Bewertung und berücksichtige keinerlei Abstufungen in den Schüleräußerungen.

Wir haben daher ein neues Bewertungsschema entwickelt, das graduelle Abstufungen in den Äußerungen durch eine Punktezuweisung berücksichtigt. Dabei geht es darum, Äußerungen, die eine größere kognitive Leistung (wie z. B. die Erklärung eines Phänomens) erkennen lassen, stärker zu gewichten als solche geringerer kognitiver Leistung (wie z. B. die bloße Zustimmung zu einer Begründung seitens eines anderen Probanden). Auch bei diesem Verfahren ist wie schon bei der Klassifizierung in A- und K- Aussagen wieder das Problem der subjektiven Einschätzung von Schüleräußerungen durch die auswertende Person anzumerken.

Mit diesem Raster wurden zwei Untersuchungssequenzen mit einem zweiten und vierten Schuljahr bewertet und das Ergebnis mit den A/K- bzw. den A/(A+K)-Werten verglichen. Wie die nachstehende Übersicht (Tab. 8) zeigt, lässt sich das Folgende aussagen:

	A/K	A/(A+K)	Punkte
2. Schuljahr			
Monika	5,50	0,85	6,54
Nadia	4,00	0,80	6,70
Jamil	1,62	0,62	5,72
Patrizia	1,35	0,57	5,10
Mittelwert	3,12	0,71	6,02
Fehler	±0,99	±0,07	±0,37
4. Schuljahr			
Stefan	2,52	0,72	5,83
Simon	1,47	0,60	5,77
Florian	3,29	0,77	5,55
Olaf	1,35	0,57	5,07
Mittelwert	2,16	0,65	5,83
Fehler	± 0,46	±0,04	±0,17

Tab. 8: Übersicht

1. Die Rangreihenfolgen ändern sich für das zweite Schuljahr in allen drei Bewertungsfällen überhaupt nicht, für das vierte Schuljahr fällt Platz 1 beim Übergang zur Punktbewertung auf Platz 3 zurück.
2. In allen drei Fällen liegen die Bewertungen für das zweite Schuljahr über denen für das vierte Schuljahr. Man vergleiche dazu das in Abschnitt 4 Ausgeführte.
3. Betrachtet man die in der Tabelle angegebenen quadratischen Fehler der Mittelwerte (vgl. dazu Abschnitt 3), so stellt man im Vergleich der drei Bewertungsmodi einen deutlichen Rückgang der prozentualen Überlappung der Werte zwischen dem zweiten und vierten Schuljahr fest:

A/K-Wert:

$$(0,46 + 0,99) - (3,12 - 2,16) = 0,49 \quad \text{und} \quad 0,49/2,64 = 18,6 \%$$

A/(A+K)-Wert:

$$(0,07 + 0,04) - (0,71 - 0,65) = 0,05 \quad \text{und} \quad 0,05/0,68 = 7,4 \%$$

Punktbewertung:

$$(0,37 + 0,17) - (6,02 - 5,56) = 0,08 \quad \text{und} \quad 0,08/5,79 = 1,4 \%$$

Wir kommen damit zu dem Ergebnis, dass die Präzisierung des Bewertungsverfahrens die Trennschärfe des Ergebnisses zwar beträchtlich erhöht, hingegen in der Rangfolge der einzelnen Schüler in den untersuchten Gruppen zu keiner grundlegenden Modifikation führt.

Man kann sodann die entsprechende Bewertung für die einzelnen Phänomene durchführen. Hier führt sowohl die A/K-Bewertung als auch die A/(A+K)-Bewertung in beiden Schuljahren zur gleichen Rangreihenfolge. Der Flaschengeistversuch steht im zweiten Schuljahr an der Spitze und die Kerzenschaukel im vierten Schuljahr (wie schon in der Nordhäuser Untersuchung) am Ende der Rangreihenskala. Die Punktbewertung führt für das vierte Schuljahr zu keiner Änderung der Rangreihenfolge. Im zweiten Schuljahr geht der „Balancierende Knopf“ vom vierten auf den zweiten Rang.

Zusammenfassend lässt sich damit feststellen, dass die vorgenommenen Präzisierungen nicht zu *grundlegend* neuen Einsichten führen und damit die Bestimmung der A/K-Werte für eine erste Bewertung hinreichend genaue Aufschlüsse sowohl über den Anteil von Verstehensoperationen bei den Schülern als auch über die intellektuelle Anregungsqualität der Versuche in den entsprechenden Präsentationen durch den Versuchsleiter geben kann.

Die vorstehenden Erläuterungen sollten vor allem das von uns verwendete Verfahren erläutern, weniger auf die damit erzielten Untersuchungsergebnisse abheben, die allein schon wegen der geringen Probandenzahlen statistisch nicht repräsentativ sind. Sollte sich jedoch z. B. der Rückgang von abstrakten Äußerungen in höheren Schulklassen auch in weiterführenden Studien bestätigen, so wäre dies sicher-

lich ein Befund, der Überlegungen zum Umgang mit Phänomenen im Sachunterricht anregen könnte.

## Literatur

- Günther, G. (1994). *Studien zum konkret-anschaulichen und abstrakt-kognitiven Vorgehen bei der Interpretation physikalischer Versuche durch Grundschüler*. Kassel: Wissenschaftliche Hausarbeit zur 1. Staatsprüfung für das Lehramt an Grundschulen an der Universität/Gesamthochschule Kassel.
- Hagstedt, H. & Spreckelsen, K. (1986). Wie Kinder physikalischen Phänomenen begegnen. *Sachunterricht und Mathematik in der Primarstufe*, 14, 318-322.
- Heitmann, A. (2001). *Analyse von Phänomenbegegnungssituationen im physikalischen Lernbereich des Sachunterrichts unter besonderer Berücksichtigung des Alters der Probanden*. Kassel: Wissenschaftliche Hausarbeit zur 1. Staatsprüfung für das Lehramt an Grundschulen an der Universität/Gesamthochschule Kassel.
- Krüger, S. (1993). *Formen selbständiger Verhaltensweisen beim physikalischen Experimentieren im Sachunterricht der Grundschule*. Kassel: Wissenschaftliche Hausarbeit zur 1. Staatsprüfung für das Lehramt an Grundschulen an der Universität/Gesamthochschule Kassel.
- Kurz, C. (2000). *Analyse und Diskussion der A/K-Werte in physikalischen Begegnungssituationen am Ende des zweiten Schuljahres*. Kassel: Wissenschaftliche Hausarbeit zur 1. Staatsprüfung für das Lehramt an Grundschulen an der Universität/Gesamthochschule Kassel.
- Lienert, G. A. & von Eye, A. (1994). *Erziehungswissenschaftliche Statistik*. Weinheim & Basel: Beltz.
- Press, H. J. (1964). *Spiel – das Wissen schafft*. Ravensburg: Otto Maier Verlag.
- Schorch, A. (2001). *Analyse von Phänomenbegegnungssituationen im physikalischen Lernbereich des Sachunterrichts unter besonderer Berücksichtigung der Charakteristika und Präsentation von Phänomenen*. Kassel: Wissenschaftliche Hausarbeit zur 1. Staatsprüfung für das Lehramt an Grundschulen an der Universität/Gesamthochschule Kassel.
- Spreckelsen, K. (1998). Bericht über das Forschungsvorhaben „Wie Kinder physikalischen Phänomenen begegnen“. In J. Kahlert (Hrsg.), *Wissenserwerb in der Grundschule* (S. 201-209). Bad Heilbrunn: Klinkhardt.

## Zur Entwicklung des begrifflichen Denkens im Grundschulalter

### 1. Theorie und Fragestellung

Grundlegend für die Orientierung in der Umwelt sind Erkennungsleistungen, die auf der Basis von Erkennungssystemen zustande kommen (Klix 1988). Wesentlicher Bestandteil der Erkennungssysteme sind Begriffe. Begriffe sind Klassifikationsergebnisse von Objekten bzw. Erscheinungen nach ihren Merkmalen und gestatten die verschiedensten unterschiedlichen Umweltinformationen in ein für das Leben relevantes Ordnungs- bzw. Orientierungssystem einzuordnen. Die Datenbasis für die Begriffe bilden klassifizierungsrelevante Merkmale bzw. Merkmalsätze. Die Klassifizierungsrelevanz dieser Merkmale ergibt sich nicht aus den Eigenschaften der Dinge und Erscheinungen an sich, sondern aus deren Bedeutung für das sie klassifizierende Subjekt. Die individuelle Ausbildung von Erkennungssystemen und die damit verbundene Zunahme an Variabilität und Flexibilität der auf dieser Basis möglichen Aktivitätsformen werden durch Lernen bedingt. Lernen ist allgemein als Prozess und Ergebnis umgebungsbezogener Verhaltensänderung in der Folge individueller Informationsverarbeitung aufzufassen (Edelmann 1996). Begriffe und Begriffssysteme, die durch Lernen erworben werden, sind flexibel variierbar, können umgelernt, neu gelernt, aber auch vergessen werden, wenn sie nicht mehr gebraucht oder unbrauchbar werden, d.h. es nicht mehr gestatten, die für das Leben des Individuums relevanten Anforderungen zu bewältigen. Damit wird klar: Wesentliche Determinanten (bzw. Faktoren) von Klassifizierungssystemen (vor allem von Begriffen und Begriffssystemen) sind a) die Besonderheiten (inhaltliche Charakteristik) der Anforderungen, für deren Bewältigung sie gebraucht werden; b) die Bewertungssysteme des Individuums (Bedeutung, Sinn), welche durch den Kontext (dessen wesentliches Moment die Aktivität des Individuums darstellt) gestiftet werden, sowie c) die Charakteristik des Lernens und die durch sie bedingte Entwicklung.

Die inhaltliche Seite, also besondere Erkennungsanforderungen sowie die Struktur und Prinzipien der internen Repräsentation des Wissens werden schwerpunktmäßig von der kognitiven Psychologie untersucht. Relativ offen bleibt dabei die Frage nach kontext- und entwicklungsabhängigen Besonderheiten der Erkennungs-

leistungen und besonders des begrifflichen Denkens. Dieser Fragestellung hat sich eine entwicklungspsychologische Forschungsrichtung zugewandt (Light & Butterworth 1992, Schnotz 1998, Carey & Gelman 1991, Sodian 1998).

Kernannahmen der den entsprechenden Forschungen zugrunde liegenden Theorien sind: Begriffe werden nicht losgelöst vom Kontext, in dem sie genutzt werden, angeeignet. Sie sind eingebettet in jeweils spezifische (naive oder aber höhere kulturelle – wissenschaftliche, künstlerische, theologische, sprachliche usw.) Theorien, die entscheidend das Vorwissen der Kinder charakterisieren (Vosniadou 1994). Die kognitive Entwicklung des Menschen verläuft nicht generalisiert, sondern domänenspezifisch, indem anforderungsbezogen auf der Basis des verfügbaren Vorwissens dieses erweitert (z. B. indem zusätzliches Wissen assimiliert – in die vorhandene kognitive Struktur eingebaut) wird, umkonstruiert oder neues Wissen bzw. neue kognitive Strukturen aufgebaut, konstruiert werden (Markman 1989, Chi et al. 1994, Spada 1994, Preub & Sink 1995). Dieses Wissen (Vorstellungen, Begriffe, domänenspezifische Methoden...) ist eingebettet in einen Theorierahmen, der eine personale Bedeutsamkeit aufweist (Hasselhorn & Mähler 1998). Die Sinnhaftigkeit (personale Bedeutung) einer Theorie hängt vor allem mit der Besonderheit der Auseinandersetzung des Individuums mit seiner Umwelt (z. B. mit dem Kontext bzw. der Tätigkeit) zusammen (Light & Butterworth a. a. O.). Die Kinder müssen im Verlauf ihrer Schulzeit einen Perspektivwechsel, Paradigmenwechsel im Denken vollziehen und ihr Denken auf die verschiedenen Domänen der höheren Kultur (vs. Alltag) richten (vgl. zur Theorie des „conceptual change“ Markman a. a. O., Carey et al. 1994, Chi et al. 1994, Spada 1994, Caravita and Hallden 1994, Schnotz 1998). Bis heute ist allerdings empirisch weitgehend ungeklärt, welche Mechanismen den bereichsspezifischen Theoriewandel im kindlichen Denken hervorrufen (Hasselhorn & Mähler a. a. O.). Relativ klar ist lediglich, dass ein solcher Paradigmenwechsel durch den Übergang vom Alltagsdenken zum wissenschaftlichen Denken, vom Alltagsbegriff zur Bildung wissenschaftlicher Begriffe im Sinne des Aufbaus einer neuen begrifflichen Ebene, gekennzeichnet ist.

Wissenschaftliche Begriffe zeichnen sich gegenüber den Alltagsbegriffen durch:

- Normierung (Intention, Extension festgelegt – Definition)
- Einbettung in einen theoretischen Zusammenhang und die
- Begründungspflicht (Behauptung, Erklärung, Rechtfertigung) aus (Merten 1999).

Da sich Kinder und Erwachsene weniger in der Struktur und den Prinzipien der internen Repräsentation des Wissens, sondern vor allem durch die kognitive Kapazität und die Fähigkeit zur metakognitiven Kontrolle unterscheiden (Schrempf & Sodian 1999, Hasselhorn & Mähler a. a. O.), dürfte gerade hierdurch die Entwicklung wissenschaftlicher Begriffe, die in besonderem Maße der metakognitiven



Kontrolle bedürfen und erfahrungsgemäß auch ein hohes Maß an kognitiver Kapazität beanspruchen, maßgeblich bedingt sein.

Wygotski (1987) unterscheidet vier Phasen in der Entwicklung von Begriffen: Synkretie, Komplex, Pseudobegriff und Begriff (wobei diese Stufe nicht notwendig erreicht werden muss). Synkretie, Komplex und Pseudobegriff werden durch Verallgemeinerung äußerer, sinnlich konkreter Merkmale (Erscheinungen) gebildet und unterscheiden sich durch die Abstraktionshöhe, Schärfe, Prägnanz sowie Stabilität jener der Klassenbildung zugrunde gelegten Merkmale bzw. Merkmalsätze untereinander. Leider hat Wygotski diese Phasen nur beschrieben und z. T. empirisch belegt, nicht aber den Übergang vom Vorbegriff (Synkretie, Komplex, Pseudobegriff) zum Begriff detailliert untersuchen können.

An diesen Ansatz von Wygotski anknüpfend und unter Bezugnahme auf das von ihm formulierte Gesetz der Verschiebung (1985, 1987) nehmen wir an, dass hinter aufeinander folgenden Entwicklungsetappen eine allgemeinere Gesetzmäßigkeit steht. So vermuten wir, dass diese Etappen nicht nur bei der Bildung der Alltagsbegriffe, sondern auch bei der Bildung wissenschaftlicher Begriffe auftreten werden. Analoge Hinweise auf Entwicklungsetappen, die den Übergang zwischen naiven zu wissenschaftlichen Theorien markieren, wurden u. a. von Vosniadou (1994) gefunden. Dieser Wechsel zeigt zwar individuelle Ausprägungen, ist jedoch für alle Individuen verpflichtend und betrifft insofern einen generellen Paradigmenwechsel, der etwa in den Klassen 3, 4 vollzogen wird (vgl. Schrempf & Sodian 1999). Weitgehend Einigkeit herrscht darüber, dass schulischer Unterricht im Sinne einer bedeutenden kontextstiftenden Bedingung ein Haupteinflussfaktor für diesen Begriffs- bzw. Theoriewandel darstellt (vgl. hierzu auch Edelmann a. a. O., Metz 1995, Schäfer 1999 und zu Defiziten des Unterrichts hierbei Mähler 1999, Grzesik 1992). Daher richten wir unsere Bemühungen darauf, den Einfluss des Unterrichts auf die Entwicklung des begrifflichen Denkens der Kinder zu untersuchen.

Auf dem Hintergrund der gekennzeichneten theoretischen Positionen fragten wir daher, welche Besonderheiten der Begriffsbildung bei Schülerinnen und Schülern in den Klassen 1 bis 4 festzustellen sind und wie diese durch unterschiedliche Unterrichtsstrategien beeinflusst wird.

## 2. Methode

Unter der Annahme einer weitgehend kontextuierten Kognition dürften Kontextvariablen wie Alltag und Unterricht einen entscheidenden Einfluss auf die Begriffsbildung haben. Da diese in der Regel auf Grund ihrer Komplexität kaum im Experiment gestaltet werden können, fehlen entsprechende Untersuchungen weitgehend. Bedingt durch den Transformationsprozess in den neuen Bundesländern haben sich zwei für die Begriffsbildung als bedeutsam angenommene Faktoren (un-

abhängige Variablen) verändert: a) der gesellschaftliche Alltag (Alltagshandeln) sowie b) die Unterrichtskonzeption (Merkmale des Lehrens und der Lerntätigkeit). Hieraus erwächst die Möglichkeit, die Wirkung dieser Variablen auf die Begriffsbildung und das begriffliche Denken zu prüfen.

Für die Untersuchung der Begriffsbildung und Begriffsverwendung stehen eine Reihe von Methoden zur Verfügung (vgl. hierzu Mähler 1999, Kluwe 1988). Wir haben uns für ein strukturiertes Interview entschieden, im Rahmen dessen Fragen nach einem a priori Modell der hypothetisch angenommenen Wissensstruktur formuliert wurden, und dies mit der Methode der Bildwahl und anschließender Begründung gekoppelt. Mit Blick auf die Analyse wissenschaftlicher Begriffe, die bewusstseinsfähig, wenn nicht bewusstseinspflichtig sind, erschien uns dies als die Methode der Wahl.

Für die Untersuchung von Entwicklungsprozessen sind Längsschnittuntersuchungen unerlässlich, auf Grund des mit ihnen verbundenen Aufwands jedoch sehr selten anzutreffen. Unsere Untersuchung verknüpft zwei Längsschnittuntersuchungen zur Begriffsbildung mit einer (kultur-) vergleichenden Studie (DDR 1988-1990 und Land Brandenburg 1997-2000). Hierzu konnten Daten einer 1988 in der ehemaligen DDR begonnenen Erhebung (strukturierte Interviews in den Klassen 1 bis 4 (1988-1991) – Einzeluntersuchungen, n= 30 Kinder x 4 Jahre) genutzt werden. In den Jahren 1997-2000 wurde eine analoge Untersuchung (n= 60 x 4 Jahre) im Land Brandenburg durchgeführt. Gegenstand der strukturierten Interviews waren Anforderungen der begrifflichen Identifikation und der Reflexion darüber sowie Inferenzleistungen über der begrifflichen Struktur bezogen auf für den Sachunterricht prototypische Begriffe („Pflanze“ und „Arbeit“ – hier nicht Gegenstand der Darstellung).

### 3. Ergebnisse

#### 3.1. *Entwicklung begrifflicher Identifikation*

Wir fragten, wodurch die Entwicklung der begrifflichen Identifikation von Pflanzen im Verlaufe der 4 Grundschuljahre gekennzeichnet ist. Den Kindern wurden dazu zehn Fotos (8 Vertreter typischer Pflanzenarten, 1 Tier, 1 Gegenstand aus dem Haushalt – Vase) präsentiert. Die Kinder wurden gebeten, die Bildinhalte zu identifizieren („Was siehst du auf dem Bild?“), zu entscheiden, ob der abgebildete Gegenstand eine Pflanze ist und schließlich die Entscheidung zu begründen. Hierdurch war zu ermitteln, auf Grund welches aktuell präsenten Merkmalsatzes die Entscheidung erfolgte. Schließlich wurden den Kindern die Bilder noch einmal vorgelegt, und sie wurden gebeten, nochmals (nach bewusster Reflexion bezüglich der Merkmale) zu entscheiden, ob die abgebildeten Gegenstände Pflanzen sind. Es zeigte sich, dass die abgebildeten Pflanzenarten bzw. Abteilungen (Moos, Farn,

Gras, Kraut, Strauch, Baum, Kaktus, Pilz) von Klasse 1 bis Klasse 4 im Trend zunehmend besser identifiziert werden. Damit konnten Ergebnisse querschnittlich angelegter Voruntersuchungen bestätigt werden (Giest 1999). Signifikante Unterschiede (die statistische Prüfung der dargestellten Ergebnisse erfolgte mit dem Friedman-Test/SPSS – nichtparametrische Tests für k verbundene Stichproben) zeigten sich bei: Kaktus  $p = .025$ ; Kraut (Narzisse)  $p < .001$ ; Farn  $p < .001$ ; Gras (Getreide)  $p = .001$ ; Strauch (Forsythie)  $p < .001$ ; Moos  $p = .004$ . Auffällig war, dass die Identifikation von „Frühblühern“ (diese wurden zumindest in der DDR traditionell in Klasse 2 behandelt, und diese Tradition scheint fortzuleben, was Unterrichtsbeobachtungen bestätigten) in dieser Klassenstufe offenbar deutlich besser gelingt.

Bei der Entscheidung, welche auf den Bildern präsentierten Objekte Pflanzen sind, ergaben sich insgesamt hochsignifikante ( $p < .001$ ) Entwicklungseffekte, die vor allem für Farn ( $p = .019$ ), Baum ( $p = .035$ ), Pilz ( $p < .001$ ) sehr deutlich ausfielen. Beeinflusst das bewusste Begründen der Entscheidung (Bewusstmachen des Merkmalsatzes) die Identifikationsleistung? Aus Abb. 1 ist ersichtlich, dass dies der Fall ist. Die Identifikation der bildlich präsentierten Objekte erfolgt nach der Begründung dieser Entscheidung hochsignifikant ( $p < .001$  – Wilcoxon Test) besser. Das Bewusstmachen des einer begrifflichen Identifikation zugrundeliegenden Merkmalsatzes erhöht die Trefferquote bei der Identifikation. (Aus Platzgründen und auf Grund fehlender konsistenter Effekte werden die Ergebnisse einer differentiellen Varianzanalyse bezüglich der Variablen Geschlecht, Schule, Leistungsgruppe an dieser Stelle nicht mitgeteilt.)

Auffällig ist die Überlagerung von Effekten des Unterrichts (Artenkenntnisse in Klasse 2 sind besser als in 3 und 4) und der allgemeinen kognitiven Entwicklung (kontinuierlicher Anstieg der adäquaten Identifikation als Pflanze in den Klassen 1 bis 4).

Wir haben die Merkmale, mit denen begründet wurde, warum ein Objekt eine Pflanze ist, entsprechend unserer Hypothese in die Kategorien vorbegrifflicher und begrifflicher Stadien eingeordnet:

Vorbegriffliche Stadien der Begriffsbildung sind:

Synkretie – gekennzeichnet durch situativ zufällige Begründungen (Merkmalsnennung) bzw. fehlende explizit zu nennende Merkmale (In die Begründungen geht die Ähnlichkeit zum Primärbegriff – „weil es so aussieht“ und die Objektklassifikation (Tautologie) „weil es eine Pflanze, ein Baum...ist“ ein.);

Komplex – gekennzeichnet durch weniger zufällige dem Objekt zugeordnete oft anschauliche Merkmale (grün, in der Erde, rot, gelb, für Mutti, duften, in der Natur – Lokation);

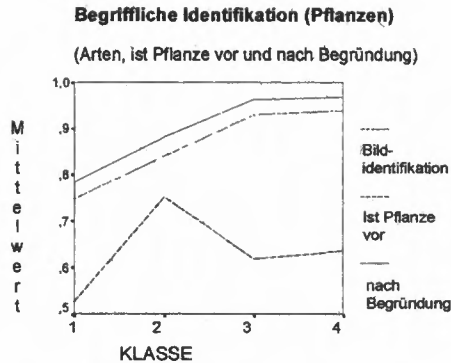


Abb. 1: Entwicklung begrifflicher Identifikation von Pflanzen in den Klassen 1 bis 4  
(Werte entsprechen 1/10 je Treffer)

Pseudobegriff – gekennzeichnet durch einseitige Hervorhebung (oft im konkreten Zusammenhang unwesentlicher aber) allgemeiner, abstrakter Merkmale („weil es Lebewesen sind“, „weil sie kein Gegenstand sind“ bzw. Nennung abstrakter Materialeigenschaften).

Begriffliche Stadien der Begriffsbildung sind:

„empirischer“ Begriff – kennzeichnend sind morphologische Merkmale: Blüte, Blätter, Stiel (Sprossachse), Wurzel und

„theoretischer“ Begriff – kennzeichnend sind physiologische bzw. Merkmale des Lebens: Wachstum, Bewegung, Fortpflanzung, Ernährung – Stoffwechsel im Falle der Pflanzen die Photosynthese.

Es bleibt anzumerken, dass die gekennzeichneten psychologischen Kategorien des Begriffs dem biologischen Begriff „Pflanze“ nicht voll entsprechen, da innerer und äußerer Bau der Lebewesen, einschließlich der durch diesen Bau gestifteten Funktionsmerkmale, in das natürliche System der Lebewesen eingehen. Ferner ist anzumerken, dass nicht vor Einsetzen des fachwissenschaftlichen Lernens mit der psychischen Existenz des wissenschaftlichen Begriffs gerechnet werden kann. Es kann sich hier nur um Vorformen in Analogie zu den synthetischen Modellen der Wirklichkeit handeln, bei denen Initialmodell (durch den natürlichen Realismus der Kinder gestiftet) und wissenschaftliche Vorstellung eine Symbiose eingehen (Vosniadou 1994). Die Analyse dieser Kategorien lässt eine deutliche Entwicklung erkennen: Vorbegriffliche Stadien nehmen ab und begriffliche Stadien zu (vgl. Abb. 2a).

Die statistische Prüfung ergab bezogen auf alle 4 Messzeitpunkte hochsignifikante Unterschiede (vorbegriffliche Stufe –  $p < .001$ ; begriffliche Stufe  $p < .001$ ). Betrachtet man die einzelnen Kategorien begrifflicher Entwicklung (vgl. Abb. 2b), so zeigt

sich ein ähnliches Bild. Auch hier lässt sich statistisch ein hochsignifikanter Entwicklungstrend bestätigen ( $p < .001$ ). Was hier aus Platzgründen nicht gezeigt werden kann, ist eine detaillierte Fallanalyse, die sowohl Kinder zeigt, die im Entwicklungstrend liegen, aber auch Entwicklungsverzögerungen, Entwicklungsvorsprünge und „Ausreißer“ (vornehmlich in Klasse 2) werden deutlich: Wurden bei Kindern in Klasse 1 besonders viele Synkretien festgestellt, so konnte beobachtet werden, dass in Klasse 4 weniger morphologische Merkmale die Begriffsbildung kennzeichnen, und umgekehrt, wurden in Klasse 1 viele morphologische Merkmale in den Begründungen und wenige Synkretien registriert, so verläuft die Entwicklung hin zur Bildung adäquater Begriffe deutlicher ausgeprägt.

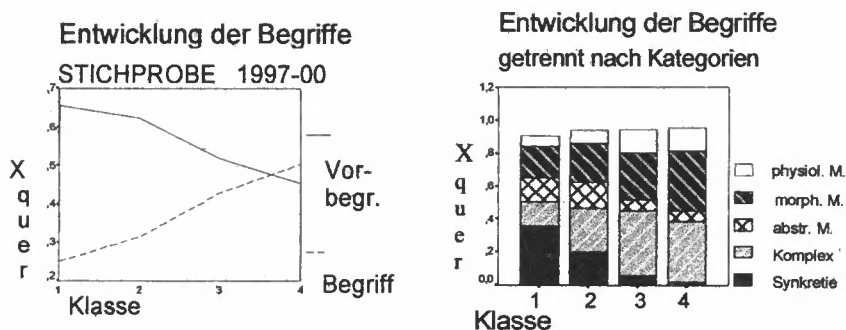


Abb. 2a, b: Entwicklung begrifflicher Kategorien in den Klassen 1 bis 4 (2a – siehe Anmerkung Abb. 1; 2b – Aus Gründen der Übersichtlichkeit wurde die Kategorie „ohne oder andere Angaben“ nicht in die Graphik eingearbeitet, wodurch die Differenz zu 1,0 bei den Stapeldiagrammen zu erklären ist.)

### 3.2. Wirkung der Variable Unterricht

Um die Wirkung des Unterrichts zu prüfen, haben wir zwei unterschiedliche Unterrichtsstrategien verglichen: Zum einen den Unterricht in der DDR, der dem Modell der direkten Instruktion folgt und eher als geschlossen zu kennzeichnen ist und andererseits den gegenwärtigen Unterricht im Land Brandenburg (Modell der indirekten Instruktion, Versuch der Öffnung von Unterricht – vgl. Giest 1999). Auch bei der Stichprobe von 1988-1991 ist ein signifikanter Entwicklungseffekt festzustellen, wenngleich dieser nicht so deutlich ausfällt, wie in der Stichprobe 97-00 (begriffliches Stadium  $p < .001$ ; vorbegriffliches Stadium  $p = .034$ ). Was hier jedoch deutlich auffällt, sind die vergleichbaren Werte in der zweiten und vierten Klasse (Stichprobe 1988/1991), welche wohl eindeutig durch den Unterricht bedingt sein dürften (vgl. Abb. 3a und b). In Klasse 2 wurde seinerzeit im Lehrplan

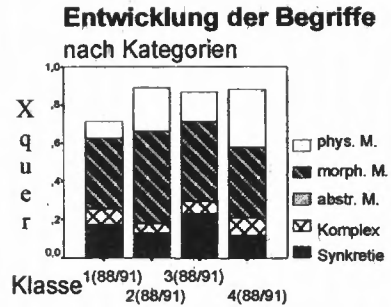
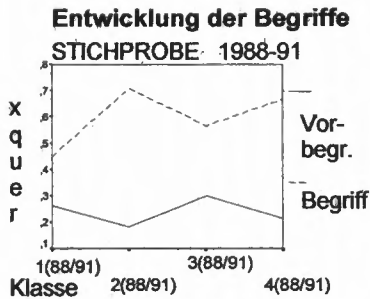


Abb. 3 a, b: Begriffliche Entwicklung in den Klassen 1 bis 4 (1988/1991) (Siehe auch Anmerkung zu Graphen bei Abbildung 2a und b.)

explizit die Behandlung der morphologischen Merkmale der Pflanzen (Blüte, Laubblatt, Stängel, Wurzel) beim Thema Frühblüher gefordert. Dieses Ergebnis deckt sich mit anderen analogen Beobachtungen (vgl. Giest 1999).

Vergleichen wir beide Stichproben, so zeigt sich ein deutlicher Unterschied (vgl. Abb. 4). Die Prüfung der Unterschiede zwischen beiden Stichproben (Mann-Whitney-U-Test) erbrachte hochsignifikante Unterschiede ( $p < .001$ ) bei der vorbegrifflichen wie auch bei der begrifflichen Stufe.

#### 4. Diskussion

Die dargestellten Ergebnisse bilden nur einen sehr begrenzten Ausschnitt der Gesamtuntersuchung ab. Sie machen jedoch auf wichtige Entwicklungstrends der kognitiven Entwicklung der Kinder in den ersten vier Schuljahren aufmerksam. Analog der conceptual change Theorie zeigt sich ein Perspektivwechsel: der Übergang vom Alltagsbegriff zum wissenschaftlichen Begriff. Dieser Wechsel vollzieht sich in verschiedenen Phasen, welche Analogien zur Begriffsbildung bei Kleinkindern (Alltagsbegriffe) aufweisen (Gesetz der Verschiebung). Analog zur Entwicklung wissenschaftlicher Vorstellungen von komplexen Erscheinungen (Tag, Nacht, Gestalt der Erde u. a.) lassen sich drei Grundphasen unterscheiden: Initiale Stadien der Begriffsbildung (Primärbegriffe – bezüglich wissenschaftlicher Begriffsbildung als vorbegriffliche Phase zu kennzeichnen), synthetisches Stadium (gekennzeichnet durch die Zunahme von Merkmalen wissenschaftlicher Begriffe, allerdings ohne die wissenschaftliche Begriffsbildung zu erreichen) und die wissenschaftliche Begriffsbildung. Letztere ist im betrachteten Alter, zumindest bei den untersuchten Begriffen, noch nicht voll ausgeprägt.

Obwohl deutliche Unterschiede bezüglich der durch Unterricht bedingten Entwicklungseffekte zwischen beiden Stichproben zu verzeichnen sind, sollte die

## Begriffsbildung bei Schülern 1. bis 4. Klassen

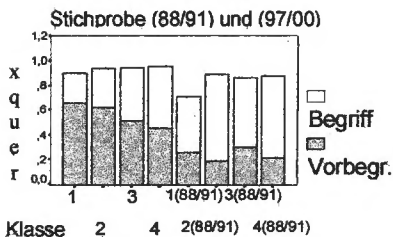


Abb. 4: Vergleich der Begriffsentwicklung in den beiden Studien (Siehe auch Anmerkung zu Graphen bei Abbildung 2b.)

Wirkung des Unterrichts in der DDR-Stichprobe hinsichtlich der Begriffsbildung durchaus kritisch beurteilt werden. Offenbar war der Effekt des Unterrichts in Klasse 2 nicht von Dauer, hatte demzufolge wenig Wirkung auf die eigentliche Begriffsentwicklung. Dennoch macht der Vergleich der beiden Stichproben deutlich, dass a) der Unterricht eine entscheidende Wirkung auf die Begriffsbildung haben kann und b) durch Unterricht bedingte Effekte allgemeinere Entwicklungstrends überlagern sowie c) es bislang und vor allem gegenwärtig offenbar nur wenig gelingt, Unterricht mit Blick auf die Begriffsbildung bei den Kindern effektiv zu gestalten.

Die Analyse der einzelnen Entwicklungsstadien der Begriffe gestattet, die Zone der nächsten Entwicklung der Kinder zu umschreiben und Konsequenzen für einen entwicklungsorientierten Unterricht abzuleiten. Dieser muss zum einen auf die Zone der nächsten Entwicklung der Kinder (hier auf die wissenschaftliche Begriffsbildung) bewusst orientiert werden und zum anderen durch die Gestaltung adäquat kontextuierter Lernsituationen konkreter Einfluss auf die Begriffsbildung (als einen vom Kind vollzogenen Lernprozess) nehmen. Gerade in der zuletzt gekennzeichneten Richtung sind weitere Untersuchungen notwendig.

## Literatur

- Caravita, S. & Hallden, O. (1994): Re-Framing the problem of conceptual change. *Learning and Instruction*, 4, H.1, 89-112.
- Carey, S.; Gelman, R. (Hrsg.) (1991): The epigenesis of mind. Essays on biology and cognition. Hillsdale, NJ u. a.: Erlbaum.
- Carey, S. & Spelke, E. S. (1994): Domain specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Sds): Mapping the mind: Domain specificity in cognition and culture. (pp 169-200) Cambridge: Cambridge University Press.
- Chi, M. T. H.; Slotta, J. D. and de Leeuw, N. (1994): From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, 4, H.1, 27-44.

- Edelmann, W. (1996): Lernpsychologie. Weinheim: Beltz.
- Giest, H. (1999): Lernen und Lehren in der Grundschule. Empirische Erhebungen im Sachunterricht. Potsdam: Universität Potsdam.
- Grzesik, J. (1992): Begriffe lernen und lehren. Stuttgart: Klett.
- Hasselhorn, M. & Mähler, C. (1998): Wissen, auf das Wissen baut: Entwicklungspsychologische Erkenntnisse zum Wissenserwerb und zum Erschließen von Wirklichkeit im Grundschulalter. In J. Kahlert (Hrsg.), Wissenserwerb in der Grundschule (S. 73-90). Bad Heilbrunn: Klinkhardt.
- Klix, F. (1988): Gedächtnis und Wissen. In H. Mandl und H. Spada (Hrsg.), Wissenspsychologie (S. 19-54). München, Weinheim: Psychologie Verlags Union.
- Kluwe, R. (1988): Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen. In H. Mandl und H. Spada (Hrsg.), Wissenspsychologie (S. 359-385). München, Weinheim: Psychologie Verlags Union.
- Light, P. & Butterworth, G. E. (Hrsg.) (1992): Context and cognition. Ways of learning and knowing. New York u. a.: Harvester Wheatsheaf.
- Mähler, C. (1999): Naive Theorien im kindlichen Denken. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31, H.2, 55-65.
- Markman, E. M. (1989): Categorization and naming in children. Problems of induction. Cambridge, Mass. (MIT).
- Merten, R. (1999): Verständigungsprobleme. *Zeitschrift für Pädagogik*, 45, H.2, 195-208.
- Metz, K. E. (1995): Reassessment of Developmental Constraints on Children's Science Instruction. *Review of Educational Research*, 65, H.2, 93-127.
- Preub, A. & Sink, Th. (1995): Symposium on Conceptual Change. *EARLI News*, 1, 17-18.
- Schäfer, G. E. (1999): Frühkindliche Bildungsprozesse. *Neue Sammlung*, 39, 2, 213-226.
- Schnotz, W. (1998): Conceptual change. In D. H. Rost (Hrsg.), Handwörterbuch der Pädagogischen Psychologie (S. 55-59). Weinheim: Beltz, Psychologie Verlags Union.
- Schrepp, I. & Sodian, B. (1999): Wissenschaftliches Denken im Grundschulalter. Die Fähigkeit zur Hypothesenprüfung und Evidenzevaluation im Kontext der Attribution von Leistungsergebnissen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 3, H.2, 67-77.
- Sodian, B. (1998): Der Beitrag nativistischer Ansätze zur entwicklungspsychologischen Theoriebildung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 30, H.4, 174- 178.
- Spada, H. (1994): Conceptual Change or multiple Representations? *Learning and Instruction*, 4, H.1, 113-116.
- Vosniadou, S. (1994): Capturing and Modeling the Process of Conceptual Change. *Learning and Instruction*, 4, H.1, 45-70.
- Wygotski, L. S. (1985, 1987): Ausgewählte Schriften, Bd. 1 u. 2. Berlin: Volk und Wissen.



## **Lernfortschrittsdiagnosen durch Interviews – Ergebnisse einer Pilotstudie zum „Schwimmen und Sinken“ im Sachunterricht der Grundschule**

### **1. Problemstellung**

Naturwissenschaftsbezogenes Lernen in der Grundschule ist seit der sog. Wissenschaftsorientierung in den siebziger Jahren kritischen Einwänden ausgesetzt. Einer der Haupteinwände bezieht sich auf die Qualität des erworbenen Wissens. Insbesondere wird bezweifelt, dass bereits Grundschulkindern aufgrund ihrer Lern- und Entwicklungsbedingungen zu einem Verstehen naturwissenschaftlicher Phänomene in der Lage sind.

Indizien für eine kognitive Überforderung von Grundschulkindern ergaben sich aus Untersuchungen zum Erfolg der eher wissenschaftsorientierten, amerikanischen Curricula in den siebziger und achtziger Jahren. Es zeigte sich, dass der Unterricht zwar positive Wirkungen auf die Förderung methodischer Fähigkeiten und Fertigkeiten und die Steigerung des Interesses hatte, nicht aber auf das Verständnis naturwissenschaftlicher Begriffe. Auch Vertreter eines naturwissenschaftsbezogenen Elementarunterrichts in Deutschland kamen aufgrund von Untersuchungen zu dem Urteil, dass die Bemühung, begriffliches Fachwissen anspruchsvoll zu vermitteln, grundsätzlich als misslungen gelten muss: „Grundschulkindern lernen nicht, die naturwissenschaftlichen Fachbegriffe zu verstehen, sondern bestenfalls Wörter, die für sie stehen, assoziativ und grammatikalisch korrekt zu gebrauchen“ (Lauterbach 1992, S. 205).

Neuere Untersuchungen aus dem Bereich der Sachunterrichtsdidaktik zeigen aber, dass Lernfortschritte im Vorfeld der Naturwissenschaften möglich sind, wenn Unterricht angemessen gestaltet wird. Bewährt haben sich genetisch-sokratische (vgl. Köhnlein 1999) bzw. nach moderat-konstruktivistischen Gesichtspunkten gestaltete Lehr-Lernumgebungen (vgl. Möller 1999).<sup>2</sup> Allerdings bleibt die Frage bestehen, ob derartige Lehr-Lernumgebungen zu individuellen Lernfortschritten bei allen Kindern führen.

Auch muss geprüft werden, ob es sich bei dem angeeigneten Wissen um wirklich verstandenes, anwendungsfähiges Wissen handelt, nicht um ein träges Wortwissen. Schriftliche Leistungstests geben aber hierüber nur unzureichend Auskunft.

Die vorliegende Studie verfolgt das Ziel, über Einzelinterviews, die vor und nach einem moderat-konstruktivistisch orientierten Unterricht (mit instruktiven Anteilen) durchgeführt werden, Lernfortschritte im Verstehen physikalischer Basiskonzepte zu diagnostizieren. Als Unterrichtsgegenstand wurde der Bereich „Schwimmen und Sinken“ gewählt, da zu diesem Thema sowohl in der Didaktik als auch in der Entwicklungspsychologie bereits eine Reihe teilweise widersprüchlicher Untersuchungen zum inhaltlichen Verstehen vorliegen. Insbesondere gilt als ungeklärt, ob Grundschul Kinder bereits ein angemessenes Verständnis bei anspruchsvollen physikalischen Begriffen wie „Dichte“ und „Auftrieb“ erreichen können.

## 2. Theoretische Grundlagen

### 2.1 Zur Gestaltung der Lehr-Lernumgebung

In Voruntersuchungen (vgl. Möller 1999) haben sich Lehr-Lernumgebungen als verstehensfördernd erwiesen, die auf der Basis konstruktivistischer Lehr-Lerntheorien entwickelt wurden. Entsprechende, häufig als „moderat konstruktivistisch“ bezeichnete Lerntheorien (vgl. Duit 1995; Gerstenmaier & Mandl 1995) gehen von einem als aktiv, konstruktiv, sozial, selbstgesteuert und situiert zu beschreibenden Wissenserwerb aus, in dem Konzepte aktiv verändert werden; eine rezeptive Wissensaneignung ist danach nicht möglich.

Eigene Untersuchungen wie auch vorliegende Forschungsbefunde (vgl. z. B. Renkl, Gruber & Mandl 1996; Weinert & Helmke 1995; Reinmann-Rothmeier & Mandl 1999) haben allerdings gezeigt, dass eine zu starke Situietheit von Lernsituationen wie auch ein zu großer Freiraum bei der Selbstbestimmung zu Problemen führen kann. In Anlehnung an Modifizierungen konstruktivistischer Ansätze entwickelten wir sog. „*moderat-konstruktivistische Lehr-Lernumgebungen mit instruktiven Anteilen*“, die auf einen aktiven und konstruktiven Wissenserwerb ausgerichtet sind und dem Dialog unter den Kindern wie auch dem kooperativen Wissenserwerb einen hohen Wert beimessen. Ergänzt wird der Unterricht durch eine geeignete Gliederung des Unterrichtsinhalts, die wir als Sequenzierung bezeichnen, und durch eine angemessene Lehrerunterstützung; diese Elemente fassen wir unter dem Begriff „*instruktive Anteile*“ zusammen (vgl. Möller 2001a, c).

Wir erwarten von einem derartigen Unterricht, dass sich kognitive und nicht-kognitive Zielsetzungen auch bei unterschiedlichen Eingangsbedingungen (Geschlecht, Leistungsstärke) erreichen lassen.

### 2.2 Zum Themenbereich „Schwimmen und Sinken“

Sowohl in der naturwissenschaftsdidaktischen Forschung (vgl. z. B. Banholzer 1936; Thiel 1990; Kircher & Rückel 1992) als auch in der Entwicklungspsychologie (vgl. Piaget & Inhelder 1977; Carey 1991; Janke 1995) wurde das

Thema „Schwimmen und Sinken“ vielfach aufgegriffen. Die Ergebnisse zeigen, dass begriffliche Verständigungsprobleme und tiefverwurzelte Präkonzepte wie auch die notwendige Integration mehrerer Aspekte (z. B. beim Dichtebegriff das Verhältnis von Volumen und Masse) Lernschwierigkeiten erzeugen. Nur die wenigsten Kinder beziehen beim Auftrieb die Rolle des Wassers in ihre Überlegungen von sich aus ein (vgl. Möller 1999). Andererseits zeigen neuere, aus der Tradition Piagets heraustretende Untersuchungen in der Entwicklungspsychologie, dass auch jüngere Kinder unter bestimmten Bedingungen, insbesondere in sinnvollen, einsichtigen Kontexten, bereits für physikalische Kausalitäten sensibel sind, schlussfolgernd denken und über ein intuitives Dichteverständnis verfügen. Auch von Thiel (1990) und Kircher & Rückel (1992) mit Grundschulkindern durchgeführte Unterrichtseinheiten sprechen für mögliche Lernfortschritte bei einem Unterricht über „Auftrieb“.

### **3. Eine Untersuchung zur Diagnose von Lernfortschritten im Themenbereich „Schwimmen und Sinken“**

Im Rahmen dieser Untersuchung, die als Pilotstudie angelegt war, konzentrierten wir uns auf die Ermittlung kognitiver Lernfortschritte von 17 Kindern (8 Jungen, 9 Mädchen) in einem nach den oben genannten Prinzipien konzipierten Unterricht zum Schwimmen/Sinken in einer dritten Klasse.<sup>3</sup> Ziel dieser Untersuchung war es, eine Methode zur differenzierten, quantifizierbaren Erfassung von Lernfortschritten im Hinblick auf das Verstehen physikalischer Basiskonzepte zu entwickeln und zu erproben.

#### **3.1 Zur Unterrichtsintervention**

Der Unterricht umfasste in einem Zeitraum von vier Wochen acht Doppelstunden (einschließlich einer Doppelstunde im Schwimmbad) mit folgenden Teileinheiten:

- das Schwimmen und Sinken verschiedener Vollkörper, das die Überprüfung von Präkonzepten und die Ausbildung von Material- oder Dichtekonzepten initiiert
- die Erfahrung der Auftriebskraft und das Erfassen von Zusammenhängen zwischen Wasserverdrängung und Auftriebskraft
- die Integration der Erkenntnisse, um damit die komplexe Fragestellung nach dem Schwimmen eiserner Hohlkörper zu beantworten.

Zum Unterricht gehörten das Ausprobieren verschiedener Materialien (Vollkörper) im Hinblick auf Schwimmen bzw. Sinken, wobei intensiv auf nicht belastbare Konzepte eingegangen wurde (z. B. Luftkonzept, Gewichtskonzept) (1. und 2. Doppelstunde), das Durchführen von Versuchen im Klassenzimmer und im Schwimmbad, wobei das „Drücken des Wassers“ erfahren werden konnte (3. und

4. Doppelstunde), das Herstellen und Beladen von schwimmfähigen und tragfähigen Knettschiffen und die Repräsentation der gegeneinandergerichteten Kräfte „Auftriebskraft“ und „Gewichtskraft“ in einem Spiel und einer zeichnerischen Darstellung (ohne dabei allerdings diese fachlichen Begriffe zu benutzen) (5. und 6. Doppelstunde), die Integration der erworbenen Erkenntnisse zur Beantwortung der Frage nach dem Schwimmen eines eisernen Schiffes (7. Doppelstunde) und das Anwenden der erworbenen Kenntnisse beim Bauen von Schiffen aus unterschiedlichen Materialien (8. Doppelstunde).

Der Unterricht bot über Experimente, welche die Kinder selbständig durchführten, vielfältige Möglichkeiten zum aktiven Erfahrungserwerb. In Gruppen-, häufig aber auch in Klassengesprächen, die von der Lehrerin<sup>4</sup> durch geeignete Impulse unterstützt wurden, hatten die Kinder Gelegenheit, ihre Erfahrungen zu konzeptualisieren, Vermutungen aufzustellen, zu überprüfen und zu korrigieren.

### 3.2 Erhebung und Datenaufbereitung

Siebzehn der sechsundzwanzig Schülerinnen und Schüler wurden vor dem Unterricht und zehn Wochen nach dem Unterricht unter Einsatz eines teilstandardisierten Leitfadens einzeln interviewt (Interviewdauer: 30-45 Minuten). Die Interviews enthielten 18 auswertbare Aufgaben zur Teiluntersuchung „Vollkörper“ und vier auswertbare Aufgaben zur Teiluntersuchung „Auftriebskraft“.

In der Teiluntersuchung „Vollkörper“ wurden den Kindern 18 verschiedene Gegenstände gezeigt. Bei jedem Gegenstand (Holzknopf, Metalldraht, ein großes Stück Holz, ein winziges Stück Metall...) wurden die Kinder gebeten, eine Vorhersage zum Schwimmen bzw. Sinken dieses Gegenstandes abzugeben und die Vorhersage zu begründen.

In der Teiluntersuchung „Auftriebskraft“ führten die Kinder vier Experimente durch, die sie erklären sollten (ein Stein, mit einem Bindfaden an einer Angel befestigt, der ins Wasser getaucht wird und dabei spürbar leichter wird; zwei vom Volumen unterschiedliche Becher, die ins Wasser gedrückt werden und einen unterschiedlichen Widerstand erspüren lassen; ...)<sup>5</sup>

Die Integration der Erfahrungen im Hinblick auf die Unterrichtsfrage „Wie kommt es, dass ein eisernes Schiff nicht untergeht?“ (Teileinheit „Integration“) wurde im Prätest im offenen Format schriftlich, im Posttest gleichlautend mündlich erfragt.

Um Lernfortschritte zu erfassen, waren die Aufgaben in den Prä- und Post-Einzelinterviews identisch gewählt. Die Interviews wurden auditiv aufgezeichnet und vollständig wörtlich transkribiert.

### 3.3 Zur Erhebungsmethode

Einzelinterviews sind ein geeignetes Erhebungsinstrument zur Diagnostik von individuellen Verstehensprozessen (vgl. auch White & Gunstone 1999). Schriftli-

chen Tests sind sie überlegen, da Nachfragen durch den Interviewer möglich sind. Kinder können so ihre Äußerungen präzisieren und auch korrigieren. Zudem umgehen Interviews das Problem der begrenzten Lese- und Schreibfähigkeit von Grundschulkindern. Nach unseren Erfahrungen lassen sich die meisten Grundschul Kinder ausgesprochen bereitwillig auf solche Einzelinterviews ein.

Um dem konstruktiven Charakter von Lernprozessen auf der Basis von Vorerfahrungen gerecht zu werden, bietet es sich an, Lernfortschritte durch einen Vergleich von Wissensbeständen vor und nach einer Unterrichtsintervention zu ermitteln (Prä- und Post-Einzelinterviews). Durch ein solches Design lassen sich *individuelle Lernfortschritte* erheben.

Um Interviews nicht nur qualitativ, sondern auch quantitativ auszuwerten, müssen die Aufgaben im Prä- und Post-Einzelinterview konstant und durch ein zumindest teilstandardisiertes Format auch vergleichbar gehalten werden. Gegenüber standardisierten Interviews haben teilstandardisierte Interviews den Vorteil, dass eine Vergleichbarkeit in Bezug auf die gestellten Aufgaben gegeben ist und dennoch Freiräume für Nachfragen des Interviewers vorhanden sind.

### 3.4 Datenauswertung

Die Interviewtexte wurden mit Hilfe induktiv gefundener Kategorien nach dem Verfahren der strukturierenden Inhaltsanalyse (vgl. Mayring 1995) aufbereitet. Als Auswertungseinheit wurde jeweils eine Aufgabe definiert. Die Kategorien beziehen sich auf den Inhalt der geäußerten Konzepte. (Beispiel: Ein Baumstamm geht unter, weil er so schwer ist = Gewichtskonzept; weil er aus Holz ist = Materialkonzept; weil er für seine Größe leicht ist = Dichtekonzept; vgl. dazu Tab. 1).

Sämtliche Interviewantworten wurden mit Hilfe der definierten und durch Beispiel-Items konkretisierten Kategorien kodiert. Die Kodierung wurde durch eine zweite Person geprüft, wobei die Übereinstimmung bei ca. 90 % lag. Im Falle von Mehrfachantworten wurde eine Entscheidung für das Konzept getroffen, welches das Kind bei allen Aufgaben am häufigsten genannt hat. Für den Fall, dass Konzepte bei den anderen Aufgaben gleich häufig vorkamen, wählten wir im zweiten Schritt das inhaltlich auf einem höheren Erklärungsniveau stehende Konzept aus. War das Erklärungsniveau jedoch gleich, wurde das vom Kind zuerst genannte Konzept ausgewählt.

In einem deduktiven Vorgehen wurden den gefundenen Kategorien Levels in Bezug auf das erreichte Verständnis zugeordnet: *Level 1* umfasst die Kategorie „*weiß nicht*“.

Unter *Level 2* wurden „*nicht belastbare Konzepte*“ zusammengefasst, die einer Überprüfung an der Sache und im Diskurs nicht standhalten, weil sie durch nicht zutreffende Vorstellungen geprägt sind. In diesen Level wurden zum Beispiel folgende Antworten eingeordnet: Ein riesiges, großes, schweres Schiff aus Eisen geht

nicht unter, „weil da ein Kapitän drauf ist. Wenn er runtergeht vom Schiff, dann sinkt es. Deshalb geht er auch als letztes von Bord.“ – „Weil es einen Motor hat.“ – „Im Hafen wird es festgebunden, damit es nicht untergeht.“ – „Weil da Luft drin ist. Luft will immer nach oben und dann zieht die das Schiff mit hoch.“ Bei der Frage nach dem Schwimmen bzw. Sinken von Gegenständen: „Ein Stück Holz schwimmt, aber ein großer Baumstamm schwimmt nicht, weil er so schwer ist.“ – „Ein Holzknopf geht unter, weil da Löcher drin sind.“ Als „ausbaufähig“ wurden diejenigen Konzepte bezeichnet, die Anknüpfungsmöglichkeiten für die Entwicklung angemessenerer Konzepte bieten (*Level 3*). Ausbaufähig ist zum Beispiel das Materialkonzept beim Schwimmen bzw. Sinken von Vollkörpern (Holz schwimmt, Eisen geht unter), da es Anknüpfungsmöglichkeiten für das Konzept „Dichte“ als „Eigenschaft von Materialien“ bietet.

Als „Vorkonzepte“ wurden qualitative Konzepte bezeichnet, die den physikalischen Konzepten nahe kommen, weil sie Kernideen physikalischer Konzepte beinhalten (*Level 4*): Die Eisennadel geht unter, „weil sie für ihre Größe ganz schwer ist.“ „Das Schiff ist leichter als genauso viel Wasser.“ „Weil das Schiff so viel Platz im Wasser braucht, drückt das Wasser stärker.“ „Das Wasser drückt (beim Schiff) stärker als das Gewicht zieht und dann gewinnt das Wasser.“ „Weil das Schiff so groß ist, will viel Wasser zurück an seinen Platz und drückt ganz stark, stärker als das Gewicht vom Schiff nach unten zieht.“<sup>6</sup>

Den erreichten Levels wurden je Aufgabe – der ansteigenden Verstehensqualität entsprechend – Leistungspunkte zugeordnet, um einen gemeinsamen Leistungsscore je Teiluntersuchung berechnen zu können (*Level 1*: 0 Punkte, *Level 2*: 1 Punkt, *Level 3*: 2 Punkte, *Level 4*: 3 Punkte). Maximal waren daher in der Teiluntersuchung „Vollkörper“ bei 18 Aufgaben 54 Punkte, in der Teiluntersuchung „Auftriebskraft“ bei vier Aufgaben 12 Punkte zu erreichen (maximal erreichbare Leistung in Punkten).

### 3.5 Hypothesen der Untersuchung

Wir erwarten durch den Unterricht einen signifikanten Gesamtleistungsanstieg, einen signifikanten Abbau nicht belastbarer Konzepte wie auch einen signifikanten Anstieg der Vorkonzepte. Darüber hinaus erhoffen wir, dass leistungsschwache wie auch leistungsstarke Kinder, Mädchen und Jungen, vom Unterricht profitieren.

Level	Kategorie	Inhalt	Erläuterung	Ankerbeispiele
L1	A0	kein Konzept genannt		„weiß nicht“
L2	A1	Ähnlichkeitskonzept	gleiche oder ähnliche Gegenstände ohne Nennung weiterer Eigenschaften	„Der Stock schwimmt, weil im Gievenbach hab ich das auch gesehen.“
	A2	Gewichtskonzept	Leichtes schwimmt/ Schweres geht unter	„Der Baumstamm geht unter, weil der viel zu schwer ist.“
	A3	„besondere Form“-Konzept	Schwimmen und Sinken ist abhängig von der Form der Gegenstände	„... der lange Draht hält sich auf dem Wasser fest.“
	A4	Größenkonzept	Schwimmen und Sinken ist abhängig von der Größe der Gegenstände	„Die Stecknadelspitze schwimmt, weil die so winzig klein ist.“
	A5	Wasser-drücktherunterkonzept	Gegenstände mit Löchern sinken, wenn das Wasser durch das Loch kommt (und den Gegenstand nach unten drückt)	„... weil der Holzknopf solche Löcher hat, wenn da Wasser draufkommt, dann könnte der auch untergehen.“
	A6	Luft-ziehtnach obenkonzept	Luft will nach oben	„... da sind Luftbläschen im Holz, die Luft will nach oben.“
	A7	Gewichtsverteilungskonzept	bei flachen Dingen verteilt sich das Gewicht und deshalb ist es leicht	„... eine Eisenplatte schwimmt, weil es so verteilt ist, das Eisen auf dem Wasser.“
L3	A8	Materialkonzept (auch latent)	Schwimmen und Sinken ist abhängig vom Material, aber unabhängig von Form und Größe	„Der Baumstamm schwimmt, weil er aus Holz ist.“
L4	A9	Dichtekonzept	Volumen und Masse werden in einen Zusammenhang gebracht	die Nadelspitze geht unter, ... weil die Nadelspitze, die ist so klein und 'n bisschen schwerer“

Tab. 1: Kategorisierung des Konzeptinhaltes in Teiluntersuchung 1 (Vollkörper)

## 4. Darstellung der Ergebnisse

### 4.1 Lernfortschritte in den Teiluntersuchungen

Tabelle 2 zeigt die erreichten Leistungen und die Leistungszuwächse in beiden Teiluntersuchungen. Zum Beispiel steigern die Kinder durchschnittlich bei der Vollkörper-Untersuchung ihre Präleistung, die sich aus den Punkten für die erreichten Level je Aufgabe zusammensetzt, um 48,7 %; bei der Auftriebsuntersuchung fällt der Anstieg mit 58,5 % noch deutlicher aus. Entsprechend ist bei der Vollkörper-Untersuchung der Leistungserfolg wie auch der Leistungszuwachs im Verhältnis zur maximal möglichen Leistung<sup>7</sup> (Totalleistungsmaßstab) geringer als bei der Auftriebskraft-Untersuchung. Der erreichte Wert von 60,0 % der maximal möglichen Leistung in der Vollkörper-Untersuchung lässt den Schluss zu, dass die Klasse im Durchschnitt in dieser Teiluntersuchung nicht über den Level 3 (Materialkonzept) hinausgelangt ist. Bei der Auftriebskraft-Untersuchung schneidet der Durchschnitt der Klasse mit dem erreichten Wert von 73,0 % der maximal möglichen Leistung deutlich über dem Level 3 ab.<sup>8</sup> In beiden Teiluntersuchungen ist aber ein Leistungszuwachs ersichtlich; die Kinder verbessern sich im Durchschnitt um mehr als einen halben Level.

Die Tabellen 3a und b enthalten die Häufigkeitsverteilungen ausgewählter, dominierender Konzepte. In der Vollkörper-Untersuchung (Tab. 3a) dominierte im Prätest das sog. Gewichtskonzept (alles Schwere sinkt, alles Leichte schwimmt) unter den nicht belastbaren Konzepten auf dem Level 2; neun Kinder hatten im Prätest ein (relativ) stabiles<sup>9</sup> Gewichtskonzept. Im Posttest war das Gewichtskonzept nur noch bei einem Kind (relativ) stabil; drei Kinder hatten es völlig abgebaut, weitere 12 Kinder nannten dieses Konzept nur noch in maximal der Hälfte der Aufgaben. In der Auftriebsuntersuchung (Tab. 3b) war der Abbau noch erfolgreicher: Während das Luftkonzept (ein Ball schwimmt, weil Luft darin ist und Luft zieht nach oben) noch im Prätest von 11 Kindern in mindestens einer Aufgabe genannt wird, nennen es 14 Kinder im Posttest in keiner Aufgabe.

Auch der Aufbau angemessener Konzepte verlief relativ erfolgreich (vgl. Tab. 3a und 3b): In der Vollkörper-Untersuchung verfügen sieben Kinder (von 16) im Posttest über ein stabiles Materialkonzept (Level 3); im Prätest gilt dieses nur für zwei Kinder. Allerdings haben nur zwei Kinder im Posttest ein (relativ) stabiles Dichtekonzept (Level 4) aufbauen können; bei acht Kindern ist aber zumindest ein (relativ) labiles Dichtekonzept vorhanden. Dieses Ergebnis ist beachtlich, da im Prätest nur ein Kind in zwei Aufgaben ein Dichtekonzept geäußert hat.

In der Auftriebsuntersuchung haben dagegen acht Kinder im Posttest ein (relativ) stabiles Druckkonzept auf dem höchsten Level (Level 4) erreichen können; im Prätest galt dieses nur für drei Kinder. Entsprechend höher ist hier die erreichte



durchschnittliche Leistung im Verhältnis zur maximal möglichen Leistung (vgl. Tabelle 1).

Tabelle 4 gibt die durchschnittliche Nennung einzelner Konzepte (mit Standardabweichungen) und die entsprechenden Levels in der Vollkörper- und Auftriebskraft-Untersuchung wieder, bezogen auf 18 Aufgaben bei der Teiluntersuchung „Vollkörper“ und auf vier Aufgaben bei der Teiluntersuchung „Auftriebskraft“.

Die Werte zeigen, dass in der Teiluntersuchung „Vollkörper“ bei den nicht belastbaren Konzepten das im Prätest am häufigsten vertretene Gewichtskonzept stark zurückging: Während es im Prätest noch durchschnittlich in 10,06 Aufgaben genannt wurde, trat es im Posttest nur noch in durchschnittlich 3,5 von 18 Aufgaben auf. Gleiches gilt in der Teiluntersuchung „Auftriebskraft“ für das am häufigsten genannte Luftkonzept: Dieses nicht belastbare Konzept wurde im Prätest in durchschnittlich einer Aufgabe als Erklärung herangezogen; im Posttest wurde es im Durchschnitt nur noch bei 0,18 von vier Aufgaben genannt.

Inferenzstatistische Prüfungen ergaben Folgendes: Mit dem Wilcoxon-Test konnten in beiden Teiluntersuchungen signifikante Anstiege (einseitig  $p \leq 0.01$ ) der Gesamt-Leistungsscores im Vergleich von Prä- und Posttest nachgewiesen werden. Der Abbau nicht belastbarer Konzepte (Level 2) und der Aufbau von Vorkonzepten (Level 4) waren in beiden Teiluntersuchungen ebenfalls signifikant (Wilcoxon-Test einseitig,  $p \leq 0.01$ ).

#### 4.2 Lernfortschritte im Subgruppenvergleich

Tabelle 2 enthält ebenfalls die Zuwächse zu den Anfangsleistungen in Abhängigkeit vom Geschlecht. Mädchen und Jungen unterscheiden sich in den Eingangsleistungen nur geringfügig; dieser Unterschied wird durch den Leistungszuwachs in der Vollkörper-Untersuchung verringert, in der Auftriebs-Untersuchung leicht vergrößert. Insgesamt ergeben Mann-Whitney-Tests keine signifikanten Geschlechtsunterschiede in beiden Teiluntersuchungen, weder im Leistungszuwachs noch in der erreichten Postleistung. Dieses Ergebnis ist vor dem Hintergrund geschlechtsspezifischer Differenzen im Fach Physik bei den Leistungserfolgen bzw. -fortschritten von Mädchen und Jungen an weiterführenden Schulen durchaus bemerkenswert.

Die ebenfalls in Tabelle 2 ausgewiesenen Unterschiede zwischen Extremgruppen mit hoher und geringer Leistungsstärke sind weder in der Postleistung noch im Leistungszuwachs signifikant. Allerdings zeigen die erreichten Werte einen höheren absoluten Lerngewinn der leistungsstarken Kinder in beiden Teiluntersuchungen; in der Auftriebsuntersuchung fällt die Differenz im Lernzuwachs zwischen beiden Gruppen noch stärker aus als in der Vollkörper-Untersuchung.

### 4.3 Lernfortschritte bei der Wissensintegration

Die vollzogenen Konzeptveränderungen wirken sich deutlich aus bei der Integration des Wissens in der Beantwortung der Unterrichtsfrage („Wie kommt es, dass ein eisernes Schiff nicht untergeht?“). Abbildung 1 zeigt die erreichten Konzepte im Vergleich von Prä- und Posttest: Während im Prätest (schriftlich) Konzepte auf dem nicht belastbaren Niveau (Level 2) dominierten und kein Kind den höchsten Level erreichte, formulierten im Posttest (mündlich) elf Kinder ein Konzept auf dem höchsten Level. Im Prätest dominierte das Antriebskonzept eindeutig bei acht Kindern (ein Schiff schwimmt wegen des Motors, wegen der Segel ...); auch das Hohlkörperkonzept wurde von sechs Kindern genannt (ein Schiff schwimmt, weil es hohl ist).

Im Posttest argumentierten immerhin 11 Kinder auf dem Level 4-Niveau: Das Schiff schwimmt, weil das Wasser drückt (Druckkonzept), weil das Schiff so viel Wasser verdrängt (Verdrängungskonzept), weil das Schiff leichter ist als genauso viel Wasser (Dichtekonzept). Obwohl die Kinder in den Experimenten beobachten konnten, dass keineswegs *alle* Hohlkörper schwimmen, bleibt bei drei Kindern das Hohlkörperkonzept auch im Posttest noch bestehen. Bei einem Kind ist eine Hybridvorstellung entstanden: „Weil das Schiff hohl ist, ist das Material so weit auseinandergezogen, das ist so wie die Teilchen beim Holz, deshalb schwimmt das Schiff.“

### 4.4 Analyse von Levelwechseln

Tabelle 5 zeigt die Veränderungen in den maximal erreichten Levels zwischen Prä- und Posttest, aufgeschlüsselt nach der Zahl der Kinder. Im Falle von Levelwechseln wurden die Werte durch Fettdruck hervorgehoben.

Insgesamt 11 Kinder erreichten in der Vollkörper-Untersuchung im Posttest mindestens einmal einen höheren Level als im Prätest: Zwei Kinder steigerten sich von Level 2 auf Level 3, neun Kinder von Level 3 auf Level 4. Bei fünf Kindern fand in dieser Teiluntersuchung kein Wechsel zu einem höheren Level statt; vier Kinder erreichten sowohl im Prä- wie auch im Posttest maximal Level 3; ein Kind hatte bereits im Prätest Level 4 erreicht.

Auftriebskraft-Untersuchung: Sieben Kinder haben hier Veränderungen in den höchsten erreichten Levels erreicht. Sechs Kinder hatten bereits im Prätest den höchsten Level, also das Druckkonzept, in mindestens einer Aufgabe erreicht. Zwei Kinder wechselten bei mindestens einer Aufgabe sogar von Level 2 auf Level 4. Nur vier Kinder blieben auf den im Prätest erreichten Levels 2 bzw. 3.

Integrationsfrage: Hier haben 14 von 17 Kindern einen Wechsel im Level erreicht. Sechs Kinder wechselten sogar von Level 2 auf Level 4. Insgesamt 11 von 17 Kindern haben den höchsten Level erreicht; nur ein Kind ist auf dem nicht belastbaren Level 2 verblieben.

Die Ergebnisse der Tabelle dürfen allerdings nicht überinterpretiert werden, da es sich nur um die in mindestens einer Aufgabe erreichten Levels handelt. Eine Aussage über die Stabilität des Wechsels von einem Level zum anderen in Bezug auf die gestellten Aufgaben ist aus diesen Daten nicht zu entnehmen. Allerdings zeigen die Daten, dass die höheren Levels von der Mehrzahl der Kinder zumindest in einer Aufgabe genannt werden. Insbesondere die Auswertung der Integrationsfrage deutet auf einen sehr guten Lernerfolg bei 11 von 17 Kindern hin.

#### 4.5 Anteil der Konzeptbewegungen am Lernfortschritt

Lernfortschritte können durch verschiedenartige Konzeptveränderungen erreicht werden. Einen großen Anteil am Lernfortschritt nimmt der Abbau nicht belastbarer Konzepte ein. In unseren Teiluntersuchungen waren vor allem der Abbau des Gewichtskonzeptes bei der Vollkörper-Untersuchung und der Abbau des Luftkonzeptes bei der Auftriebskraft-Untersuchung zu beobachten.

Lernfortschritte machen Kinder auch, wenn angemessenere Konzepte, das heißt ausbaufähige Konzepte oder Vorkonzepte, neu aufgebaut werden.

Werden ausbaufähige Konzepte oder Vorkonzepte im Posttest häufiger als im Prätest genannt, so sprechen wir von einer Verstärkung angemessenerer Konzepte. Auch diese Verstärkungen führen zu einem Lernfortschritt.

Die Abbildung 2 zeigt, dass in der Vollkörper- und Auftriebskraft-Untersuchung der Leistungszuwachs im Posttest primär auf den Aufbau von neuen Konzepten zurückzuführen ist (Neuaufbau); an zweiter Stelle folgt der Abbau von Präkonzepten, an dritter die Verstärkung bereits vorhandener Konzepte.

In der Auftriebskraft-Untersuchung steigert allein der Neuaufbau von Konzepten das Ausgangsniveau der Leistung um 86,6 %. Die Prä-Leistung wurde also überwiegend durch neu aufgebaute Konzepte im Posttest gesteigert. Neu aufgebaut wurden belastbare Konzepte (ein Kind) und Vorkonzepte (sechs Kinder) (vgl. Tabelle 5). Weitergehende Analysen zeigen, dass im Falle solcher Levelwechsel die höherwertigen Konzepte bei über einem Drittel der Post-Aufgaben genannt wurden.

In der Vollkörper-Untersuchung konnte das Ausgangsniveau durch den Neuaufbau von Konzepten um 62,6 % gesteigert werden. Neu aufgebaut wurden in dieser Teiluntersuchung belastbare Konzepte (zwei Kinder) und Vorkonzepte (neun Kinder) (vgl. Tabelle 5). Auch hier zeigen weitergehende Analysen, dass der Neuaufbau relativ stabil ist: In der Gruppe der Wechsler von Level 2 nach 3 entfallen im Posttest durchschnittlich 15,5 von 18 Aufgaben auf Level 3. Daraus kann geschlossen werden, dass das Materialkonzept von den entsprechenden Levelwechslern relativ stabil aufgebaut wurde. Der Wechsel von Level 3 nach 4 führte allerdings zu weniger stabilen Nennungen des Vorkonzepts; im Mittel wird nur bei 3,4 von 18 Aufgaben Level 4 erreicht.

Zusammenfassend lässt sich festhalten, dass in beiden Teiluntersuchungen Konzepte auf Level 2 (nicht belastbare Konzepte) zugunsten von Konzepten auf Level 3 (ausbaufähige Konzepte) bzw. auf Level 4 (Vorkonzepte) und Konzepte auf Level 3 zugunsten von Konzepten auf Level 4 abgebaut wurden, wobei die Stabilität der neu aufgebauten Konzepte unterschiedlich ausgeprägt war.

Der Leistungsanstieg lässt sich ebenfalls auf die Verstärkung von Präkonzepten im Posttest zurückführen, wenn auch in geringerem Maße. In der Vollkörper-Untersuchung ist die Verstärkung von Konzepten mit einem Leistungsanstieg um 41,8 % verbunden und damit relativ bedeutsam. Tabelle 4 zeigt, dass die Verstärkung in erster Linie auf die fünf Kinder zurückzuführen ist, welche das Materialkonzept (Level 3) und das Vorkonzept (Level 4) bereits im Prätest genannt hatten. In der Auftriebskraft-Untersuchung fällt das Niveau der Leistungssteigerung durch Verstärkung geringer aus (23,4 %), hier verstärkten sechs Kinder das bereits vorhandene Vorkonzept „Druck“, drei Kinder verstärkten belastbare Konzepte (vgl. Tabelle 4).

Auffällig ist auch der hohe Anteil der abgebauten Konzepte am Leistungserfolg in beiden Teiluntersuchungen. Auch wenn dieser Abbau nicht bei allen Kindern stabil war, weist doch der hohe Anteil der abgebauten Konzepte am Leistungserfolg auf erhebliche Konzeptwechsel hin.

## 5. Interpretation der Ergebnisse

Die Hypothesen der Untersuchung wurden im Prä-Postvergleich durch den signifikanten Anstieg des Leistungsscores, den signifikanten Anstieg der Vorkonzepte und den signifikanten Abbau nicht belastbarer Konzepte bestätigt.

Aufgrund der deutlichen Lernfortschritte kann vermutet werden, dass die untersuchte Gruppe mit der Unterrichtsthematik nicht überfordert war. Der Einfluss von Geschlecht und Leistungsstärke ist gering; auch Mädchen profitieren von dem Unterricht, auch leistungsschwache Kinder machen erhebliche Lernfortschritte. Zwar kann der Unterricht die bestehenden Unterschiede nicht ausgleichen, doch sind die auftretenden Schereneffekte für eine derart anspruchsvolle inhaltliche Thematik erstaunlich gering.

Besonders eindrucksvoll ist der Lernerfolg bei der Integrationsfrage. Während im Prätest kein Kind den höchsten Level erreichte, äußerten 11 Kinder (von 17 Kindern) im Posttest relativ angemessene Vorkonzepte. Dieses Ergebnis ist beachtlich, wenn man bedenkt, dass der Posttest zehn Wochen nach dem Unterricht durchgeführt wurde.

Die durchweg besseren Ergebnisse in der Teiluntersuchung „Auftriebskraft“ können vielleicht darauf zurückgeführt werden, dass das Vorkonzept „Druck“ für Grundschul Kinder eine hohe Plausibilität und Fruchtbarkeit durch entsprechende

Versuche und Alltagserfahrungen besitzt, so dass Konzeptwechselprozesse erleichtert werden. Das Konzept der Dichte dagegen scheint dem Verständnis der Grundschul Kinder nicht in gleichem Maße zugänglich zu sein; hier wären ggfs. weitere Lernhilfen zu erproben.

Neben dem Neuaufbau von Konzepten waren maßgeblich der Abbau von lernbehindernden Konzepten am Leistungsanstieg beteiligt. Das Ziel des Unterrichts, aktive Konzeptwechselprozesse anzuregen, kann in hohem Maße als erreicht gelten, auch wenn nicht alle Kinder Konzepte auf dem höchsten Level erreicht haben. Die Ergebnisse zeigen in Übereinstimmung mit der Conceptual-Change- und Präkonzeptforschung auch, dass Konzeptveränderungen sehr fragile Prozesse sind. Nur wenigen Kindern gelingt ein völlig stabiler Konzeptwechsel, häufig verursachen situative Bedingungen noch die Labilität von Konzepten.

Zu vermuten ist aber, dass ein Übergang von stabilen, nicht belastbaren Konzepten zu ausbaufähigen Konzepten oder Vorkonzepten auch dann einen wichtigen Lernfortschritt beinhaltet, wenn dieser Wechsel noch nicht stabil ist. Die Instabilität des Denkens im Lernprozess scheint geradezu eine Vorbedingung für den stabilen Aufbau angemessenerer Konzepte zu sein.

## 6. Zur Methode der Interviewdiagnostik

Das Ziel der Untersuchung, ein quantifizierbares Verfahren für die Auswertung inhaltsreicher, „weicher“ Daten zu entwickeln, konnte erreicht werden.

Der Vergleich von Prä- und Posttest-Ergebnissen ermöglicht eine Diagnostik erreichter Lernfortschritte auf der Basis von Ausgangsvorstellungen. So wird der individuell erreichte Wissensstand nicht nur am Endziel (also am höchsten erreichbaren Level), sondern auf der Basis des vor dem Unterricht vorhandenen Wissens gemessen. Die vorgestellte Methode der Wissensdiagnostik eignet sich daher insbesondere, um Konzeptwechsel und Konzeptdifferenzierungen zu erfassen.

Mit der Differenzierung von Levels wird es möglich, auch graduelle Annäherungen im Verstehen messbar zu machen und Lernfortschritte als Weiterentwicklung von Ausgangsvorstellungen differenziert zu erfassen.

In Bezug auf die Validität des Verfahrens bleiben dennoch einige Fragen offen. Zwar scheinen Interviews schriftlichen freien Äußerungen überlegen zu sein, weil die mündliche Kommunikation Grundschulkindern leichter fällt als schriftliche Äußerungen. Allerdings kann nicht aus dem mündlich Geäußerten sicher auf interne Kognitionen geschlossen werden. Zumindest muss berücksichtigt werden, dass die Befragten nicht ihr gesamtes Wissen in jeder Aufgabe präsentieren. Zum Beispiel bemerkten wir in den Interviews, dass die Kinder trotz aufgebautem Dichtekonzept in den Aufgaben zum Schwimmen und Sinken von Vollkörpern mit dem Materialkonzept antworteten; vielleicht weil es dem Allgemeinverständnis

nis entspricht, dass Holz in der Regel schwimmt, Metallstücke aber untergehen. Es ist also durchaus möglich, dass die Befragten zum Beispiel über das Dichtekonzept verfügten, auch wenn sie dieses nicht äußern. In weiteren Untersuchungen sollen deshalb gezielte Fragen nach dem Aufbau dieses Konzeptes in den Interviewleitfaden aufgenommen werden.

Eine weitere Schwierigkeit liegt darin, die sprachlichen Äußerungen der Kinder angemessen zu deuten. So war es z. B. nicht leicht, Äußerungen mit dem Begriff „schwer“ zu deuten, da Kinder mit diesem Wort sowohl die absolute Masse als auch die Dichte im Sinne eines „schweren“ Materials bezeichnen. Hier wären gezielte Nachfragen erforderlich gewesen.

Einen anderen Unsicherheitsfaktor bilden sogenannte Ad-Hoc-Konstruktionen. In der Interviewsituation wurden in manchen Situationen Antworten genannt, die auf Verlegenheitsantworten oder auch auf sozial erwünschte Antworten hindeuten. Diese Unsicherheiten lassen es sinnvoll erscheinen, nach Verfahren zu suchen, die eine Validierung der ermittelten Ergebnisse ermöglichen.<sup>10</sup> Möglich wären z. B. ergänzende schriftliche Multiple-Choice-Tests oder auch die Überprüfung der Gültigkeit der aus den Interviews diagnostizierten Konstruktionen in nachträglichen Gesprächen der Versuchsleiter mit den Befragten.

Die vorgestellte Pilotstudie beschränkt sich aus methodischen Gründen auf die Diagnostik kognitiver Lernfortschritte; nicht-kognitive Zielsetzungen blieben in dieser Erhebung (nicht im Unterricht) unberücksichtigt. Die beobachteten Reaktionen der Schülerinnen und Schüler deuteten jedoch auf eine hohe Lernzufriedenheit hin. In weiterführenden Untersuchungen soll ergänzend die Vereinbarkeit von kognitiven und nicht-kognitiven Zielsetzungen im Sinne einer multi-kriterialen Zielerreichung geprüft werden.

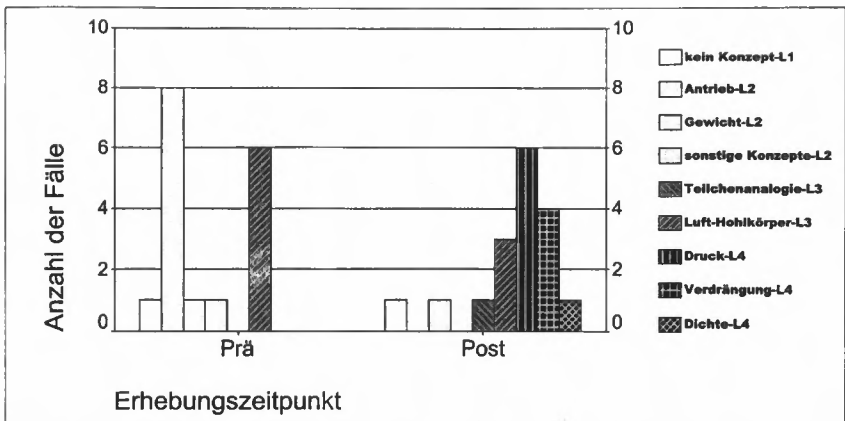


Abb. 1: Konzepte und Level bei der Integrationsfrage im Prä-Post-Vergleich (N=17)

Mittlerer Leistungserfolg im Verhältnis zur Prä-Basis (Gesamtdurchschnitt); (Mittlerer Leistungserfolg im Verhältnis zu totalen Leistungs-Maßstäben <sup>2)</sup> : in Prozent <sup>3</sup>					
Teilerhebung	gesamt	männlich (n = 9)	weiblich (n = 7;8) <sup>4</sup>	hohe Leistungsstärke (n = 6;7) <sup>4</sup>	geringe Leistungsstärke (n = 7)
Vollkörper Prä	100,0 (40,3)	107,6 (43,2)	90,9 (36,5)	115,3 (46,3)	89,0 (35,7)
Vollkörper Zuwachs Post	48,7 (19,6) 148,7 (59,9)	47,7 (19,1) 155,3 (62,3)	50,1 (20,1) 141,0 (56,6)	55,4 (22,2) 170,7 (68,5)	41,5 (16,7) 130,5 (52,4)
Auftriebskraft Prä	100,0 (46,1)	102,5 (47,2)	97,2 (44,8)	105,9 (48,8)	90,4 (41,7)
Auftriebskraft Zuwachs Post	58,5 (27,0) 158,5 (73,1)	62,3 (28,7) 164,8 (75,9)	54,3 (25,0) 151,5 (69,8)	74,9 (34,5) 180,8 (83,3)	43,9 (20,2) 134,3 (61,9)

- 1) nach Noten in Mathematik und Deutsch (hoch: 1+ bis 2+; leistungsschwach: 3- bis 4-): Die mittlere Gruppe (n = 3) mit den Noten 2 bis 3 bleibt unbeachtet (Extrem-Gruppenvergleich)
- 2) in Klammern; Skalen-Wertebereich für Vollkörper: 0 bis 54, für Auftriebskraft: 0 bis 12 Punkte
- 3) Abweichungen durch Rundungen
- 4) In der Vollkörper-Erhebung wurde ein leistungsstarkes Mädchen weniger befragt als in der Auftriebskraft-Untersuchung.

Tab. 2: Vollkörper-(N = 16) und Auftriebskraft-Untersuchung (N = 17) Prä-Leistungen & Post-Zuwächse/Gesamtleistungen (Gesamt, nach Geschlecht und nach Leistungsstärke [Mathematik/Deutsch])<sup>1)</sup>

Konzepte		Zahl der Nennungen																		
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Gewichtskonzept	Prä	0	1	1	1	0	0	1	0	0	3	1	3	0	0	0	3	1	0	1
	Post	3	2	3	2	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0
Materialkonzept	Prä	2	3	1	3	2	2	0	1	0	0	0	1	0	0	1	0	0	0	0
	Post	1	2	0	1	0	1	1	1	1	1	1	2	0	0	1	2	1	0	0
Dichtekonzept	Prä	15	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Post	6	3	1	0	1	1	0	0	2	0	1	0	1	0	0	0	0	0	0

Tab. 3a: Häufigkeitsverteilung der Nennungen ausgewählter Konzepte im Prä- und Posttest in der Teiluntersuchung „Vollkörper“ (N = 16, 18 Aufgaben)

		Zahl der Nennungen				
		0	1	2	3	4
Luftkonzept	Prä	6	6	4	1	0
	Post	14	3	0	0	0
Wasserdruckkonzept	Prä	11	3	0	3	0
	Post	5	1	3	5	3

Tab. 3b: Häufigkeitsverteilung der Nennungen ausgewählter Konzepte im Prä- und Posttest in der Teiluntersuchung „Auftriebskraft“ (N = 17, 4 Aufgaben)

Teilerhebung	Level	Konzepte	Prätest Mittelwert (Standardabweichung)	Posttest Mittelwert (Standardabweichung)
Vollkörper (N = 16; 18 Aufgaben)	Level 1	Kein Konzept genannt	0,50 (0,97)	0,44 (0,73)
	Level 2	Ähnlichkeits-Konzept	0,81 (1,11)	0,44 (1,03)
		Gewichts-Konzept	10,06 (5,11)	3,50 (3,25)
		Besondere-Form-Konzept	1,94 (2,82)	1,50 (1,79)
		Größen-Konzept	0,25 (0,77)	0,01 (0,25)
		Wasser drückt herunter-Konzept	0,25 (0,77)	0,50 (1,10)
		Gewichtsverteilungs-Konzept	0,01 (0,25)	0,01 (0,25)
		<b>Gesamt (Level 2)</b>	<b>13,32 (4,18)</b>	<b>5,96 (4,20)</b>
Level 3	Material-Konzept	4,00 (3,89)	8,25 (5,30)	
Level 4	Vorkonzept „Dichte“	0,13 (0,50)	3,25 (4,10)	
Auftriebskraft (N = 17; 4 Aufgaben)	Level 1	Kein Konzept genannt	0,65 (0,86)	0,24 (0,56)
	Level 2	Eigenkraft-Konzept	0,35 (0,79)	0,29 (0,69)
		Luft-Konzept	1,00 (0,94)	0,18 (0,39)
		Andere Konzepte	0,53 (1,01)	0,29 (0,59)
		<b>Gesamt (Level 2)</b>	<b>1,88 (1,22)</b>	<b>0,76 (0,97)</b>
	Level 3	Hohlraum-Konzept	0,18 (0,53)	0,12 (0,33)
		Leichtigkeit im Wasser-Konzept	0,41 (0,51)	0,35 (0,49)
		Widerstands-Konzept	0,18 (0,53)	0,53 (1,01)
		<b>Gesamt (Level 3)</b>	<b>0,77 (0,75)</b>	<b>1,00 (1,06)</b>
	Level 4	Druckkonzept	0,71 (1,16)	2,00 (1,54)

Tab. 4: Konzepte und Levels: Mittelwerte und Streuungen in der Vollkörper- und Auftriebskraft-Untersuchung



Teilerhebung erreichte Level-Spitze	Vollkörper (N=16)	Auftriebskraft (N=17)	Integrationsfrage Eiserner Hohlkörper (N=17)
Höchster Prä-Level	2 3 3 4	2 2 2 3 3 4	1 2 2 2 3 3
Höchster Post-Level	3 3 4 4	2 3 4 3 4 4	1 2 3 4 3 4
Anzahl der Fälle	2 4 9 1	1 1 2 3 4 6	1 1 3 6 1 5

1) Werte fett markiert im Fall eines Levelwechsels

Tab. 5: Häufigkeiten von Level-Spitzen

Prä-Post-Vergleich nach Teilerhebungen<sup>1</sup> (höchster, mindestens einmal erreichter Level in den gestellten Aufgaben)

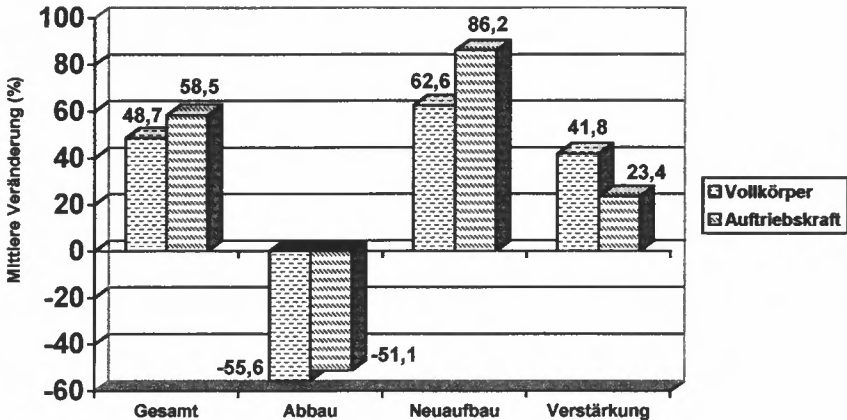


Abb. 2: Vollkörper (N=16) und Auftriebskraft-Erhebung (N=17)

## Anmerkungen

- <sup>1</sup> An der Auswertung waren ebenfalls beteiligt: Inken Büsch, Kristina Dohm, Nadine Prenger-Berninghoff, Ute Westerwalbesloh. Die Erhebungen wurden von Studierenden des Sachunterrichtsstudienganges im Rahmen von Examensarbeiten durchgeführt.
- <sup>2</sup> Lehr-Lernumgebungen, die nach den genannten Prinzipien gestaltet werden, haben viele Gemeinsamkeiten (vgl. Möller 2001 b).
- <sup>3</sup> Der Unterricht wurde unter annähernd normalen Klassenbedingungen mit 26 Schülerinnen und Schülern durchgeführt. 17 dieser Kinder wurden zufällig für die individuellen Lerndiagnosen ausgewählt.

- 4 Der Unterricht wurde von einer Forschergruppe in unserem Institut vorbereitet (Walburga Henry, Ingrid Nachtigäller, Susanne Lips, Kornelia Möller). Ein Mitglied der Gruppe führte den Unterricht durch.
- 5 In der Erhebung wurden 24 (Teiluntersuchung I) bzw. 6 Aufgaben (Teiluntersuchung II) gestellt; sechs Aufgaben in der Teiluntersuchung I, zwei Aufgaben in der Teiluntersuchung II mussten aus der Auswertung herausgenommen werden, da Missings durch Missachten des Interviewleitfadens durch die Interviewer entstanden waren.
- 6 Die hier genannten Aussagen sind eigensprachliche Formulierungen der Kinder.
- 7 Als maximal mögliche Leistung wird die maximal erreichbare Punktzahl bezeichnet. In der Vollkörperuntersuchung handelt es sich bei 18 Aufgaben auf Level 4 um maximal 54 Punkte, in der Auftriebsuntersuchung bei vier Aufgaben auf Level 4 um maximal 12 Punkte (pro Aufgabe gibt es drei Punkte bei erreichtem Level 4).
- 8 Wäre bei allen Aufgaben Level 3 erreicht worden, so entspräche das 66,6 % der maximal erreichbaren Leistung.
- 9 Als relativ stabil werten wir ein Konzept, wenn es in mehr als der Hälfte der gestellten Aufgaben als Konzept genannt wird. Als relativ labil stufen wir ein Konzept ein, wenn es in der Hälfte oder weniger als der Hälfte der Aufgaben als Konzept genannt wird.
- 10 In unserer Untersuchung führten wir probeweise zusätzliche qualitative Lerndiagnosen durch, die den Lernerfolg der Schülerinnen und Schüler im Hinblick auf die Integrationsfrage aufgrund von Unterrichtsbeobachtungen und Zwischeninterviews beschrieben. Die Ergebnisse der Lerndiagnosen waren nahezu identisch mit den Ergebnissen der quantifizierenden Interviewdiagnostik.

## Literatur

- Banholzer, A. (1936). *Die Auffassung physikalischer Sachverhalte im Schulalter. Inaugural-Dissertation zur Erlangung der Doktorwürde einer hohen philosophischen Fakultät der Eberhard Karls-Universität zu Tübingen*. Stuttgart: Widmann.
- Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (Hrsg.), *The epigenesis of mind. Essays on biology and cognition* (S. 257-291). Hillsdale, N. J.: Erlbaum.
- Duit, R. (1995). Zur Rolle der konstruktivistischen Sichtweise in der naturwissenschaftsdidaktischen Lehr-Lernforschung. *Zeitschrift für Pädagogik*, 41, 905-923.
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41, 867-887.
- Janke, B. (1995). Entwicklung naiven Wissens über den physikalischen Auftrieb: Warum schwimmen Schiffe? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 17, 122-138.
- Kircher, A. & Rückel, B. (1992). Warum Eisenschiffe schwimmen. In K. H. Wiebel (Hrsg.), *Zur Didaktik der Physik und Chemie. Probleme und Perspektiven. Vorträge auf der Tagung für Didaktik der Physik, Chemie in Hamburg, September 1991* (S. 101-103). (= Gesellschaft für Didaktik der Chemie und Physik, Bd. 12) Alsbach, Bergstraße: Leuchtturm-Verlag.
- Köhnlein, W. (1999). Vielperspektivität und Ansatzpunkte naturwissenschaftlichen Denkens. Analyse von Unterrichtsbeispielen unter dem Gesichtspunkt des Verstehens. In W. Köhnlein, B. Marquardt-Mau & H. Schreier (Hrsg.), *Vielperspektives Denken im Sachunterricht* (S. 88-124). (= Forschungen zur Didaktik des Sachunterrichts, Bd. 3) Bad Heilbrunn: Klinkhardt.
- Lauterbach, R. (1992). Naturwissenschaftlich orientierte Grundbildung im Sachunterricht. In Riquarts, K. u. a. (Hrsg.), *Naturwissenschaftliche Bildung in der Bundesrepublik Deutschland. Bd. 3: Didaktik* (S. 191-256). Kiel: IPN.

- Mayring, P. (1995): *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 5. Aufl. Weinheim: Deutscher Studien Verlag.
- Möller, K. (1999). Konstruktivistisch orientierte Lehr-Lernprozessforschung im naturwissenschaftlich-technischen Bereich des Sachunterrichts. In W. Köhnlein, B. Marquardt-Mau & H. Schreiber (Hrsg.), *Vielperspektives Denken im Sachunterricht* (S. 125-191). (= Forschungen zur Didaktik des Sachunterrichts, Bd. 3) Bad Heilbrunn: Klinkhardt.
- Möller, K. (2001a). Wissenserwerb und Wissensqualität im naturwissenschaftsbezogenen Sachunterricht. In J. Kahlert & E. Inckemann (Hrsg.), *Wissen, Können und Verstehen – über die Herstellung ihrer Zusammenhänge im Sachunterricht* (S. 115-126). (= Probleme und Perspektiven des Sachunterrichts, Bd. 11) Bad Heilbrunn, Obb.: Klinkhardt.
- Möller, K. (2001b): Genetisches Lehren und Lernen – Facetten eines Begriffs. In: D. Cech, B. Feige, J. Kahlert, G. Löffler, H. Schreiber, H.-J. Schwier & U. Stoltenberg (Hrsg.), *Die Aktualität der Pädagogik Martin Wagenscheins für den Sachunterricht. Walter Köhnlein zum 65. Geburtstag* (S. 15-30). Bad Heilbrunn, Obb.: Klinkhardt.
- Möller, K. (2001c). Konstruktivistische Sichtweisen für das Lernen in der Grundschule? In K. Czerwenka, K. Nölle & H.-G. Roßbach (Hrsg.), *Jahrbuch Grundschulforschung, Bd. 4*. (S. 16-31) Opladen: Leske und Buderich.
- Piaget, J. & Inhelder, B. (1977). *Von der Logik des Kindes zur Logik des Heranwachsenden*. Olten u. a.: Walter.
- Reinmann-Rothmeier, G. & Mandl, H. (1999). *Unterrichten und Lernumgebungen gestalten*. (= Forschungsbericht, Bd. 60, überarbeitete Fassung) München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.
- Renkl, A., Gruber, H. & Mandl, H. (1996). *Situated learning in instructional settings: From euphoria to feasibility*. (= Forschungsbericht Nr. 74) München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.
- Thiel, S. (1990). GrundschulKinder zwischen Umgangserfahrung und Naturwissenschaft. In M. Wagenschein (Hrsg.), *Kinder auf dem Wege zur Physik* (S. 90-180). (= Pädagogische Bibliothek, Bd. 2) Weinheim, Basel: Beltz.
- Weinert, F. E. & Hemke, A (1995). Learning from wise mother nature or big brother instructor: The wrong choice as seen from an educational perspective. *Educational Psychologist*, 41, 135-142.
- White, R. & Gunstone, R. (1999). Alternativen zur Erfassung von Verstehensprozessen. *Unterrichtswissenschaft*, 27, 128-134.

## **Selbstbestimmungsempfinden in offenen Lernsituationen Eine Pilotstudie zum Sachunterricht**

### **1. Einleitung und Fragestellung**

Die Feststellung, dass es zur Öffnung von Unterricht zwar viele interessante, kreative und anregende Erfahrungsberichte gibt, dass jedoch die empirische Forschungslage immer noch als unbefriedigend anzusehen ist, ist nicht neu, doch leider gilt sie immer noch (vgl. z. B. Einsiedler 1997; von Saldern 1998; Hanke 2001). Es gibt zwar einige spannende Einzeluntersuchungen mit z. T. sehr bemerkenswerten Ergebnissen, von kumulativem empirischen Wissen kann jedoch bei weitem noch nicht gesprochen werden. Dies gilt in besonderer Weise für den Sachunterricht (vgl. Einsiedler in diesem Band). Denn auch wenn sich manche Ergebnisse aus anderen Fachgebieten übertragen lassen, so ist es m. E. dennoch erforderlich, aufgrund der spezifischen Besonderheiten des Sachunterrichts auch eigenständige Forschung zu betreiben.

Neben der geringen Quantität der Forschung zu Öffnung von Unterricht gibt es jedoch auch noch einen inhaltlichen Grund, warum über offene Lernsituationen vergleichsweise wenig gesichertes empirisches Wissen existiert: Hildegard Kasper hat bereits 1989 darauf hingewiesen, dass Öffnung von Unterricht nicht als „Totalkonzept“ zu sehen ist. Unterricht ist also nicht ‚ganz oder gar nicht‘ ‚offen‘. Definiert man Öffnung von Unterricht ganz allgemein dadurch, dass unterrichtsrelevante Entscheidungen von Schüler/innen getroffen oder zumindest mitgetroffen werden, so wird deutlich, dass es verschiedene Formen einer solchen Öffnung gibt (vgl. z. B. Hartinger 1997, S. 152, Peschel 1996). Peschel unterscheidet z. B. zwischen der Grundausswahl des Inhaltes, der individuellen Inhaltswahl über die konkreten Inhalte einer Stunde, der Methodenwahl, der Differenzierung, der Wahl von Sozialformen, der Wahl von Zeitpunkt und -dauer sowie der Wahl des Arbeitsortes (1996, S. 19). Sämtliche dieser Entscheidungen können von dem/r Lehrer/in oder von den Schüler/innen getroffen werden. Bei der systematischen Betrachtung offener Lernsituationen muss man aus diesem Grund differenziert unterscheiden: Es ist nicht zu erwarten, dass z. B. die freie Zeiteinteilung durch die Kinder die

gleichen Effekte zeigt wie eine freie Auswahl der Themen. Dieses Problem wurde schon in der Metaanalyse von Giaconia & Hedges 1982 deutlich (vgl. dazu auch Schöll 1996).

Ein erster Schritt, um die Datenlage transparenter zu machen, wird getan, wenn die jeweiligen Aspekte der Öffnung präzise dargelegt werden. Dies geschieht gerade in neueren Untersuchungen durchwegs (vgl. z. B. Lipowsky 1999, S. 124). Ein zweiter Schritt müsste jedoch sein, gezielt einzelne Bereiche der Öffnung herauszunehmen und zu untersuchen. Gerade *die systematische Variation genau einer Variablen* könnte helfen, die Bedeutung einzelner methodischer Maßnahmen einzuschätzen. Und damit könnten solche Daten vermutlich wohl auch wertvolle Hilfen für täglich zu treffende methodische Einzelentscheidungen von Lehrer/innen sein. Gleichzeitig könnte die Forschung zur Öffnung von Unterricht mehr Transparenz gewinnen.

Eine zentrale Entscheidung unterrichtlichen Handelns ist die über die *Inhalte des Unterrichts*. Wenn Schüler/innen die Themen des Unterrichts mitbestimmen, so ist dies eine klare Öffnung des Unterrichts. In Phasen Freier Arbeit sind z. B. die Inhalte häufig sehr frei; hier wird zumeist auch nicht vorgegeben, aus welchem Fachgebiet die Schüler/innen ihre Arbeit wählen sollten. Gerade im Sachunterricht ist es aber auch durchaus gut vorstellbar, dass zwar ein bestimmtes Rahmenthema vorgegeben wird, dass dann jedoch die Kinder Fragen stellen oder Unterthemen nennen, die sie als besonders spannend empfinden. Diese Fragen können dann zum Ausgangspunkt für den Unterricht werden (vgl. Götz 1991). Es ist plausibel zu vermuten, dass ein solches Aufgreifen von Interessen der Kinder zu positiven Effekten führt – systematische empirische Daten fehlen für den Sachunterricht jedoch.

Ein gängiges Argument – gewonnen nicht zuletzt aus positiven Erfahrungsberichten – für die Öffnung von Unterricht ist, dass die Schüler/innen in offenen Lernsituationen besonders motiviert sind (vgl. z. B. die Lehreraussagen bei Jürgens 1999). Betont wird v. a. die *intrinsische Motivation* (also aus der Sache selbst erwachsen) sowie das *Interesse an der Thematik*. Interesse und intrinsische Motivation sind wichtige Variablen für den Lernerfolg (vgl. die Metaanalysen von Schiefele & Schreyer 1994; Schiefele, Krapp & Schreyer 1993; zur theoretischen Konzeptionierung von Interesse vgl. z. B. Krapp 1999). Bzgl. der verstärkten (intrinsischen) Motivation bzw. der positiven Interessenentwicklung bei Öffnung von Unterricht gibt es einige empirische Evidenzen (vgl. z. B. Swann & Pittman 1977; Zuckerman, Porac, Lothin, Smith & Deci 1978; Green & Foster 1986; Monty & Perlmutter 1987; Deci, Eghari, Patrick & Leone 1994; Hartinger 1997). Zudem können die theoretischen Begründungen für die positive Wirkung von Öffnung des Unterrichts auf den Aufbau und Erhalt intrinsischer Motivation als recht fundiert gelten (vgl. z. B. Deci & Ryan 1993; Ryan, Deci & Grolnick 1995). Aber

auch hier fehlen Untersuchungen für die Grundschule, die gezielt einzelne Elemente der Öffnung betrachten. Die vorliegenden Studien vergleichen entweder recht allgemein ‚offeneren Unterricht‘ mit ‚traditionellerem‘ (vgl. z. B. Hartinger 1997) oder isolieren zwar einzelne Aspekte wie z. B. die Wahlfreiheit, dies jedoch unter Laborbedingungen (vgl. z. B. Swann & Pittman 1977).

Intrinsische Motivation sollte dazu führen, dass die Schüler/innen entsprechend engagiert im Unterricht mitarbeiten. Mitarbeit im Unterricht bzw. die *Lernzeitnutzung (time on task)* gilt allgemein als eine der wichtigsten Bedingungen für Lernerfolg und ist eine der am intensivsten erforschten Unterrichtsvariablen (für einen Überblick vgl. Karweit 1989). Die Ergebnisse der Forschung zu ‚time on task‘ scheinen jedoch auf den ersten Blick eher gegen eine Öffnung von Unterricht zu sprechen (vgl. hierzu auch Heckt & Jürgens 1999), denn gut organisierte Instruktion führt in Grundschulklassen zu deutlich höheren Aufmerksamkeitszeiten als selbstständiges Arbeiten der Schüler/innen (vgl. z. B. Burns 1984). Allerdings sind diese Studien allesamt im Rahmen traditionellen Unterrichtens durchgeführt worden. Die Frage, ob z. B. die Mitarbeit der Schüler/innen höher ist, wenn sie im Unterricht selbstgewählte Themen bearbeiten, ist demnach noch nicht beantwortet.

Die vorliegenden Untersuchungen zu offenen Lernsituationen beziehen sich immer auf eine ‚objektiv beobachtbare‘ Variation des Unterrichts. Die subjektiv empfundene *Wahrnehmung von Selbstbestimmung* oder Autonomie wird in diesen Studien nicht berücksichtigt. Dies ist insofern problematisch, da z. B. genau diese Wahrnehmung von Selbstbestimmung als zentrales Element für den Aufbau und den Erhalt intrinsischer Motivation gilt (vgl. Deci & Ryan 1993; vgl. auch Kasser & Ryan 1993, 1996). Und umgekehrt ist zu konstatieren, dass die Forschergruppe um Edward Deci, die in vielen Studien genau dieses subjektive Selbstbestimmungsempfinden von Schüler/innen oder Studierenden erfasste, bislang keine Untersuchungen zu Unterrichtsmethoden durchgeführt hat. Einige Arbeiten gehen zwar ähnliche Fragestellungen an; es wird dort jedoch das Selbstbestimmungsempfinden lediglich indirekt über den Motivationsstil erhoben (vgl. Ryan & Connell 1989); zudem sind stärker die Bedeutung von Leistungsüberprüfungen (vgl. Grolnick & Ryan 1987) bzw. die Lehrereinstellungen (vgl. Deci, Schwartz, Sheinman & Ryan 1981; Reeve, Bolt & Cai 1999) im Fokus und weniger unterrichtsmethodische Variationen.

Mit den Untersuchungen zur Öffnung von Unterricht sowie mit den Arbeiten zum Selbstbestimmungsempfinden als grundlegendem Bedürfnis existieren also zwei Forschungsrichtungen, die sich gegenseitig ergänzen können, da ihre Aussagen durchaus kompatibel sind. Die Schnittstelle, also die Frage, inwieweit sich eine Öffnung des Unterrichts auf das Selbstbestimmungsempfinden von Schüler/innen auswirkt, ist empirisch jedoch noch nicht geklärt. Es ist zwar plausibel, dass Zu-

sammenhänge bestehen; differenziertere Erkenntnisse sind jedoch schon aus dem Grund erforderlich, da festzustellen ist, dass sich nicht alle Kinder in einer Klasse als gleichermaßen selbstbestimmt empfinden (vgl. Hartinger 2001). Es stellt sich damit die Frage, ob sich solch individuelle Unterschiede innerhalb derselben Unterrichtssituation auswirken, indem Schüler/innen, die sich – unabhängig von dem durchgeführten Unterricht – als sehr selbstbestimmt empfinden, z. B. mehr Interesse an den Themen des Unterrichts entwickeln als diejenigen, die sich – aus welchen Gründen auch immer – als wenig selbstbestimmt einschätzen.

Die vorliegende Studie versucht nun, für folgende Fragen erste empirische Hinweise zu finden:

1. Führt eine thematische Öffnung des Unterrichts im Vergleich zu einem Unterricht, in dem vorgegebene Themen behandelt werden, zu
  - a) höherem Selbstbestimmungsempfinden der Kinder?
  - b) verstärkter Mitarbeit der Schüler/innen?
  - c) höherem Interesse an der Thematik des Unterrichts?
2. Wirkt sich – unabhängig von der Variation des Unterrichts – das subjektive Selbstbestimmungsempfinden der Schüler/innen auf das Interesse an der Thematik des Unterrichts aus?

## 2. Methode

### *a) Untersuchungsaufbau*

Die Untersuchung ist als Pilotstudie angelegt. Dies wird schon an der relativ kleinen Stichprobe deutlich: Insgesamt nahmen 164 Schüler/innen (84 Mädchen und 80 Jungen) der dritten Jahrgangsstufe teil. Die Kinder waren zwischen 8 und 10 Jahre alt. Es waren sieben Klassen aus drei verschiedenen Schulen in und um Regensburg beteiligt. Der Unterricht wurde von mir durchgeführt, die Lehrerpersönlichkeit dürfte als Erklärung für Unterschiede daher keinen Einfluss besitzen.

In drei Klassen ging der Unterricht von den Fragen der Kinder aus (diese Klassen werden im Folgenden als Experimentalgruppe (EG) bezeichnet), wobei das Rahmenthema 'So lebten Kinder früher' vorgegeben war. In einer Einführungsstunde erzählten die Kinder zunächst in einem Sitzkreis, was sie bereits über das Thema wissen; um die spätere Entscheidung etwas zu vereinfachen, wurden die Gebiete, die hier genannt wurden (z. B. ‚Schule früher‘, ‚Kleidung früher‘, ‚Kinder im Krieg‘ o. Ä.), gleich an der Tafel notiert. Anschließend wurde den Kindern erklärt, dass in der kommenden Woche zwei Doppelstunden Unterricht zum Thema ‚So lebten Kinder früher‘ stattfinden werden, dass sie sich individuell aussuchen könnten, was sie hier besonders interessiert, und dass jedes Kind dann seinen eigenen Fragen nachgehen könne. Die Kinder überlegten daher noch weitere Fragestellungen oder

inhaltliche Spezifizierungen (z. B. ‚Kinder in der Steinzeit‘, ‚Kinderarbeit‘ o. Ä.), die alle notiert wurden. Abschließend entschied sich jedes Kind für ein Thema. In der darauffolgenden Woche erhielt jedes Kind Materialien (in erster Linie historische Texte und Bilder) zu dem von ihm gewählten Thema. Gleichzeitig bekamen sie die Aufgabe, diese Materialien durchzuarbeiten, zu entscheiden, was sie für wichtig erachten und danach die wesentlichen Erkenntnisse auf einem Plakat festzuhalten, um sie später den anderen Schüler/innen der Klasse (die ja zum größten Teil andere Themen bearbeiteten) zu zeigen. Hatten mehrere Kinder das gleiche Thema gewählt, so wurden Kleingruppen gebildet. Zum Durcharbeiten der Materialien und zum Erstellen der Poster hatten die Kinder vier Unterrichtsstunden Zeit. Abschließend wurden die Poster im Sitzkreis vorgestellt. Jeder EG-Klasse wurde eine Parallelklasse der gleichen Schule als Kontrollgruppe (KG) zugeteilt. (In einer Schule wollten gerne alle drei Parallelklassen an der Untersuchung teilnehmen, so dass hier eine EG und zwei KG existieren). Der Unterricht in den KG unterschied sich lediglich in der ersten Stunde. Der Klasse wurde auch hier das Rahmenthema und das Erstellen der Poster erklärt; im Anschluss daran bekamen die Schüler/innen dann die Themen vorgegeben, welche die Kinder aus einer EG-Klasse ausgewählt hatten. Hatten z. B. vier Kinder aus einer EG-Klasse ‚Kinder in der Ritterzeit‘ gewählt, so erhielten vier Kinder aus der KG-Klasse das gleiche Thema. Die Verteilung erfolgte, soweit wie möglich auch nach Geschlechtern parallelisiert. Der Unterricht selbst lief völlig identisch zu dem der EG ab, die Schüler/innen bearbeiteten das gleiche Material und erstellten Poster zu ihrem Thema.

### *b) Instrumente*

*Skala Selbstbestimmungsempfinden:* Auf einer vierstufigen Ratingskala (zwischen 1 und 4) schätzten die Schüler/innen ihre Selbstbestimmung ein. Der Fragebogen enthält 11 Items wie z. B. „Beim Unterricht zu ‚So lebten Kinder früher‘ wählte ich oft selbst etwas aus“ oder „Beim Unterricht zu ‚So lebten Kinder früher‘ bestimmen die Schüler mit“. Die Skala hat sehr gute Kennwerte ( $\alpha = .91$ ). In einer Vorerhebung wurde der gleiche Fragebogen durchgeführt, hier jedoch allgemeiner auf den Unterricht der Klassenlehrerin/des Klassenlehrers bezogen ( $\alpha = .82$ ).

*Skala Interesse:* Das Interesse der Kinder am Thema wurde ebenfalls anhand einer 4-stufigen Ratingskala erhoben (10 Items;  $\alpha = .81$ ). Fragen waren z. B. „Beim Unterricht zu ‚So lebten Kinder früher‘ lernte ich Sachen, mit denen ich mich auch zu Hause beschäftigen will“ oder „Ich würde in meiner Freizeit gerne ein Buch darüber lesen, wie die Kinder früher gelebt haben“.

*Mitarbeit im Unterricht:* In Anlehnung an die Arbeiten von Wagner und Schöll (1992) sowie Helmke und Renkl (1992) wurde ein Beobachtungsinventar mit drei Kategorien entwickelt: „On-task aktiv“, „On-task inaktiv“ sowie „Off-task“. „On-



task inaktiv” ist dabei besonders für Gruppenarbeiten gedacht, wenn z. B. bei der Arbeit an einem Poster ein Kind ein anderes beim Schreiben oder Malen beobachtet. Die Werte beziehen sich stets auf die ganze Klasse: Alle 15 Sekunden wurde reihum ein anderes Kind beobachtet. Dies geschah während der zwei Doppelstunden jeweils 45 Minuten lang, so dass für jede Klasse insgesamt 360 Werte vorliegen – je nach Klassengröße waren das etwa sieben Werte pro Kind. Da bei einführenden Sitzkreisen u. Ä. keine Unterschiede zwischen den Experimental- und Kontrollgruppen erwartet werden konnten, wurden lediglich die direkten Arbeitszeiten der Kinder in der aktiven Beschäftigung mit ihrem Thema beobachtet. Die Beobachter/innen wurden anhand von Videoaufnahmen trainiert; die Beobachterübereinstimmung – ebenfalls über Videosequenzen erhoben – ist mit einem Kappa von .80 gut vertretbar.

### 3. Ergebnisse

#### a) *Mitarbeit im Unterricht:*

Das Ergebnis ist eindeutig (vgl. Abb. 1). In den Klassen, in denen der Unterricht von den Fragen der Kinder ausging, arbeiteten die Kinder im Durchschnitt aktiver mit (74,2% im Vgl. zu 66,6%) und zeigten weniger Off-task-Verhalten (11,0% zu 17,4%). Diese time on task Werte sind insgesamt sehr hoch, da bei nicht-lehrer-gesteuerter Arbeit üblicherweise ein Wert von 70% der Zeit erwartet werden kann (vgl. Karweit 1989). Keine bemerkenswerten Differenzen gab es bzgl. inaktiven Verhaltens. Berechnet man die Unterschiede zwischen EG und KG mit Chi-Quadrat-Verfahren (vgl. Clauß & Ebner 1985; sehr anschaulich auch bei Wagner & Schöll 1992, S. 31ff), so ergeben sich zwischen EG und KG signifikante Unterschiede ( $\chi^2 = 22,90$ ;  $p < .001$ ).

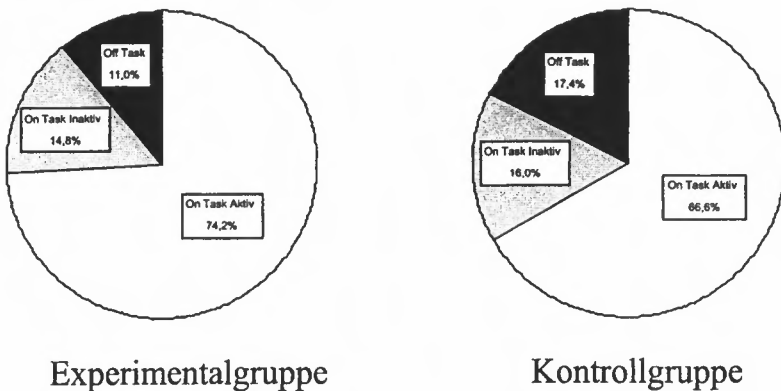


Abb. 1: Mitarbeit im Unterricht (Vergleich EG - KG);

### *b) Wahrgenommene Selbstbestimmung*

Die Einteilung der Klassen in EG bzw. KG war u. a. auch organisatorischen Bedingungen unterworfen. Rückblickend gesehen ergab sich daraus das Problem, dass sich die Klassen bereits in der Vorerhebung bzgl. des Selbstbestimmungsempfindens unterschieden. Allerdings ist m. E. nicht logisch eindeutig zu prognostizieren, welche Auswirkungen dies hat. Die These, dass Selbstbestimmungsempfinden ein relativ festes Persönlichkeitsmerkmal ist und dass daher Kinder mit hohem Selbstbestimmungsempfinden in der Vorerhebung und auch in der Nacherhebung höhere Werte haben, ist vertretbar. Genauso plausibel ist es jedoch, dass die Kinder in der Nacherhebung schwerpunktmäßig Unterschiede zum gewohnten Unterrichtsablauf werteten. Dann sind höhere Werte zu erwarten, wenn das Selbstbestimmungsempfinden in der Vorerhebung gering war. Aus diesem Grund ist es nicht sinnvoll, die Daten der Vorerhebung als Kovariate in die Berechnung einzubinden.

Um dennoch einen Vergleich zu haben, der nicht durch die Vorerfahrungen der Schüler/innen verfälscht ist, wurden nur Kinder in die folgenden Berechnungen einbezogen, wenn sich in der Parallelgruppe (EG bzw. KG) ein Pendant mit dem gleichen Selbstbestimmungswert in der Vorerhebung fand. Gab es in einer Gruppe mehr Kinder mit einem bestimmten Wert als in der anderen, so wurde zufällig ausgewählt. Insgesamt blieben dadurch noch 96 Kinder für die weiteren Berechnungen (48 EG; 48 KG). Alle weiteren Berechnungen beziehen sich nun auf diese Stichprobe.

Abb. 2 zeigt, dass in dieser Stichprobe Unterschiede bzgl. des Selbstbestimmungsempfindens bestehen. Die Schüler/innen aus der Experimentalgruppe empfanden sich während des Unterrichts zu 'So lebten Kinder früher' als selbstbestimmter als die Kinder der KG. Die Differenzen sind signifikant ( $t = 2,25$ ;  $p < .05$ ).

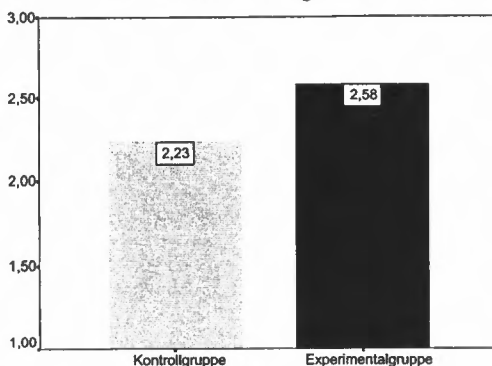


Abb. 2: Selbstbestimmungsempfinden Nacherhebung (Vergleich EG - KG; gematchtes Sample;  $n = 96$ );

### c) Interesse

Auch das Interesse an der unterrichteten Thematik unterscheidet sich signifikant zwischen Experimental- und Kontrollgruppe ( $t = 2.28; p < .05$ ) (vgl. Abb. 3). Die Schüler/innen, die im Unterricht ihre eigenen Fragen bzw. ihre selbstgewählten Themen bearbeiteten, äußerten danach mehr Interesse an der gesamten Thematik als die Kinder, die vorgelegte Themen bearbeiteten.

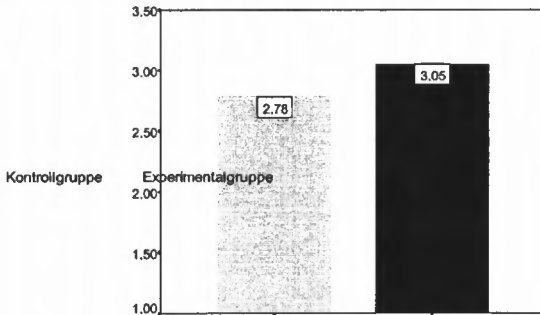


Abb. 3: Interesse an 'So lebten Kinder früher' Nacherhebung (Vergleich EG - KG; gematchtes Sample;  $n = 96$ );

Interessant ist dabei auch, dass hier nicht nur Unterschiede zwischen Kontroll- und Experimentalgruppe existieren. Es gibt auch deutliche Zusammenhänge zwischen dem Selbstbestimmungsempfinden der Schüler/innen und dem Interesse. Die Korrelation (Spearman's Koeffizient) liegt bei  $.221$  ( $p < .05$ ). Allerdings ist festzustellen, dass dieser Zusammenhang v. a. durch die Kontrollgruppe entsteht (vgl. Abb. 4). Besonders wenig Interesse entwickeln die Schüler/innen, die in der KG unterrichtet wurden und die sich zusätzlich als wenig selbstbestimmt empfinden. In der EG sind die Unterschiede zwischen den Kindern mit hohem und geringem Selbstbestimmungsempfinden nicht so markant.

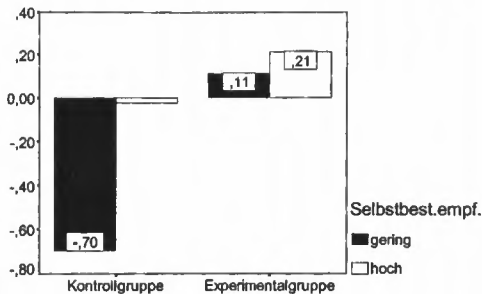


Abb. 4: Interesse an der Thematik: Extremgruppenvergleich (Terzile) Kinder mit hohem und geringem Selbstbestimmungsempfinden in EG und KG (Werte z-standardisiert)

#### 4. Diskussion und Ausblick

Bei der Einschätzung der Ergebnisse sind einige Einschränkungen zu betonen. Zum einen ist aufgrund der recht kleinen Stichprobe (v. a. hinsichtlich der Anzahl der Klassen) selbstverständlich nur wenig Repräsentativität gegeben. Hinzu kommt, dass dadurch manche mehrfaktorielle oder hierarchische Berechnungen, die aufgrund der Fragestellung erforderlich wären (z. B. im Vergleich der Bedeutung von objektiver zu subjektiv empfundener Selbstbestimmung), nicht sinnvoll durchgeführt werden können. Zum zweiten fehlen Vorerhebungsdaten bzgl. des Interesses der Schüler/innen vor dem Unterricht. Aufgrund des Aufbaus der Studie (es wäre m. E. recht problematisch, die Kinder zunächst nach ihrem Interesse zu fragen und dann in den Kontrollgruppen andere Themen vorzusetzen) habe ich dies nicht erhoben. Da die Ergebnisse zur Interessenentwicklung aber mit denen einer früheren Studie (vgl. Hartinger 1997) gut kompatibel sind, sind sie jedoch m. E. dennoch präsentierbar und bemerkenswert.

Zusammenfassend deuten die Daten auf zwei Ergebnisse hin: Zum einen scheint die hier durchgeführte inhaltlich-thematische Öffnung tatsächlich dazu beizutragen, dass die Schüler/innen höheres Interesse entwickeln, aktiver im Unterricht mitarbeiten und sich auch als selbstbestimmter empfinden. Dies gilt zumindest für das vorgestellte Thema. Inwieweit sich dieses Ergebnis verallgemeinern lässt, müsste durch weitere Untersuchungen geklärt werden.

Zum zweiten wird deutlich, dass nicht nur die ‚objektiv durchgeführte‘ Variation des Unterrichts (hier mehr oder weniger Selbstbestimmung durch die thematische Öffnung des Unterrichts), sondern auch das ‚subjektive‘ Selbstbestimmungsempfinden Einfluss auf die Interessenentwicklung hat. Aufgrund der vorliegenden Daten sind jedoch kausale Zusammenhänge nicht eindeutig identifizierbar; nicht zuletzt aufgrund der relativ kleinen Stichprobe ist auch vor Verallgemeinerungen zu warnen. Subgruppenunterschiede können hier nicht ausgeschlossen werden – sie deuten sich ja im Vergleich EG – KG bereits an (vgl. Abb. 4).

Die Frage, warum sich Schüler/innen im Rahmen desselben Unterrichts als mehr oder weniger selbstbestimmt empfinden und wie sich dies auswirkt, bedarf dringend weiterer Erhebungen. Hier wird es auch wichtig werden, Persönlichkeitsmerkmale von Schüler/innen zu identifizieren, von denen eine Auswirkung auf das Selbstbestimmungsempfinden zu erwarten ist, wie z. B. Ambiguitätstoleranz, Attributionsstil oder Ängstlichkeit. Untersuchungen, in denen dies genauer erhoben wird, sind in Vorbereitung.

## Literatur

- Burns, R. B. (1984). How Time is Used in Elementary Schools: The Activity Structure of Classrooms. In L. Anderson (Ed. ), *Time and School Learning* (pp. 91-127). London & Canberra: Croom Helm.
- Clauß, G. & Ebner H. (1985). *Statistik für Soziologen, Pädagogen, Psychologen und Mediziner. Bd. 1: Grundlagen*. 5. Aufl. Thun & Frankfurt a. M. : Harri Deutsch.
- Deci, E. L., Eghari, H., Patrick, B. C. & Leone, D. R. (1994). Facilitating Internalization: The Self-Determination Theory Perspective. *Journal of Personality*, 62, No. 1, 119-142.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39, H. 2, 223-238.
- Deci, E. L., Schwartz, A. J., Sheinman, L. & Ryan, R. M. (1981). An Instrument to Assess Adults' Orientation toward Control versus Autonomy with Children: Reflections on Intrinsic Motivation and Perceived Competence. *Journal of Educational Psychology*, 73, 642-650.
- Einsiedler, W. (1997). Empirische Grundschulforschung im deutschsprachigen Raum - Trends und Defizite. *Unterrichtswissenschaft*, 25, 291-315.
- Giaconia, R. M. & Hedges, L. V. (1982). Identifying Features of Effective Open Education. *Review of Educational Research*, 52, 579-602.
- Götz, M. (1991). Weiß die Ameise, daß sie Ameise heißt? Überlegungen zur pädagogischen und didaktischen Bedeutung von Kinderfragen im Sachunterricht. *Grundschule*, 23, H. 11, 51-54.
- Green, L. & Foster, D. (1986). Classroom Intrinsic Motivation: Effects of Scholastic Level, Teacher Orientation, and Gender. *Journal of Educational Research*, 80, No. 1, 34-39.
- Grolnick, W. S. & Ryan, R. M. (1987). Autonomy in Children's Learning. An Experimental and Individual Difference Investigation. *Journal of Personality and Social Psychology*, 52, No. 5, 890-898.
- Hanke, P. (2001). Forschungen zur inneren Reform der Grundschule am Beispiel der Öffnung des Unterrichts. In H. G. Roßbach, K. Czerwenka & K. Nölle (Hrsg. ), *Jahrbuch Grundschulforschung. Bd. 4: „Forschung zu Lehr- und Lernkonzepten für die Grundschule*. (S. 46-62) Bad Heilbrunn: Klinkhardt.
- Hartinger, A. (1997). *Interessenförderung. Eine Studie zum Sachunterricht*. Bad Heilbrunn: Klinkhardt.
- Hartinger, A. (2001). Selbstbestimmung im Unterricht. Die Sicht der Schüler/innen. In H. G. Roßbach, K. Czerwenka & K. Nölle (Hrsg. ), *Jahrbuch Grundschulforschung. Bd. 4: „Forschung zu Lehr- und Lernkonzepten für die Grundschule*. (S. 93-101) Bad Heilbrunn: Klinkhardt.
- Heckt, D. H. & Jürgens, E. (1999). Zur Qualität offener Lernsituationen. *Grundschule*, 31, H. 7-8, 62-64.
- Helmke, A. & Renkl, A. (1992). Das Münchener Aufmerksamkeitsinventar (MAI): Ein Instrument zur systematischen Verhaltensbeobachtung der Schülaufmerksamkeit im Unterricht. *Diagnostica*, 38, 130-141.
- Jürgens, E. (1999). Freiarbeit im Spiegel der Praxis. *Grundschule*, 31, H. 7-8, 46-48.
- Karweit, N. (1989). Time and Learning: A Review. In R. E. Slavin (Ed. ), *School and Classroom Organization* (pp. 69-95). Hillsdale: Lawrence Erlbaum.
- Kasper, H. (1989). Offener Unterricht in der Diskussion. In H. Kasper u. a. (Hrsg. ), *Laßt die Kinder lernen* (S. 12-21). Braunschweig: Westermann.
- Kasser, T. & Ryan, R. M. (1993). A Dark Side of the American Dream: Correlates of Financial Success as a Central Life Aspiration. *Journal of Personality and Social Psychology*, 65, No. 2, 410-422.
- Kasser, T. & Ryan, R. M. (1996). Further Examining the American Dream: Differential Correlates of Intrinsic and Extrinsic Goals. *Personality and Social Psychology Bulletin*, 22, No. 3, 280-287.

- Krapp, A. (1999). Intrinsische Lernmotivation und Interesse. Forschungsansätze und konzeptionelle Überlegungen. *Zeitschrift für Pädagogik*, 45, No. 3, 387-406.
- Lipowsky, F. (1999). *Offene Lernsituationen im Grundschulunterricht. Eine empirische Studie zur Lernzeitnutzung von Grundschulern mit unterschiedlicher Konzentrationsfähigkeit*. Frankfurt a. M.: Lang.
- Monty, R. A. & Perlmutter L. C. (1987). Choice, Control, and Motivation in the Young and Aged. In M. L. Maehr & D. A. Kleiber (Eds.), *Advances in Motivation and Achievement: Enhancing Motivation*, Vol. 5 (pp. 99-122). Greenwich & London: JAI Press Inc.
- Peschel, F. (1996). *Offener Unterricht am Ende - oder erst am Anfang?* Bericht Nr. 2 Projekt OASE. Primarstufe, FB 2 der Universität-Gesamthochschule, Siegen.
- Reeve, J., Bolt, E. & Cai, Y. (1999). Autonomy-Supportive Teachers: How They Teach and Motivate Students. *Journal of Educational Psychology*, 91, 537-548.
- Ryan, R. M. & Connell, J. P. (1989). *Self-Regulation Questionnaire*. o. O.
- Ryan, R. M., Deci, E. L. & Grolnick, W. S. (1995). Autonomy, Relatedness and the self: Their Relation to Development and Psychopathology. In D. Cicchetti & D. Cohen (Eds.), *Developmental Psychopathology. Vol. 1: Theory and Methods* (pp. 618-655). New York: John Wiley & Sons.
- Saldern, M. von (1998). Die Aufgabenfülle der Grundschule und ihrer Pädagogik. Eine Sammelbesprechung zur Grundschulpädagogik und Grundschulforschung. *Zeitschrift für Pädagogik*, 44, 907-924.
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 120-148.
- Schiefele, U. & Schreyer, I. (1994). Intrinsische Lernmotivation und Lernen. Ein Überblick zu Ergebnissen der Forschung. *Zeitschrift für Pädagogische Psychologie*, 8, 1-13.
- Schöll, G. (1996). Offene Lernsituationen in der Grundschule. In H. Ulonska, S. Kraschinski & T. Bartmann (Hrsg.), *Lernforschung in der Grundschule* (S. 197-218). Bad Heilbrunn: Klinkhardt.
- Swann, W. B. Jr. & Pittman, T. S. (1977). Initiating Play Activity of Children: The Moderating Influence of Verbal Cues on Intrinsic Motivation. *Child Development*, 48, 1128-1132.
- Wagner, G. & Schöll, G. (1992). *Selbständiges Lernen in Phasen freier Aktivitäten - Entwicklung eines Beobachtungsinventars und Durchführung einer empirischen Untersuchung in einer 4. Grundschulklasse*. Nürnberg: Berichte und Arbeiten aus dem Institut für Grundschulforschung. Bd. 70.
- Zuckerman, M., Porac, J., Lothian, D., Smith, R. & Deci, E. L. (1978). *On the Importance of Self-Determination for Intrinsically-Motivated Behavior*. *Personality and Social Psychology Bulletin*, 4, No. 3, 443-446.

## **Soziale Beziehungen zwischen Grundschulkindern – empirische Befunde zu einem wichtigen Thema des Sachunterrichts**

### **1. Einleitung**

Für die Rekonstruktion des komplexen Themas „Soziale Beziehungen von Kindern“ erscheint eine mehrperspektivische Analyse besonders angemessen. Dabei lässt sich nach den Analyseebenen *Kind – Interaktion – Gruppe – Kontext* unterscheiden, von denen aus der Gegenstand expliziert, in der Forschung analysiert und in curriculare Überlegungen eingebracht werden kann:

Auf der *Ebene des Kindes* bezieht sich die Analyse u. a. auf die soziale Entwicklung, den Erwerb sozialer Kompetenzen und Orientierungen, den Status und die Selbstbehauptung in der Gruppe sowie situative Handlungsweisen in Gruppensituationen und einzelnen Beziehungen.

Die *Ebene der Interaktion* beinhaltet u. a. Prozesse des dyadischen sozialen Wechselspiels in unterschiedlichen sozialen Situationen (Konflikt, Kooperation, Spiel) und die Betrachtung einzelner Beziehungen, die nach der Qualität des Austauschs, nach dem Ausmaß an Intimität und nach der Begegnungshäufigkeit variieren.

Auf der *Ebene der Gruppe*, als Ort des Vergleichens mit anderen, der Selbsteinschätzung und der Vergabe von Einfluss stehen u. a. soziale Strukturen (soziales Beziehungsgefüge, Teilgruppen, Rangordnung), das Sozialklima sowie Normen des Zusammenlebens im Mittelpunkt der Betrachtung.

Auf der *Ebene des Kontextes* richtet sich der Wahrnehmungsschwerpunkt auf außer- und innerschulische Faktoren, die u. a. von situativen Einflussgrößen in einzelnen Schulklassen über formelle und informelle Strukturen der Schule bis zu Entwicklungsdeterminanten in Familien und Peergroups reichen.

Neben der generellen Bedeutung sozialen Lernens als zentralen Aspekt grundlegender Bildung lassen Veränderungen in den Entwicklungsbedingungen der Kinder und daraus resultierende Defizite im Umgang mit Gleichaltrigen (vgl. Bohardt 1999, Herzberg 1992) die Schule zum „Dreh- und Angelpunkt sozialer Interaktionen und der Aushandlung sozialer Regeln“ (Preuss-Lausitz 1999, S.166) werden und verlangen nach geeigneten pädagogischen Konzepten.

## 2. Soziale Beziehungen im Sachunterricht

Im Sachunterricht kommt sozialem Lernen die Funktion zu, fachliche Inhalte mit sozialen Elementen zu verbinden (Katzenberger 1995; Schreier 1994) und soziale Beziehungen im Sinne einer „sozialen Wende“ (Meier 1993) auf den verschiedenen Ebenen zu einem bedeutsamen Gegenstand des Unterrichts zu machen (Kaiser 2000). Dabei geht es vor allem um Lernsituationen, die über das Entwickeln theoretischer Lösungswege und das Verbalisieren sozialer Einsichten hinaus Möglichkeiten der Selbstregulation innerhalb der Gleichaltrigengruppe, der individuellen Erprobung interaktionalen Handelns und der gemeinsamen Auseinandersetzung über Sachverhalte und Probleme fördern sollen (vgl. Petillon 1993b; Speck-Hamdan 1997), ohne dass die Eigendynamik der Gruppe in bedeutsamen Bereichen eingeschränkt wird. Fähigkeiten, auf die der Unterricht angewiesen ist, entstehen im Sozialleben der Schülerinnen und Schüler (Krappmann 2001).

Soziales Lernen als Gegenstand des Unterrichts und als erzieherischer Schwerpunkt nimmt im Sachunterricht unter verschiedenen Aspekten eine Sonderstellung ein, die zu einer konsequenten Realisierung von Erfahrungsbezug, Nachhaltigkeit, Ko-Konstruktion des Wissens, Reflexion des Lernweges und Fachbezug herausfordert:

- Die Gruppe der Gleichaltrigen ist gleichzeitig alltäglicher sozialer Erfahrungsraum mit hoher Eigendynamik, Lernort und Ausgangspunkt für *erfahrungshaltige* unterrichtliche Prozesse. Daraus resultierendes *erfahrungs*offenes verantwortliches Handeln, das aus der Logik des sozialen Geschehens unmittelbar erwächst sowie *erfahrungs*orientiertes Evaluieren von eigenen Gestaltungsversuchen verweisen auf ein breites Spektrum an Lernmöglichkeiten: Interessen aufeinander abstimmen, Argumente entwickeln, Beziehungen „lesen“, Einfühlungsvermögen entfalten u. a. Bei der Bearbeitung dieses Themenbereichs wird der metakommunikative Austausch in besonderer Weise gefördert. In der Gruppe erarbeitete „theoretische Konzepte“ müssen sich dabei einer praktischen Bewährung und entsprechenden Verifizierungsprozessen stellen.
- Das Prinzip der Nachhaltigkeit, als Hinweis auf die langfristige Speicherung und Abrufbarkeit des Wissens, Identifikation mit dem Gelernten und hohe Alltagsrelevanz, nimmt im Bereich des sozialen Lernens eine Schlüsselfunktion ein. Auseinandersetzungen mit Widersprüchen und Bemühungen andere zu überzeugen, tragen wesentlich dazu bei, in ein Thema tief einzudringen und vermitteln auf nachhaltige Weise Einsichten in Sachverhalte, Denkmuster und Gesetzmäßigkeiten (vgl. Krappmann 2001).

In Untersuchungen zur Perspektive der Kinder wird deutlich, dass Beziehungen zu Gleichaltrigen in der kindlichen Lebenswelt von zentraler Bedeutung sind (vgl. Petillon 1993a; Stöckli 1997). Im Verlauf der Grundschulzeit finden



sich deutliche Hinweise auf die abnehmende Orientierung an Maßstäben der Eltern und die zunehmende Bedeutung von Peers, die den Heranwachsenden notwendige Räume für ihre Entwicklung öffnen (Stöckli 1997). Darüber hinaus verweisen andere Befunde auf die nachhaltige Wirkung sozialer Erfahrungen mit Peers für das spätere Zurechtkommen im Erwachsenenalter (vgl. Parker & Asher 1987; Hymel, Rubin, Rowden & LeMare 1990).

- Aus der Sicht des „sozialen Konstruktivismus“ (vgl. Gerstenmaier & Mandl 1995) können sich Kinder auf der Ebene prinzipiell gleichberechtigter Partner in Prozessen der *Ko-Konstruktion* (vgl. Youniss 1980) mit unterschiedlichen Perspektiven auseinandersetzen. Sie können individuelle Konstruktionen in Beziehung setzen, Lösungswege aushandeln, im Prozess der gemeinsamen Vergewisserung „angereicherte“ Zugänge zu Sachverhalten gewinnen, gemeinsam Vorhaben entwerfen, zu einem „kritischen Bewusstwerden von Denkmöglichkeiten“ (Köhnlein 2000, S.139) gelangen. Viele „fruchtbare Momente“ für die kognitive Entwicklung entstehen aus Widersprüchen zwischen der eigenen und einer fremden Perspektive, wobei im Sinne einer „Dezentrierung“ (Piaget 1972) aus dem Bewusstsein der Nichtübereinstimmung und der Suche nach Gemeinsamkeit interne Prozesse der Umstrukturierung in Gang gesetzt werden können. Es wird angenommen, dass Konflikte eher die kognitive Problemlösung fördern, während Prozesse des sozialen Ausgleichs die Entwicklung sozialer Kompetenz unterstützen (Damon 1982). Die unerlässliche Grundlage für einen offenen Austausch, für eine produktive Auseinandersetzung und damit für eine hohe Lernqualität sind Strukturen, Normen und das Sozialklima in der Gruppe, die von hoher Gesprächs- und Kooperationsbereitschaft, „Toleranz gegenüber anderen Lernwegen mit Fehlern und Umwegen“ (Möller 1999) geprägt sind.
- Offene Methoden des Sachunterrichts sind auf soziale Kompetenzen der Kinder angewiesen und gleichzeitig ein Ort sozialen Lernen und metakognitiven Austauschs, bei dem der Weg und die Methode zur Erlangung von Wissen in den Mittelpunkt der Betrachtung rückt. Das Kind in seiner Bewährung in einzelnen sozialen Interaktionen und in der Auseinandersetzung mit der Gruppe wird zum Thema, wenn sich aus der „pädagogischen Logik“ der Situation ein Klärungs- und Handlungsbedarf entwickelt. Der Lernweg zu der angestrebten Lösung insbesondere im Hinblick auf Kommunikations- und Kooperationsprozesse wird dabei zu einem zentralen Gegenstand gemeinsamer Bemühungen. Wie Untersuchungen zeigen, kann Unterricht an Lernqualität gewinnen, wenn die Kinder mehr Gelegenheiten zur Kooperation, zur gegenseitigen Unterstützung und zum Austausch ihrer Gedanken haben (vgl. Topping & Ehly 1998).
- Die Herstellung eines Fachbezugs und damit auch die Berücksichtigung des

Anliegens der Sachangemessenheit (vgl. Schreier 2000) fordert für das Thema „soziale Beziehungen“ den Einbezug von Bezugswissenschaften (Psychologie, Soziologie, Sozialpsychologie und Erziehungswissenschaft). Dabei sind entsprechende theoretische und empirische Befunde für Lehrerinnen und Lehrer in dreifacher Hinsicht hilfreich: (1) als fachlicher Hintergrund für die unterrichtliche Gestaltung, (2) als Wissensbasis, die über die differenzierte, sensible Wahrnehmung sozialer Phänomene in der eigenen Schulklasse hinaus die Autorität eines Moderators, Anregers und Helfers unterstützen kann und (3) im Hinblick auf vorliegende Untersuchungsverfahren als Möglichkeit, Ausgangsbedingungen und Entwicklungen in der Schülergruppe sorgfältiger zu diagnostizieren.

Im Weiteren wird auf empirische Untersuchungen zu dem Themenbereich „soziale Beziehungen von Kindern“ eingegangen. Aus der Vielzahl vorliegender Befunde wählen wir im Rahmen dieser Ausführungen eher diejenigen aus, die mit entsprechenden Themenstellungen in der Praxis des Sachunterrichts korrespondieren. Wir gehen dabei von den Lehrplänen für die einzelnen Bundesländer aus. In der folgenden Synopse vorliegender Themenvorschläge im Sachunterricht wird nach den o.g. Ebenen „Kind“, „Interaktion“ und „Gruppe“ differenziert. Als einfaches „Suchraster“ soll diese Übersicht (vgl. Tab. 1 S. 189/90) lediglich in Stichworten auf die einzelnen Schwerpunktsetzungen verweisen.

### 3. Empirische Untersuchungen zu sozialen Beziehungen von Grundschulkindern – ein Überblick

Die wissenschaftlichen Zugänge zu Kindern und die Rekonstruktion von kindlichen Perspektiven werden in bedeutsamer Weise von Forschungsmethoden, die das Erheben von Daten und deren systematische Analyse betreffen, beeinflusst (Heinzel 2000). In empirischen Untersuchungen zu sozialen Beziehungen von Kindern lassen sich drei grundlegende Zugangsweisen unterscheiden:

(1) *Beobachtung*: Standardisierte Prozessbeobachtungen, sog. Interaktionsanalysen (vgl. Bales 1950), vermitteln einen relativ zuverlässigen Einblick in das interaktionale Geschehen von Teilgruppen (vgl. z. B. Anderka, Hackenberg & Helbig 1997), sind aber im Hinblick auf Reichweite und Differenziertheit der Wahrnehmung deutlich begrenzt. Demgegenüber stehen teilnehmende Beobachtungen, die sich bemühen, in der „natürlichen Umwelt“ der Kinder (im Klassenzimmer, auf dem Pausenhof, bei Ausflügen) Beobachtungen in Form von „dichten Beschreibungen“ zu protokollieren (vgl. Beck & Scholz 2000), um danach durch ein systematisches Vergleichen vieler beobachteter Ereignisse Zusammenhänge und Muster zu entdecken (Oswald 2000). Häufig werden diese Beobachtungen durch Videoaufnahmen ergänzt, oder das Videografieren bildet die Grundlage für einzelne

Bundesland	Kind	Interaktion	Gruppe
Bayern 1981	soziale Kompetenz (einfühlen, zuhören, Gefühle ausdrücken, Toleranz), Selbstdarstellung	gegenseitige Hilfe, Ermutigung, Trost, sich einigen, Kompromiss, Mädchen - Jungen	Erleben von Gemeinschaft, Lernen in der Gruppe, Regeln für das Zusammenleben, Entscheidungen beraten
Baden-Württemberg 1994	soziale Kompetenzen, Selbstvertrauen, Selbstbehauptung, Umgang mit Problemen	Konflikte und Lösungen, Hilfsbereitschaft und Rücksichtnahme, Freundschaft, Mädchen - Jungen	zusammen leben und lernen, Umgangsformen und Regeln erarbeiten, Rollenklischees
Berlin 1990	soziale Kompetenzen, soziale Entwicklung	Konflikte und Lösungen, Mädchen - Jungen	Gemeinschaft erleben, Kooperation lernen, Regeln
Brandenburg 1991	sich selbst und andere darstellen, Gefühle anderer entschlüsseln, Problemsituation	miteinander gleichberechtigt spielen, Konflikte und Lösungen, Freundschaft, Mädchen - Jungen	Gleichberechtigung, Umgangsformen spielen und üben
Bremen 1984	Gefühle ausdrücken, Wertschätzung für andere entwickeln, Toleranz, den eigenen Standpunkt vertreten	Konflikte und Lösungen, Rollenaufgaben: Jungen und Mädchen, partnerschaftliche Arbeitsaufgaben	Zusammenarbeit in Gruppen, soziale Ordnungen, Integration von Neulingen, Gleichberechtigung
Hamburg 1979	soziale Kompetenzen (soziale Sensibilität, Kontaktbereitschaft)	Konflikte und Lösungen, gegenseitige Hilfe, Mädchen - Jungen	Gemeinschaft erleben, Kooperation lernen, Regeln
Hessen 1995	soziale Kompetenzen, soziale Erfahrungen bearbeiten	Kontakte aufnehmen Absprachen treffen, Streit ohne Gewalt, Mädchen - Jungen	demokratische Spielregeln einüben, Arbeitsteilung lernen
Mecklenburg-V. 1996	Schlüsselqualifikationen (Team- Argumentations- und Kritikfähigkeit)	Freunde finden, gewaltfreie Konfliktlösungen, Mädchen - Jungen	Normen und Regeln für das Leben und Lernen in der Klassengemeinschaft, Klassenrat, Mitbestimmung

Abb. 1: Thematisierung sozialer Beziehungen in Sachunterrichtslehrplänen

Bundesland	Kind	Interaktion	Gruppe
Niedersachsen 1982	soziale Grunderfahrungen und Grundkenntnisse, Selbstbehauptung, Umgang mit Gefühlen	Ursachen von Streit und Lösungsmöglichkeiten, Freundschaft, Partnerschaft: Jungen und Mädchen	Zusammenarbeit, gemeinsam erarbeitete Ordnungen, Mitbestimmung
Nordrhein- Westfalen 1985	Einfühlungsvermögen, mit Problemen sorgfältig und kritisch auseinander- setzen	Konflikte und Lösungen, Vertrauen, Rolle: Mädchen - Junge	Dienste für die Klassen- gemeinschaft, Regeln des Miteinander entwickeln
Rheinland- Pfalz 1984	sich selbst wahrnehmen, einen Platz in der Gruppe finden, Verantwortung über- nehmen, Selbstbehauptung	Konflikte und Lösungsmög- lichkeiten, Freundschaft schließen, Mädchen - Jungen sind verschieden	gemeinschaftsbezogene Lebens- und Arbeits- formen, Regeln des Zusammen- lebens
Saarland 1992	soziale Kompetenzen, soziale Sensibilität	Kontakte aufnehmen, Kon- flikte und Lösungsmöglich- keiten	Gemeinschaft erleben, Kooperation lernen, Regeln
Sachsen 1992	soziale Kompetenzen, soziale Sensibilität, Umgang mit Problemen	Gegenseitigkeit, Freundschaft, Mädchen - Jungen	Gemeinschaft erleben, Kooperation lernen, Regeln
Sachsen- Anhalt 1993	soziale Kompetenzen, Selbstbehauptung, soziale Sensibilität	Problemlösungen suchen, Freundschaft, Mädchen - Jungen	Gemeinschaft erleben, Kooperation lernen, Regeln
Schleswig- Holstein 1993	soziale Schlüssel- qualifikationen, Identitätsentwicklung	Konflikte und Lösungen, Mädchen - Jungen, Zärtlichkeit	Leben in der Gemein- schaft, Integration in die Gruppe, demokratische Handlungsweisen ein- üben
Thüringen 1993	soziale Kompetenzen, Selbstbewußtsein, Selbstbehauptung	gegenseitige Hilfe, auf neue Kinder eingehen, Mädchen - Jungen	Gemeinschaft erleben, Kooperation lernen, Regeln

Analysen (vgl. Huhn, Dittrich, Dörfler & Schneider 2000).

2) *Befragung*: Das Spektrum schriftlicher Befragungsformen reicht von standardisierten, auf Validität und Reliabilität überprüften *Fragebögen* zur Perspektive einzelner Kinder mit Skalen zu Sozialerfahrungen mit Mitschülern, zu Sozialinteresse, Kontaktbereitschaft (Petillon 1984), sozialer Angst, Schüchternheit (Asendorpf 1993), Einsamkeit (von Salisch 1991), sozialer Unzufriedenheit und interpersonalem Vertrauen (Stöckli 1997) bis zu eher informellen schriftlichen Befragungen ohne Überprüfung der Gütekriterien (Preuss-Lausitz 1999). In halbstrukturierten *Interviews* werden Kinder zu sozialen Beziehungen (Freundschaft, Einbindung in die Gruppe, Streit, Prozesse des Aushandelns) befragt (vgl. Krappmann & Oswald 1995); bei offenen Interviews bietet sich Kindern die Möglichkeit, beispielsweise auf der Grundlage von Bildkarten zu einzelnen Emotionen bedeutsame soziale Episoden zu schildern (Petillon 1983). Entsprechende qualitative Daten können in qualitativen oder quantitativen, auf Kategoriensysteme rekurrierende Analysen, ausgewertet werden.

*Soziometrie*: Mit Hilfe unterschiedlicher Erhebungsverfahren können bedeutsame Beziehungsphänomene einer Gruppe ermittelt werden (Dollase 1998). Zur Erfassung relationaler Daten findet sich eine Vielzahl von Erhebungsmöglichkeiten, die von Wahlverfahren und Aktionstests über das „Klassenspiel“ (Masten, Morison & Pellegrini 1985; Stöckli 1997), bei dem Mitschülern vorgegebene Rollenbeschreibungen zugeordnet werden, bis zu unterschiedlichen Formen der Beobachtung reichen. In den am häufigsten eingesetzten Wahlverfahren können Beziehungsmuster zwischen je zwei Kindern, der Status jedes Gruppenmitglieds, das Beziehungsnetz (Soziogramm) und Strukturkennzeichen (z. B. Kohäsion, Rangefälle) analysiert werden. Dabei erscheinen für die Befragung von Kindern Bildwahlverfahren besonders geeignet: Kinder wählen für bestimmte soziale Kategorien, die auf Bildkarten dargestellt sind (Freundschaft, Spielen, Bezugsperson, Konflikt, Ausschluss, Führung) aus den Fotos aller Mitschüler die jeweils zutreffenden aus (Petillon 1993a). Auch wenn Validitäts- und Reliabilitätsstudien darauf hindeuten, dass insbesondere Statuswerte die soziale Stellung eines Kindes in der Gruppe verlässlich abbilden können, sagen die Ergebnisse wenig über die Qualität der einzelnen Beziehung aus. Hier erweisen sich Verfahren, die den Kindern selbst die Möglichkeit einräumen, Beziehungen differenzierter darzustellen (vgl. Krappmann & Oswald 1995), als erfolgversprechender. Dies gilt vor allem auch für Untersuchungen, die auf verschiedene empirische Zugangsweisen zurückgreifen (vgl. Stöckli 1997; Krappmann 2001) und dabei im Sinne einer mehrperspektivischen Analyse ein differenzierteres, umfassenderes und valideres Bild von sozialen Phänomenen erhalten.

Im Folgenden werden Untersuchungsbefunde dargestellt, ohne dass jeweils auf einzelne Erhebungsverfahren eingegangen wird.

### *Perspektive: Kind*

Das Wissen über die *soziale Entwicklung* bietet eine wichtige Grundlage für die Gestaltung eines sozialerzieherischen Unterrichts, der an den sozialen Möglichkeiten der einzelnen Kinder anknüpft und sie verständnisvoll unterstützend beim Erwerb sozialer Kompetenzen begleitet. Im Verlauf der Kindheit, in der sich sowohl Qualität als auch Häufigkeit der sozialen Beziehungen zwischen Kindern verändern, vollzieht sich eine soziale Entwicklung, die bei entsprechenden förderlichen ökologischen Bedingungen die Möglichkeiten des Kindes erweitert, zielgerichtet und kompetent an seinem sozialen Umfeld zu partizipieren. Im Verlauf der Grundschulzeit kommt es dabei zu einer Ausweitung sozialer Beziehungen auf mehrere Interaktionspartner (Lewis & Rosenblum 1975; Krappmann & Oswald 1995); der normative Grundrahmen für soziale Interaktionen wird differenzierter (Halisch 1988), die Dauer von Interaktionen nimmt zu (Emmerich 1971). Mit zunehmenden kommunikativen Fähigkeiten vergrößern sich auch die Sprachanteile in Interaktionsprozessen; verbale Strategien und die Fähigkeit, die Sprache des anderen zu entschlüsseln, verfeinern sich (La Greca & Stark, 1986; Boueke & Schülein 1995). Darüber hinaus wächst die Fähigkeit zu angemessenen Prognosen über das Verhalten anderer und über die Folgen eigenen Handelns (Scarlett 1971). Strategien im Umgang mit Konflikten zwischen Gleichaltrigen werden differenzierter (Wiley 1983; Petillon 1993a); dies gilt auch für Taktiken, sich Zugang zu Gruppenaktivitäten zu verschaffen (Lubin & Forbes 1981).

Im Bereich des *sozialen Verstehens* kommt der *Perspektivenübernahme*, als Grundlage für erfolgreiches Handeln in sozialen Beziehungen, besondere Bedeutung zu. Der Begriff der Perspektivenübernahme bezeichnet dabei nicht nur die Art und Weise, in der psychologisches Wissen vom Standpunkt einer anderen Person gesehen wird, sondern geht über diesen Prozess der Rollenübernahme hinaus und umfasst „das sich entwickelnde Verständnis dafür, wie verschiedene Blickwinkel zueinander in Beziehung stehen und miteinander koordiniert werden“ (Selman 1984, S. 30). Das Konzept geht über die komplexe Koordination dezentrierter kognitiver Operation hinaus und schließt das sich entwickelnde Verständnis der intrinsischen Eigenschaften und Fähigkeiten anderer Personen mit ein.

Eine Synopse vorliegender Befunde macht deutlich, dass die Grundschulzeit als „sensible Phase“ für die soziale Entwicklung, insbesondere des sozialen Verstehens und sozialer Sensibilität, betrachtet werden kann. Angesichts hoher „Plastizität“ sozialer Verhaltens- und Beziehungsmuster erscheint die Förderung elementarer *sozialer Kompetenzen* im Bereich der Grundschule im Sinne einer ersten Anbahnung und kontinuierlichen Förderung besonders erfolgversprechend. Die Fähigkeit und Bereitschaft, sich verständlich zu machen und andere zu verstehen, vielfältige Beziehungen aufzunehmen und aufrecht zu erhalten, zu kooperieren, Konflikte argumentativ und in fairem Austausch auszutragen, mit eigenen und Gefühlen

anderer sensibel umzugehen, in der Gruppe selbstbewusst und selbstbestimmt zu handeln sowie sich Regeln des Zusammenlebens zu erarbeiten und anzuerkennen, kennzeichnen basale soziale Zielbereiche und verweisen gleichzeitig auf die Schülergruppe als zentralen Ort sozialen Lernens.

Soziale Beziehungen, in die das Kind eingebunden ist, sind eine wichtige Basis seiner Selbstkonstruktion. Insbesondere die Gruppe als sozialer Erfahrungsraum bietet Möglichkeiten, Selbstbehauptung und Selbstsicherheit zu erwerben: sich einen Status erarbeiten; Spielregeln mitbestimmen und befolgen; Beziehungen eingehen; sich vor einem „Publikum“ bewähren; Konflikte bestehen; Lösungen aushandeln; gemeinsam arbeiten und spielen; mit Andersartigkeit umgehen.

Der *soziale Status des Kindes in der Gleichaltrigengruppe* als Indikator für soziale Attraktivität oder Beliebtheit in der Klasse bestimmt in bedeutsamer Weise Qualität und Quantität sozialer Erfahrungen. Auf der Grundlage soziometrischer Erhebungen lassen sich zum sozialen Status des Kindes zusammenfassend folgende Befunde nennen:

- Statushohe Kinder erwerben sich höhere soziale Kompetenz, die wesentlich zur Fundierung ihrer sozialen Stellung beiträgt (Gottman 1975).
- Es bestehen signifikante Zusammenhänge zwischen der Popularität unter Gleichaltrigen und dem Selbstwertgefühl (Horowitz 1962; Stöckli 1997).
- Abgelehnte und isolierte Kinder zeigen häufiger allgemeine Angstsymptome (Nickel & Schlüter 1970); sie entwickeln mehr Prüfungsängste und neigen zu regressivem Verhalten (Petillon 1982).
- Zahlreiche Untersuchungen belegen einen engen Zusammenhang zwischen Schulleistung und sozialer Anerkennung (Stöckli 1997). Darüber hinaus besteht ein enger Zusammenhang zwischen dem soziometrischen Status und der Beliebtheit beim Lehrer (Petillon 1993).
- Häufig führen soziale Probleme mit anderen Kindern dazu, dass Schulleistungen weit unter dem Niveau liegen, das auf Grund der Begabung erreicht werden könnte (Kemmler 1967).

Unser besonderes Augenmerk sollte den Kindern in Außenseiterpositionen gelten. Diese Positionen besitzen eine ausgesprochen hohe Stabilität. Bereits in den ersten beiden Schuljahren scheinen sich die Kinder einer Schulklasse auf einzelne Mitschüler festzulegen, die immer wieder Opfer von aggressiven Handlungen, von Ausschluss und Hänseleien sind (Stöckli 1997; Petillon 1993a). Ein Schwerpunkt der Forschung bildet in diesem Zusammenhang das Phänomen „Bullying“, als Hinweis auf das regelmäßige Schikanieren (Mobbing) einzelner Kinder, die sich auf Grund klarer Unterlegenheit kaum zur Wehr setzen können. (Busch & Todt 1998). Die Täter sind oft die gleichen Kinder (Bullies), die nach Überlegenheit streben, zu aggressivem Verhalten eine positive Einstellung haben und mit den Opfern wenig Mitgefühl entwickeln. Die regressiven Reaktionen der Opfer stabili-

sieren Außenseiterrollen und reduzieren die Gelegenheit am Sozialleben erfolgreich zu partizipieren und damit soziale Probleme aus eigener Kraft zu bewältigen. Die vom negativen Verhalten der Mitschüler betroffenen Kinder fühlen sich in vielen Fällen einsam, leiden unter Schüchternheit (Asendorpf 1993; Petillon 1984) und sind mit ihrer Situation in hohem Maße unzufrieden (Stöckli 1997). In den einzelnen Stichproben der Grundschüler finden sich etwa 15 Prozent der untersuchten Kinder, die in verschiedenen Formen mit sozialen Schwierigkeiten und Integrationsproblemen zu kämpfen haben. Pädagogische Maßnahmen als „eine Korrektur der defizitären Integration in die Welt der Gleichaltrigen“ (Stöckli 1997, S. 120) erscheinen dringend notwendig. Dabei sind einer gezielten Intervention sowohl eine differenzierte Analyse der Verhaltensweisen von zurückgewiesenen Kindern als auch die Betrachtung der Zuschreibungsgewohnheiten der übrigen Kinder zu Grunde zu legen. Hier fallen auch dem Sachunterricht bedeutsame sozialpädagogische Aufgaben zu.

### *Perspektive: Interaktion*

Interaktionen können als wechselseitige Einwirkung, als komplexes Gewebe von Handlungen und Gegenhandlungen beschrieben werden. In ereignisbezogenen Verständigungs- und Abstimmungsprozessen versuchen die Interaktionspartner, Situationsdeutungen zu koordinieren, eigene Handlungsziele durchzusetzen oder auf den anderen abzustimmen. Intensität, Dauer sowie das Ausmaß an gegenseitiger Vertrautheit und Wertschätzung der jeweiligen Beziehung sind für die Qualität dieser Austauschprozesse von zentraler Bedeutung.

Zunächst ist davon auszugehen, dass die Schulklasse nicht das einzige Reservoir von Beziehungsangeboten ist. Etwa zwei Drittel der Kinder finden in Grundschulklassen einen Spielpartner, mit dem sie auch außerhalb der Schule häufig zusammen sind. Die anderen suchen Kontakte zu Geschwistern oder Kindern in der häuslichen Umgebung; viele spielen auch allein (Stöckli 1997). Der Rückzug in die Familie kann allerdings fehlende Peerkontakte in der Schule nur sehr begrenzt kompensieren (vgl. East & Rook 1992). Die sozial vertrauensvollsten, weniger gehemmten Kinder nutzen primär das Beziehungsangebot der Klasse. Demgegenüber sind es eher die schüchternen, wenig vertrauensvollen Kinder, die auf Alternativen außerhalb der Klasse ausweichen. Die Meidung einer so vielfältigen sozialen Ressource, wie sie die Schulklasse darstellt, weist in vielen Fällen auf tiefgreifende Beziehungsprobleme hin, während die Integration in der Klasse zum Gradmesser individueller Kontaktfähigkeit wird (Stöckli 1997, S. 139).

Von besonderer Bedeutung für die soziale Entwicklung sind intensive *Freundschaftsbeziehungen*. Das Finden eines Freundes und die Ausgestaltung der Freundschaft werden bereits in den ersten Wochen der Schulzeit zum zentralen Thema der



Kinder (Petillon 1993a). Dabei sind die entstandenen Freundschaften langlebiger als gemeinhin geglaubt wird (Oswald 1994). Besonders Mädchen bleiben in den meisten Fällen über die gesamte Grundschulzeit mit ihrer Freundin zusammen, während Jungen etwas häufiger wechseln (Petillon 1993a).

In Freundschaftsbeziehungen bietet sich für die Kinder die Möglichkeit, unter der Bedingung gegenseitiger Wertschätzung kooperative Formen des sozialen Umgangs zu erfahren. Die Analysen von Krappmann & Oswald (1988) zeigen, dass mehr als unter Nichtfreunden, bei denen das überdauernde Interesse aneinander fehlt, in gemeinsamen Anstrengungen Gesichtspunkte der anderen Seite berücksichtigt, fehlgegangene Lösungen rekonstruiert und zusätzliche Argumente aufgenommen werden; es wird intensiver „ko-konstruiert“ (Youniss 1980). Im Vertrauen auf die Loyalität des Freundes können neue Kenntnisse gemeinsam ausprobiert und erlernte Fähigkeiten kooperativ geübt werden (Biskup 1994). Die Aufmerksamkeit füreinander ist größer; in Konflikten finden Freunde öfter zu beiderseits akzeptierten „beziehungsschonenden“ Lösungen (Newcomb & Bagwell 1995; Petillon 1993a). Diskussionen und Verhandlungen münden eher in gemeinsame Vorstellungen und Ideen ein; man nimmt sich mehr Zeit füreinander und kommuniziert effektiver, da man sich besser kennt und vieles „eingespielt“ ist (Nelson & Aboud 1985).

Im Verlauf der Kindheit verändern sich die Vorstellungen davon, was Freundschaft bedeutet, wofür man einen Freund braucht und was man selbst in die Beziehung einbringen muss. Entsprechende Konzepte entwickeln sich von selbstbezogenen Orientierungen bis zu Formen eines partnerzentrierten gerechten Austauschs. Selman (1984) nennt fünf Stufen des Freundschaftsverständnisses, die sich im Verlauf der Kindheit von kurzlebigen Spielkameradschaften über einseitig konzipierte Freundschaften, krisenanfällige „Schönwetter“-Freundschaften, vertrauensvoll geteilte Beziehungen bis zur „Autonomie in gegenseitiger Abhängigkeit“ entwickeln können. In einer Längsschnittuntersuchung in den ersten beiden Grundschuljahren (Petillon 1993a) wurde sichtbar, dass mit zunehmender Zeit die Ansprüche der Grundschulkinder an das Sozialverhalten des Freundes (Vertrauen, Hilfe, Empathie) steigen. Mädchen suchen häufiger soziale Nähe und eine „Gesprächspartnerschaft“, während Jungen das gemeinsame Spielen mit Wettbewerbscharakter und die Unterstützung bei Konflikten mit anderen mehr in den Vordergrund ihres Freundschaftskonzeptes rücken. Dabei wird das gemeinsame Bemühen um einen zufriedenstellenden Ausgleich der Interessen als existentiell für den Fortbestand einer Freundschaftsbeziehung gesehen. Weiterhin finden sich Hinweise, dass in vielen Fällen Kinder solche Mitschüler zum Freund auswählen, die ein ähnliches Freundschaftsverständnis besitzen. Freundschaften mit einer solchen Übereinstimmung bleiben über einen langen Zeitpunkt stabil und begünstigen gegenseitiges Verständnis und einen beziehungsfördernden Umgang mit Konflikten.

Allerdings beschränkt sich kooperatives, auf Gleichheit beruhendes Handeln in vielen Fällen auf exklusive Freundschaftsbeziehungen und wird nur begrenzt auf Beziehungen zu anderen Kindern übertragen. Wenn Beziehungen ungeklärt sind, kooperieren Kinder häufig recht unproduktiv. Neber (1998, S. 270) nennt Effekte, die den Erfolg kooperativen Lernens im Verlauf von Gruppenarbeit gefährden: beim „free rider Effekt“ überlassen schwächere Schüler die Lernarbeit den leistungsfähigeren Gruppenmitgliedern; der „sucker Effekt“ bedeutet, dass sich Leistungsstärkere ausgebeutet fühlen und ihre Anstrengungen reduzieren; der „statusabhängige Effekt“ reduziert die lernbezogenen Partizipationschancen der statusniedrigen Mitglieder, während der „ganging up Effekt“ bewirken kann, dass sich die Gruppe auf eine Lösung mit der geringsten Anstrengung einpendelt. Nach Beobachtungen von Krappmann & Oswald (1995) verläuft Zusammenarbeit im Unterricht nur in einem Drittel aller Fälle ohne größere Spannungen und überwiegend aufgabenorientiert.

Obwohl Prinzipien der Gleichberechtigung, Nichtübertreibung und Nichtbevormundung in Beziehungen zwischen Kindern eine große Rolle spielen und auch zu heftigen Auseinandersetzungen führen, können sie sich bei der Suche nach gerechten Lösungen und dem Ausgleich von Interessen in vielen Situationen nicht durchsetzen. In weniger als der Hälfte der Fälle wird in der Grundschule in Konflikten zwischen Kindern eine von allen Beteiligten akzeptierte Lösung erreicht. Viele Kinder versuchen sehr heftig und rücksichtslos ihr Anliegen durchzusetzen, Erfahrungen des Sich-durchsetzen-Könnens oder Unterliegens dominieren. Auch von der Erfahrung misslungener Interaktionen können Entwicklungsimpulse ausgehen, indem an gebrochenen Vereinbarungen in reflektierender Auseinandersetzung weitergearbeitet und gerechte Lösungen im inneren Dialog „weiterkonstruiert“ werden (Krappmann 1999).

Bei der Beobachtung von Situationen, in denen zwischen mindestens zwei Kindern ein Dissens auftrat, der aktiv von den Beteiligten ausgetragen wurde, fanden sich in fast der Hälfte aller Fälle zwangsausübende, respektlose Strategien; eher selten (unter 10%) ließ sich eine argumentative Vermittlung beobachten. Mädchen erreichen häufiger als Jungen eine Lösung, die beidseitig akzeptiert wird. So bleibt festzustellen, dass de facto viele der beobachteten Aushandlungen nicht dazu herausfordern, „die Perspektive des anderen zu übernehmen, ein kontroverses Thema unter verschiedenen Rücksichten zu betrachten, Begründungen darzulegen, Gefühle und Selbstbilder zu schützen und gemeinsam Verantwortung zu entwickeln“ (Krappmann & Oswald 1995, S. 103). Darüber hinaus lassen eine Vielzahl „unbeglichener Rechnungen“ unbearbeitete Konflikte weiterschwelen und bei passender Gelegenheit wirksam werden, so dass für Außenstehende nur der Anlass, aber nicht die eigentliche Ursache erkannt werden kann. Vielfach berichten Kinder von körperlichen Aggressionen, die deutlich über die spielerischen Rauf- und

Tobespiele, „rough and tumble play“ (Oswald 1997) hinausgehen und die Grenze zur Gewalt überschreiten. Viele Jungen passen sich einem einmal erreichten Gewaltniveau in der Gruppe an (Petillon 1993a, S. 75). Allerdings ist unser Wissen über Gewaltanwendung und deren Funktion eher oberflächlich. Oswald & Krappmann (2000) entdeckten bei ihren Beobachtungen, dass der Einsatz der Gewalt ein Handlungsschritt ist, der einer Logik des Interaktionsgefüges und der Interaktionsbedingungen unter Kindern entspricht: „Gewalt eröffnet Erfahrungen mit sich und anderen, die Kinder intensiv suchen; Gewalt ist der Versuch, Grenzen zu errichten, die Teures und Liebes gegen respektlosen Zugriff abschirmen; Gewalt verlockt, weil sie das Verhältnis zu anderen eindeutig klärt; Gewalt bietet sich als Abkürzung zu einem Ziel an, wenn andere Vorgehensweisen versagen; Gewalt ist ein Mittel, um das Wichtigste, das eigene Selbst zu behaupten“ (S.14).

Das Thema „Konflikt und Lösungen“, wie es in vielen Lehrplänen genannt wird, stellt für die Gestaltung des Sachunterrichts eine große Herausforderung dar. Auf der Grundlage differenzierter Analysen sozialer Prozesse in der jeweiligen Schulklasse, die Konflikte unter Kindern ernst nehmen, sollte es darum gehen, Kindern alternative Möglichkeiten anzubieten, ihre Interessen zu verfolgen und mit wachsender sozialer Kompetenz neue Wege des sozialen Austauschs zu finden. Beispiele für konkrete Möglichkeiten dazu finden sich bei Balsen, Schrewe & Shoof (1997). Gegenseitiges *Vertrauen* trägt in besonderer Weise zur Festigung einer Beziehung bei. Auf die Frage, in welchem Ausmaß man den Gleichaltrigen vertrauen kann, reagieren viele Kinder eher skeptisch oder unsicher. Die Zahl der Gruppenmitglieder, denen man ein Geheimnis anvertrauen kann, ist deutlich begrenzt. Zweifel anderer Mitschüler an der eigenen Vertrauenswürdigkeit führen zu starken persönlichen und sozialen Beeinträchtigungen bei den betroffenen Kindern. Die einem Kind zugeschriebene Vertrauenswürdigkeit korreliert hoch mit seinem Beliebtheitsstatus (Wentzel 1991); abgelehnte Kinder gelten unter Gleichaltrigen als ausgesprochen wenig vertrauenswürdig. Misstrauen gegenüber einzelnen Gruppenmitgliedern begünstigt den sozialen Rückzug, das Festhalten an negativen Interpretationsmustern und verhindert alternative Erfahrungen der wenig beliebten Kinder. Mädchen sind im Hinblick auf das eigene Geschlecht häufig vertrauensvoller (Stöckli 1997).

Fast in allen Lehrplänen werden die *Beziehungen zwischen Jungen und Mädchen* als Thema des Sachunterrichts vorgegeben. In den vorliegenden Untersuchungen findet sich übereinstimmend der Hinweis, dass sich Mädchen und Jungen in der Grundschulzeit voneinander abgrenzen, eine Entwicklung, die sich bereits zum Ende der Kindergartenzeit abzeichnet (Schmidt-Denter 1988). Geschlechterpräferenzen sind bereits zu Beginn der Schulzeit angelegt. Längsschnittuntersuchungen verweisen auf eine zunehmende Separierung im Verlauf der ersten vier Grundschuljahre (Macoby & Jacklin 1987; Stöckli 1997). Auf der Grundlage soziometri-

scher Wahlen bestätigt sich das Bild der getrennten Welten in der Kindheit. Speziell im geselligen Bereich ist der Wunsch nach Kontakten mit Angehörigen des anderen Geschlechts nur bei einer Minderheit anzutreffen. Engere Freundschaften zwischen Mädchen und Jungen sind eine Seltenheit. Allerdings gelingt es einer Minderheit, die soziale Distanz im Umfeld der Trennung zu überwinden. Dabei sind es nicht die Problemkinder, die in ihrer Gruppe keinen Anschluss finden, sondern es handelt sich eher um eine „soziale Elite“, um besonders beliebte, selbstständige und sozial kompetente Mädchen und Jungen, die die Grenzlinie zwischen den Geschlechtern überwinden.

Auf die Nachfrage nach Begründungen für die Bevorzugung des eigenen Geschlechts scheinen Mädchen eher rollenspezifisch festgelegt („Jungen spielen mit Jungen, Mädchen mit Mädchen“) oder nennen negative Aspekte der Jungen, während die Jungen mit den positiven Aspekte ihrer eigenen Gruppe argumentieren. Zunehmend werden auch Aversionen gegen das andere Geschlecht genannt (Petillon 1993a).

Mädchen und Jungen entwickeln unterschiedliche „Spielkulturen“ (Petillon 1993a; Thorne 1986): *Jungen* bevorzugen Spiele mit Wettbewerbscharakter, bei denen es häufig auch um die Klärung von Rangordnungen geht und Durchsetzungsfähigkeit erprobt wird; sie spielen mehr im Freien und entwickeln dort ein höheres Mobilitätsniveau. Mädchen dagegen versuchen im Spiel ihr Bedürfnis nach Intimität und Nähe zu realisieren (Swann 1992); sie bevorzugen kommunikativ geprägte Rollen- und Regelspiele und bringen dort ihre intensivere empathische Ausrichtung ein (von Salisch 1991).

Bei „Grenzüberschreitungen“ beteiligen sich Mädchen an Jungenspielen in ernsthafter Weise, während die Jungen die Aktivitäten von Mädchen häufiger durch „Invasionen“ rücksichtslos stören (Stöckli 1997). Oft sind es spielerisch getarnte Handlungen, die auf den Wunsch nach Annäherung verweisen. In den meisten Fällen verhindert die aus den „eigenen Reihen“ hervorgehende soziale Kontrolle eine zwanglose Annäherung: „Neckereien und Abwertungen liegen oft näher als Kooperation und Wertschätzung“ (Stöckli 1997, 215). Vieles deutet darauf hin, dass Separierung als generelles Prinzip der sozialen Orientierung eine Situation der Ungleichheit schafft, die stereotype Voreinstellungen verstärkt und den Rückzug in die geschlechtshomogene Gruppe begünstigt. Stöckli (1997) plädiert auf der Grundlage dieser Befunde für pädagogische Interventionen, die eine kooperative Annäherung zwischen Jungen und Mädchen fördert, um durch eine qualitative Ausdehnung sozialer Möglichkeiten dem Zusammenleben der Geschlechter neue Perspektiven zu eröffnen. Wie Interventionsstudien zeigen, lassen sich Erfolge in der Förderung eines offeneren, kooperativeren Verhältnisses zwischen Mädchen und Jungen erzielen (Bigler & Liben 1992; Katz und Walsh 1991). Auch hier ist der Sachunterricht im Hinblick auf eine „Normalisierung“ andersgeschlechtlicher

Beziehungen in besonderer Weise gefordert und bietet wichtige Ansatzpunkte, wenn sich Mädchen und Jungen gemeinsam auf eine intensive Auseinandersetzung mit einer „Sache“ einlassen.

### *Perspektive Gruppe*

Der Eintritt bestimmter Sozialereignisse, deren Verlauf und Prozesse der Bewältigung werden von *Normen, Strukturen und dem Sozialklima* der Gruppe mitbestimmt, die als Rahmen für das Sozialleben der Kinder hemmende und förderliche Möglichkeiten bereitstellen.

Als Ergebnis von Interaktionen entstehen unter den Kindern gruppenspezifisch geprägte *Normen*, als gemeinsame „Spielregeln“ des Zusammenlebens, durch die verlässliche Regelmäßigkeiten in Interaktionsabläufen entstehen können. Es entwickelt sich bei Grundschulkindern ein Wissen über soziale Regulierungen, das von dem frühesten Niveau rein situational verstandenen Regulierungen über die Kalkulation von Sanktionen bis zur Einsicht in den Eigenwert von Regeln für die Aufrechterhaltung sozialer Ordnungen reicht (Damon 1982). Es wird angenommen, dass sich unter Kindern eine Vorstellung von Gleichheit ausbildet, die im Umgang miteinander angestrebt wird (Krappmann & Oswald 1995, S. 88). Dabei steht das Prinzip der Gegenseitigkeit, d.h. ein gerechter Austausch von Geben und Nehmen, als Ideal ausgewogener Beziehungen im Hintergrund des interaktionalen Geschehens. Auch wenn die Inhalte dieses Austausches von begünstigten Handlungen (zusammen spielen) über gegenseitiges Vertrauen bis zum Anvertrauen und Bewahren von geheimen Gedanken variieren, bleibt die Funktion von Gegenseitigkeit gleich: „wenn eine Partei ihren Beitrag zurückhält, wird die Interaktion unbalanciert und damit die Beziehung gefährdet“ (Schmidt-Denter 1988, S. 287). Allerdings zeigen differenzierte Beobachtungen, dass in vielen Fällen außerhalb von Freundschaftsbeziehungen „Gleichheit eher ein Prinzip ist, das eingeklagt wird, weil sehr oft dagegen verstoßen wird, nicht aber ein Anzeichen alltäglicher Realität“ (Krappmann & Oswald 1995, S. 22) darstellt.

*Soziale Strukturen* kennzeichnen mehr oder weniger stabile Muster unterschiedlichen Einflusses, gegenseitiger Anerkennung und Sympathie, Kommunikationsnetze, Erwartungsmuster und Hierarchien. Vieles deutet darauf hin, dass *Rangordnungen* in sich neu bildenden Kindergruppen schnell geklärt werden. Wie eine Längsschnittuntersuchung zeigt, gelang es 80% der Kinder, die zu Beginn des ersten Schuljahres in einflussreichen Positionen waren, auf der Grundlage ihrer „Monopolstellung“ über zwei Jahre ihren Einfluss zu behaupten (Petillon 1993a). Nach einem ersten Stadium der Kontaktsuche und der Klärung von Rangordnungen scheint sich in vielen Schulklassen sehr bald nach Schulbeginn ein Stadium „erster Institutionalisierung“ zu entwickeln, das den sozialen Handlungsspielraum für die einzelnen Kinder im Hinblick auf Einfluss und Anerkennung vorstrukturiert.

Innerhalb der Klasse lassen sich unterschiedliche *Gruppierungsformen* feststellen. Auf der Basis bestimmter Gemeinsamkeiten finden sich Kinder zu *Teilgruppen (Cliques)* zusammen, in denen sich neben einer deutlichen Einflussdifferenzierung, gemeinsame Normen und spezifische Formen und Themen des Kommunizierens entwickeln, die den Umgang der Mitglieder untereinander erleichtern, dagegen den Zugang für Außenstehende erschweren. Intensivere Kontakte fördern einen offenen Austausch von „persönlichen Themen“ und schaffen damit ein gewisses Maß an Sicherheit, das man nur innerhalb der eigenen „Clique“ findet. Neben dieser Gruppierung beobachteten Krappmann & Oswald 1995 *Beziehungsflechte*, die einen weniger ausgeprägten Gruppencharakter besitzen. Sie haben ebenfalls einen beschreibbaren Kreis von Zugehörigen und gemeinsame Themen, aber keine klare Außengrenze und kein dichtes Netz enger, dauerhafter Beziehungen eines jeden Mitglieds mit jedem anderen. Die Restgruppe der „nichtangebundenen“, ausgegrenzten Kinder findet sich in *Interaktionsfeldern* zusammen, die außer einer erhöhten Interaktionsdichte keine Gruppenmerkmale besitzen. Es bildet sich dort kein Gruppenzusammenhang heraus, „weil die Kinder sich nur als situative Partner annehmen und ansonsten vergeblich nach anderen Beziehungen Ausschau halten“ (S. 54).

Das *Sozialklima* charakterisiert die „emotionale Grundtönung“ (Eder 1998, S. 424) der sozialen Atmosphäre in einer Schulklasse als Hinweis auf die emotionale Befindlichkeit der Mitglieder einer Gruppe. Bei entsprechenden Befragungen werden die Aussagen einzelner Kinder durch eine statistische Mittelwertbildung zu einem „aggregierten Klimawert“ zusammengefasst. In einem der elaboriertesten Fragebogenverfahren werden folgende Indikatoren genannt: Ausmaß der Cliquenbildung, Hilfsbereitschaft, Aggressionen, Diskriminierung, Zufriedenheit und Konkurrenzverhalten (vgl. von Saldern & Littig 1987).

Für den Primarbereich liegen bisher keine empirischen Befunde zum Sozialklima in Schulklassen vor. Allerdings lässt sich aus den o.g. Befunden zu anderen Gruppenphänomenen annehmen, dass das Sozialklima in einzelnen Klassen sehr unterschiedlich ausgeprägt ist. Darüber hinaus ist davon auszugehen, dass das Erleben von Zugehörigkeit, Zufriedenheit und sozialer Sicherheit in bedeutsamer Weise vom Status des Kindes in der Gruppe mitbestimmt wird. Es spricht vieles dafür, an die Stelle eines unscharfen, als Gruppenindex eher ungeeigneten und in seinem komplexen Anspruch überforderten Klimabegriffes beobachtbare Fakten zum Beziehungsgeschehen zu setzen und auf dieser Grundlage die emotionalen Befindlichkeiten einzelner Kinder zu erschließen. Ähnlich unklar und für subjektive Deutungen offen wie das Sozialklima ist der Begriff der Gemeinschaft, der als Hinweis auf atmosphärische Aspekte der Gruppe in fast allen Lehrplänen als Gegenstand oder Zielgröße des Sachunterrichts genannt wird. Für eine nüchterne Betrachtung sozialer Beziehungen und Überlegungen zur schulpraktischen Förde-

rung sozialen Lernens erscheint dieser Begriff wenig geeignet. Auch hier wären konkrete Hinweise auf Ansatzpunkte für spezifische Fördermöglichkeiten hilfreicher als die Nennung eines unbestimmten Zielhorizonts.

Viele der vorgelegten Befunde verdeutlichen, wie notwendig und sinnvoll es wäre, durch gezielte schulische Maßnahmen zur Verbesserung der Gleichaltrigenbeziehungen beizutragen. Allerdings wissen wir aus den wenigen Studien zur sozialerzieherischen Praxis, dass geplantes Lernen im Sozialbereich nur eine sehr untergeordnete Rolle spielt, und Lehrer wenig über das Sozialleben der Schüler wissen (Petillon 1983). Auf der anderen Seite bestätigen einzelne Studien, dass mit Interventionsprogrammen nachweislich positive Effekte erzielt werden können (Stöckli 1997; Beelmann, Pfingsten & Loesel 1994; Petermann & Petermann 1997; Petillon & Flor 1997). Darüber hinaus ist festzustellen, dass Lehrer in der Grundschule großen Einfluss auf den sozialen Status einzelner Kinder und die sozialen Beziehungen nehmen können. In einer Beobachtungsstudie ließ sich beispielsweise feststellen, dass in Klassen in denen sich der Lehrer sehr wertschätzend gegenüber den Kindern verhielt, weniger körperliche Auseinandersetzungen auftraten, weniger Kinder in Außenseiterpositionen waren, die Wahlbereitschaft im soziometrischen Test und die Kohäsion (Dichte des Netzes gegenseitiger, erwidelter Wahlen) höher waren (Petillon 1982).

Wie vorliegende Studien zu sozialen Beziehungen von Kindern zeigen, kommt dem Sachunterricht ein Themenbereich zu, der eine besondere Herausforderung an das didaktische und pädagogische Gestaltungsvermögen stellt. Dabei sind „das von einem Lehrplan gesteuerte und von Lehrenden überwachte Lernen und die eigenständigen, von Kindern selbst betriebenen sozialen Prozesse“ in eine „ausgehandelte Ordnung“ zu integrieren (Krappmann 2001, S. 114). Der Sachunterricht gewinnt an Lernqualität und Nachhaltigkeit, wenn er unmittelbar aus der Logik des sozialen Geschehens erwächst und dabei den Kindern Freiräume öffnet für Kooperation, gegenseitige Unterstützung und intensiven Austausch.

## Literatur

- Anderka, A., Hackenberg, R. & Helbig, P. (1997). *Soziale Integration und soziales Lernen im kooperativen Unterricht*. Nürnberg: Berichte und Arbeiten aus dem Institut für Grundschulforschung der Universität Erlangen-Nürnberg, Nr. 84.
- Asendorpf, J. B. (1993). Personality effects on children's speech in everyday life. *Journal of Personality and Social Psychology*, 64, 1072-1083.
- Bales, R. K. (1950). *Interactions process analysis*. Reading: Addison-Wesley.
- Balsen, H., Schrewe, H. & Schoof, N. (1997). *Schulprogramm: Gewaltprävention: Ergebnisse aktueller Modellversuche*. Neuwied: Luchterhand.
- Beck, G. & Scholz, G. (2000). Teilnehmende Beobachtung von Grundschulkindern. In F. Heinzel (Hrsg.), *Methoden der Kindheitsforschung* (S. 147-170). Weinheim: Juventa.

- Beelmann, A., Pfungsten, U. & Lösel, F. (1994). Effects of training social competence in children. *Journal of Clinical Child Psychology*, 23, 260-271.
- Bigler, R. S. & Liben, L. S. (1992). Cognitive mechanisms in children's gender stereotyping. *Child Development*, 63, 1351-1363.
- Biskup, C. (1994). Was Kinder in Zusammenarbeit mit anderen lernen. *Grundschule*, 26, H. 4, 13-15.
- Bohrhardt, R. (1999). *Ist wirklich die Familie schuld? Familialer Wandel und soziale Probleme im Lebenslauf*. Opladen: Leske + Budrich.
- Boueke, D. & Schülein, F. (1995). *Wie Kinder erzählen - Untersuchungen zur Erzähltheorie und zur Entwicklung narrativer Fähigkeiten*. München: Wilhelm Fink.
- Busch, L. & Todt, E. (1998). Gewalt in der Schule. In D. H. Rost (Hrsg.) *Handwörterbuch Pädagogische Psychologie* (S. 168-172). Weinheim & Basel: PVU.
- Damon, W. (1982). Zur Entwicklung der sozialen Kognition des Kindes. In: W. Edelstein & M. Keller (Hrsg.), *Perspektivität und Interpretation* (S.110-145). Frankfurt: Suhrkamp.
- Dollase, R. (1998). Soziometrie. In D.H. Rost (Hrsg.) *Handwörterbuch Pädagogische Psychologie* (S.488-492). Weinheim & Basel: PVU.
- East, P. L. & Rook, K. S. (1992). Compensatory patterns of support among children's peer relationship: A testing using friends, nonschool friends, and siblings. *Developmental Psychology*, 28, 163-172.
- Eder, F. (1998). Schul- und Klassenklima. In D. H. Rost (Hrsg.) *Handwörterbuch Pädagogische Psychologie* (S. 424-430). Weinheim: PVU.
- Emmerich, W. (1971). Differentiation and the development of social norms. *Journal of Personality and Social Psychology*, 18, 128-134.
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41, 867-888.
- Gottman, J. (1975). Social interaction, social competence and friendship of children. *Child Development*, 46, 709-718.
- Halisch, F. (1988). Empathie, Attribution und die Entwicklung des Hilfehandelns. In H.W. Bierhoff & L. Montada (Hrsg.). *Altruismus* (S. 79-103). Göttingen: Hogrefe.
- Heinzel, F. (Hrsg.) (2000). *Methoden der Kindheitsforschung*. Weinheim: Juventa.
- Herzberg, I. (1992). Kinderfreundschaften und Spielkontakte. In DJI (Hrsg.), *Was tun Kinder am Nachmittag? Ergebnisse einer empirischen Studie zur mittleren Kindheit* (S.75-126). Weinheim/ München: Juventa.
- Horrowitz, F.D. (1962). The relationship of anxiety, self-concept and sociometric status-among fourth, five and six grade children. *Journal of Abnormal Social Psychology*, 65, 212-214.
- Huhn, N., Dittlich, G., Dörfler, M. & Schneider, K. (2000). Videografieren als Beobachtungsmethode in der Sozialforschung am Beispiel eines Feldprojekts zum Konfliktverhalten von Kindern. In F. Heinzel (Hrsg.), *Methoden der Kindheitsforschung* (S.185-202). Weinheim: Juventa.
- Hymel, S., Rubin, K. H., Rowden, L. & LeMare, L. (1990). Children's peer relationships: Longitudinal prediction of internalizing and externalizing problems from middle to late childhood. *Child Development*, 61, 2004-2041.
- Kaiser, A. (2000). Sachunterricht der Vielfalt - implizite Strukturen der Integration. In G. Löffler, V. Möhle, D. von Reeken, & V. Schwier (Hrsg.), *Sachunterricht - Zwischen Fachbezug und Integration* (S.91-107). Bad Heilbrunn: Klinkhardt.
- Katz, P.A. & Walsh, P.V. (1991). Modification of children's gender stereotype behavior. *Child Development*, 62, 338-351.
- Katzenberger, L. (1995). Der soziokulturelle Lernbereich des Sachunterrichts der Grundschule. In: H. R. Becher & J. Bennack (Hrsg.), *Taschenbuch Grundschule* (S. 210-231). Baltmannsweiler: Schneider Verlag Hohengehren.



- Kemmler, L. (1967). *Erfolg und Versagen in der Grundschule*. Göttingen: Hogrefe.
- Köhnlein, W. (2000). Vielperspektivität, Fachbezug und Integration. In G. Löffler, V. Möhle, D. von Reeken, & V. Schwier (Hrsg.), *Sachunterricht - Zwischen Fachbezug und Integration* (S. 134-146). Bad Heilbrunn: Klinkhardt.
- Krappmann, L. (1999). Die Reproduktion des Systems gesellschaftlicher Ungleichheit in der Kinderwelt. In: M. Grundmann (Hrsg.), *Konstruktivistische Sozialisationsforschung* (S. 228-239). Frankfurt: Suhrkamp.
- Krappmann, L. (2001). Soziales Leben und Lernen im Klassenzimmer. In H. Merken & J. Zinnecker (Hrsg.), *Jahrbuch Jugendforschung* (S.99-116). Opladen: Leske + Budrich.
- Krappmann, L. & Oswald, H. (1988). Probleme des Helfens unter Kindern. In H. W. Bierhoff & L. Montada, (Hrsg.), *Altruismus* (S. 206-223). Göttingen: Hogrefe.
- Krappmann, L. & Oswald, H. (1995). *Alltag der Schulkinder*. Weinheim & Basel: Juventa.
- La Greca, A. M. & Stark, P. (1986): Naturalistic Observation of Children's Social Behavior. In M. Strain, M. Guralnick & H. M. Walker (Hrsg.), *Children's Social Behavior* (S. 181-213). London: Academic Press.
- Lewis, M. & Rosenblum, L. A. (1975). *Friendship and peer relation*. New York: Wiley.
- Lubin, D. & Forbes, D. (1981). Motivational and peer culture issues in reasoning-behavioral relations. Paper presented at the biennial meeting of the Society for Research in Child Development, Boston.
- Macoby, E. E. & Jacklin, C. N. (1987). Gender segregation in childhood. In H. W. Reese (Hrsg.), *Advances in child development and behavior*. New York: Academic Press.
- Masten, A. S., Morison, P. & Pellegrini, D. S. (1985). A revised class play method of peer assessment. *Developmental Psychology*, 21, 523-533.
- Meier, R. (1993). Dimensionen des Zusammenlebens. In R. Lauterbach, W. Köhnlein, H. Kiper & I.-A. Koch (Hrsg.), *Dimensionen des Zusammenlebens* (S. 19-44). Kiel: IPN.
- Möller, K. (1999). *Konstruktivistische Sichtweisen für das Lernen in der Grundschule*. Vortrag auf der 8. Jahrestagung Grundschulforschung an der Universität Lüneburg.
- Neber, H. (1998). Kooperatives Lernen. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 361-366). Weinheim: PVU.
- Nelson, J. & Aboud, F. E. (1985). The resolution of social conflicts between friends. *Child Development*, 56, 1009-1017.
- Newcomb, A. F. & Bagwell, C. L. (1995). Children's Friendship Relations: A meta-analytic review. *Psychological Bulletin*, 117, 306-347.
- Nickel, H. & Schlüter, H. (1975). Angstwerte bei Hauptschülern und ihr Zusammenhang mit Leistungs- sowie Verhaltensmerkmalen, Lehrerurteil und Unterrichtsstil. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 7, 125-136.
- Oswald, H. (1994). *Gruppen- und Freundschaftsbeziehungen. Grundschule*, 26, H. 4, 10-12.
- Oswald, H. (1997). Zur sozialisatorischen Bedeutung von Kampf- und Tobespiele (Rough and tumble play). In E. Renner (Hrsg.), *Spiele der Kinder – Interdisziplinäre Annäherung* (S. 154-167). Weinheim: DSV.
- Oswald, H. (2000). Geleitwort. In F. Heinzel, (Hrsg.), *Methoden der Kindheitsforschung* (S. 9-15). Weinheim: Juventa.
- Oswald, H. & Krappmann, L. (2000). Phänomenologische und funktionale Vielfalt von Gewalt unter Kindern. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 49, 3-15.
- Parker, J. G. & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk? *Psychological Bulletin*, 102, 357-389.
- Petermann, F. & Petermann, U. (1997). *Training mit aggressiven Kindern*. Weinheim: PVU.
- Petillon, H. (1982). *Soziale Beziehungen zwischen Lehrern, Schülern und Schülergruppen*. Wein-

- heim: Beltz.
- Petillon, H. (1984). *Sozialfragebogen für Schüler, SFS 4-6*. Weinheim: Beltz Test.
- Petillon, H. (1993a). *Das Sozialleben des Schulanfängers*. Weinheim: PVU.
- Petillon, H. (1993b). *Soziales Lernen in der Grundschule. Anspruch und Wirklichkeit*. Frankfurt: Diesterweg.
- Petillon, H. & Flor, D. (1997). Wissenschaftliche Evaluation des Modellversuchs „Lern- und Spielschule“ in Rheinland-Pfalz. In R. Fatke & R. Valtin (Hrsg.), *Sozialpädagogik in der Grundschule* (S. 157-174). Frankfurt: Arbeitskreis Grundschule.
- Piaget, J. (1972). *Sprechen und Denken des Kindes*. Düsseldorf: Schwann.
- Preuss-Lausitz, U. (1999). Schule als Schnittstelle moderner Kinderfreundschaften - Jungen und Mädchen im Austausch von Distanz und Nähe. *ZSE, 19*, 163-187.
- Saldern, M. v. & Littig, K. E. (1987). *Landauer Skalen zum Sozioklima*. Weinheim: Beltz.
- Salisch, M. v. (1991). *Kinderfreundschaften. Emotionale Kommunikation in Konflikten*. Göttingen: Hogrefe.
- Scarlett, H. H., Press, A. N. & Crockett, W. H. (1971): Children's descriptions of peers. *Child Development, 42*, 439-453.
- Schmidt-Denter, U. (1988). *Soziale Entwicklung*. München: PVU.
- Schreier, H. (1994). *Der Gegenstand des Sachunterrichts*. Bad Heilbrunn: Klinkhardt.
- Schreier, H. (2000). Zwischen Fachbezug und Integration - Orientierungen für den Sachunterricht. In G. Löffler, V. Möhle, D. von Reeken & V. Schwier (Hrsg.), *Sachunterricht – Zwischen Fachbezug und Integration* (S.13-19). Bad Heilbrunn: Klinkhardt.
- Selman, R. L. (1984). *Die Entwicklung des sozialen Verstehens*. Frankfurt: Suhrkamp.
- Speck-Hamdan, A. (1997). Soziales Lernen und die Bedeutung der Lerngruppe. In R. Meier, H. Unglaube, & G. Faust-Siehl (Hrsg.), *Sachunterricht in der Grundschule* (S. 104-114). Frankfurt: Grundschulverband e.V.
- Stöckli, G. (1997). *Eltern, Kinder und das andere Geschlecht. Selbstwertung in sozialen Beziehungen*. Weinheim: Juventa.
- Swann, J. (1992). *Girls, boys and language*. Oxford: Blackwell.
- Thorne, B. (1996). *Gender play. Girls and boys in school*. New Brunswick, N. J.: Rutgers University Press.
- Topping, K. & Ehly, S. (Hrsg.) (1998). *Peer-assisted learning*. Mahwah, NJ.
- Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development, 62*, 1066-1078.
- Wiley, P. D. (1983). *Development of strategies for coping with peer conflict in children from first through fifth grade*. Paper presented at the Annual meeting of Association for Behavior Analysis. Milwaukee.
- Youniss, J. (1980). *Parents and peers in social development*. Chicago: University Press.

## **Inhaltsstruktur und Tendenzen der Inhalte im Sachunterricht**

**Eine empirische Bestandsaufnahme der Inhalte des Sachunterrichts in den  
70er, 80er und 90er Jahren anhand von Unterrichtslehrwerken**

### **1. Inhalte des Sachunterrichts**

#### **1.1 Probleme der Inhalte des Sachunterrichts**

Vor dreißig Jahren wurde die Heimatkunde vom Sachunterricht als Unterrichtsfach der Grundschule abgelöst. Seitdem vertritt die Didaktik des Sachunterrichts als fachdidaktische Wissenschaftsdisziplin die Belange dieses Unterrichtsbereichs, wobei sie dadurch gekennzeichnet ist, dass ihr im Gegensatz zu den anderen Fachdidaktiken die eine zugeordnete universitäre Bezugswissenschaft fehlt. Die Sachunterrichtsdidaktik unterhält statt dessen Verbindungen zu mehreren Sach-Wissenschaften, über deren Auswahl und Anzahl jedoch kein allgemeiner Konsens besteht. Dabei stellen aber nicht die Inhalte der wissenschaftlichen Bezugswissenschaften per se die Inhalte des Sachunterrichts dar. Vielmehr ist eine „allgemein-didaktische Sichtweise“ (Kaiser 1997, S. 147) leitend für die Inhaltsauswahl. Dabei „bedient“ sich der Sachunterricht u. a. von den Angeboten der Inhalte der wissenschaftlichen Bezugswissenschaften, da traditionell die Wissensbestände der Gesellschaft wissenschaftlichen Disziplinen zugeordnet werden.

Der Sachunterricht weist neben dem Korrespondenzverhältnis zu den Bezugswissenschaften und ihren Didaktiken auch Beziehungen zur Erziehungswissenschaft/ Psychologie/Soziologie/Philosophie und zur Schulpädagogik/Allgemeinen Didaktik auf. Aber auch zur Grundschulpädagogik als Stufendidaktik und zu den fächerübergreifenden Erziehungsbereichen bestehen ebenso Verbindungen wie zu den heimatkundlichen Wurzeln des Sachunterrichts. In diesen vielfältigen Bezügen liegt das besondere Kennzeichen gegenüber den anderen Fachdidaktiken. Es verleiht der Didaktik des Sachunterrichts „den höchsten Grad an Komplexität“ (Klafki 1992, S. 11). Parallel dazu steht der Sachunterricht im Gefüge der drei Bezugsgrößen von „Kind“, „Sache“ und „Gesellschaft“, die den Sachunterricht in vielfältiger Weise bestimmen.

Die Inhalte des Sachunterrichts sind von diesen vielfältigen Vernetzungen, die der Sachunterricht um sich spannt, beeinflusst. Da sowohl die Disziplin-Bezüge als auch die Bezugsgrößen unübersehbaren Veränderungen aufgrund vielfältiger Faktoren in den vergangenen 30 Jahren unterliegen, ist zu erwarten, dass sich auch die Inhalte des Sachunterrichts verändert haben. Bei den Inhalten des Sachunterrichts handelt es sich also um kein „Standardmodell“ (Kaiser 1997, S. 10), das über Jahrzehnte unverändert bleibt, sondern es sind dynamisch-vernetzte Tendenzen zu erwarten. Der Wandel der Inhalte des Sachunterrichts ist dabei als etwas Selbstverständliches zu begreifen (Meiers 1994, S. 66).

Eine Negativdefinition zu den Inhalten des Sachunterrichts (vgl. Kaiser 1997, S. 145) gibt nur eine unbefriedigende Antwort auf die Frage, welche konkreten Inhalte im Sachunterricht tatsächlich verhandelt werden. Vielmehr muss eine Positivdefinition geschaffen werden, die Auskunft über die tatsächliche Inhaltsstruktur des Sachunterrichts gibt, um auf diese Weise die konkreten Inhalte des Sachunterrichts benennen zu können.

In diesem Beitrag rücken daher zwei Aspekte in den Mittelpunkt. Auf der einen Seite geht es um die Erfassung und Beschreibung der Inhaltsstruktur des Sachunterrichts und auf der anderen Seite um die inhaltlichen Tendenzen des Faches in den vergangenen 30 Jahren.

## 1.2 Untersuchungen zu Inhalten des Sachunterrichts

Seit den 60er Jahren sind einzelne empirische Arbeiten zu den Inhalten des Sachunterrichts bzw. der Heimatkunde nachweisbar. In diesen Untersuchungen wird typischerweise mit Hilfe des inhaltsanalytischen Verfahrens versucht, anhand aufgestellter Kategoriensysteme die Inhaltsstruktur des Sachfaches der Grundschule zu ermitteln.

Die verschiedenen Untersuchungen lassen sich aufgrund ihrer Intentionen zwei Gruppen zuordnen. Bei einigen Arbeiten wird alleinig das Ziel verfolgt, eine Momentaufnahme der inhaltlichen Verteilung zu erhalten (vgl. u.a. Höcker 1968; Haug 1969; Thun in Katzenberger 1972; Mitzlaff 1985; Breitschuh 1997; Einsiedler 1998; Strunck, Lück & Demuth 1998), während bei zwei weiteren Untersuchungen zusätzlich auch die Entwicklungslinien für einen ausgewählten Zeitabschnitt im Mittelpunkt stehen (Schreier 1979; Einsiedler & Schirmer 1986).

Dabei berücksichtigen die verschiedenen inhaltsanalytischen Arbeiten unterschiedliche Datenmaterialien (Lehrpläne, Schulbücher, Lehrberichte oder Schülerarbeitsmappen), verschiedene Klassenstufen sowie unterschiedliche Regionen (einzelne Städte, Landkreise, Bundesländer, Bundesrepublik Deutschland). Differenzen liegen auch bei der Quantität des Untersuchungsmaterials vor; umfangreicher angelegte Untersuchungen (u.a. Schreier 1979; Mitzlaff 1985) und kleinere Vorha-

ben mit sehr begrenztem Datenmaterial (u.a. Höcker 1968; Breitschuh 1997) existieren parallel.

### 1.3 Ansatz und Fragestellung

Die vorliegende Arbeit<sup>1</sup> ordnet sich in die Tradition der oben erwähnten inhaltsanalytischen Untersuchungen ein. Dadurch, dass sowohl Entwicklungslinien aufgezeigt werden als auch eine absolute Erfassung der Verteilung geleistet wird, kann sie als eine zeitliche Aktualisierung und Fortschreibung der Arbeiten von Schreier und Einsiedler & Schirmer verstanden werden.

Als zeitlicher Rahmen für die inhaltlichen Tendenzen wurde ein 30jähriger Untersuchungszeitraum gewählt. Dadurch wird ein vollständiger Überblick über die Inhaltsstruktur des gesamten Sachunterrichts seit der Einführung dieses Faches um 1970 bis heute angestrebt.

Die Entscheidung für die Methode der Inhaltsanalyse erfolgt sowohl aufgrund traditioneller als auch anhand methodologischer Erwägungen. Sie zeichnet sich vor allem dadurch aus, dass Inhalte quantitativ erfasst werden können, so wie es für das Ziel dieser Untersuchung erforderlich ist. Obwohl Datenmaterial ab Ende der 60er Jahre über die Inhalte des Sachunterrichts bzw. der Heimatkunde vorliegt (vgl. Abschnitt 1.2), ist eine Vergleichbarkeit des existierenden Datenmaterials, um inhaltliche Tendenzen des Faches seit der Etablierung des Sachunterrichts in der Grundschule vor 30 Jahren zu ermitteln, nicht gegeben. Als problematisch erweisen sich vor allem die verschiedenen Kategorienschemata, die bei den einzelnen Inhaltsanalysen unterschiedlich aufgestellt und umgrenzt sind. Zuordnungsunterschiede bei den konkreten Inhalten zu den einzelnen Inhaltsbereichen des Kategoriensystems sind darüber hinaus vorhanden. Größtenteils werden in den Untersuchungen keine genauen Festlegungen der einzelnen Kategorien mitveröffentlicht, so dass nicht mehr nachvollzogen werden kann, welche konkreten Inhalte in welchem Inhaltsbereich erfasst worden sind. Auch sorgen das unterschiedlich gewonnene Datenmaterial, die regionale Umgrenzung, die Klassenstufenauswahl und der Umfang des Untersuchungsmaterials dafür, dass eine Vergleichbarkeit der Ergebnisse der vorliegenden inhaltsanalytischen Arbeiten mit kaum lösbaren Schwierigkeiten verbunden ist.

Nur eine neue, nach übereinstimmenden Kriterien vorgenommene Datenerfassung kann diese Problemlage überwinden. Deshalb werden in dieser Arbeit sowohl einheitliche Inhaltsbereiche für den gesamten Untersuchungszeitraum festgelegt als auch deren inhaltlicher Rahmen genau bestimmt. Zu bedenken ist auch, dass die konkreten Aussagen für einen Inhaltsbereich nur dann Aussagekraft haben, wenn berücksichtigt wird, aus welchen einzelnen Teilbereichen sich ein Inhaltsbereich zusammensetzt. Beispielsweise muss klar benannt werden, welche konkre-

ten Inhalte im historischen Inhaltsbereich erfasst werden. Die bereits vorliegenden Untersuchungen sind hier unterschiedlich vorgegangen (z. B. wird das Thema „Zeit“ unterschiedlich dem historischen Bereich oder einem anderen Bereich zugeordnet).

Diese empirische Arbeit über die Inhalte des Sachunterrichts stellt zwei Fragestellungen in den Mittelpunkt. Auf der einen Seite geht es um die Erfassung der inhaltlichen Verteilung im Sachunterricht. Fortführend steht aber die Frage nach den Tendenzen der Inhalte des Sachunterrichts in den vergangenen 30 Jahren im Mittelpunkt. Dabei verstehe ich diese Arbeit als einen Beitrag zur „Anatomie und Morphologie“ (Löffler, Möhle, von Reeken & Schwier 2000, S. 220) des Sachunterrichts.

## 2. Methodologische Aspekte

### 2.1 Lehrwerke als Datenmaterial

Bei der Festlegung des Untersuchungsmaterials fiel die Entscheidung auf Unterrichtslehrwerke des Sachunterrichts.

Andere, zunächst geeignet erscheinende Untersuchungsmaterialien (z. B. Lehrpläne, Lehrberichte, Schülerarbeitsmappen) kamen aus unterschiedlichen Gründen nicht in Frage. Lehrpläne werden nur in unregelmäßigen Zeiträumen in den einzelnen Bundesländern veröffentlicht und können somit keine kontinuierlichen Tendenzen aufzeigen. Ferner sind Lehrpläne der Bundesländer untereinander nur schwer vergleichbar, da sie neben unterschiedlichen Strukturen auch regionalen Einflüssen und Traditionen unterliegen. Die Analyse von Lehrberichtsheften für einen Zeitraum von 30 Jahren ist aufgrund der problematischen Beschaffung dieser Hefte nicht möglich, da diese in der Regel nur einige Jahre von den Schulen bzw. Schulaufsichtsbehörden aufbewahrt werden und somit vor allem für die 70er und 80er Jahre kaum mehr in ausreichender Anzahl beschafft werden könnten. Das Problem der Materialbeschaffung existiert ebenso für die Schülerarbeitsmappen.

Hingegen sind Lehrwerke des Sachunterrichts der vergangenen 30 Jahre in Bibliotheken und Schulbuchsammlungen auffindbar und damit als Datenmaterial leicht zugänglich. Zudem nehmen die untersuchten Unterrichtswerke für sich in Anspruch, die Inhalte eines gesamten Schuljahres zu umfassen. Die Inhalte des Sachunterrichts sind in den Schulbüchern konkret enthalten, so wie es für die inhaltsanalytische Zuordnung unabdingbar ist.

Ausgewählte Sachunterrichtslehrwerke bilden in meiner Untersuchung das Datenmaterial, um Informationen über die inhaltliche Struktur des Sachunterrichts und ihre Veränderungen im Untersuchungszeitraum zu erhalten (vgl. Abschnitt 2.4).

Dieses Datenmaterial kann zwar nicht die Inhalte der differenten Unterrichtsgestaltung in der Praxis des Sachunterrichts in der täglichen Grundschularbeit repräsentieren, aber die Schulbücher können als Spiegel der Sachunterrichts-Diskussion herangezogen werden, um historische Tendenzen der inhaltsbezogenen Diskussion der Sachunterrichtsdidaktik zu erhalten.

## 2.2 Auswahl der Schulbücher

Eine Auswahl von Schulbüchern für die Analyse ist unerlässlich, da aufgrund des umfangreichen Angebots der Schulbuchverlage nicht alle Werke, die für den Sachunterricht im Untersuchungszeitraum konzipiert wurden, untersucht werden können. Ein vollständiger Überblick über den Schulbuchbestand des Sachunterrichts für den Untersuchungszeitraum (1970-1999) ist daher für eine erforderliche Stichprobenentnahme notwendig. Dieser Überblick sollte Aufschluss geben können über den Verbreitungsgrad und den Veröffentlichungszeitpunkt einzelner Lehrwerkreihen. Da aber für Schulbücher weder „Best-“ noch „Longseller-“ Listen existieren, die als Grundlage herangezogen werden könnten, noch Verlagsprospekte Aufschluss über diese Aspekte geben, wurden die offiziellen Schulbuchverzeichnisse der Bundesländer der vergangenen 30 Jahre verwendet, um einen Bestandsüberblick zu erhalten.

In allen Bundesländern müssen Schulbücher, die nicht Verbrauchsmaterial sind, von den Kultusministerien der Länder zugelassen werden. In Schulbuchverzeichnissen geben die Ministerien der einzelnen Bundesländer die Werke bekannt, die im Sachunterricht eingesetzt werden dürfen. Diese Listen erscheinen bereits seit vielen Jahren (unterschiedlich in den einzelnen Bundesländern) und decken weitgehend den zu berücksichtigenden Zeitraum ab. Die Zusammenstellung der einzelnen Listen gibt einen vollständigen Überblick über die zugelassenen Sachunterrichtswerke.

Aus dieser Übersicht muss eine Auswahl getroffen werden. Ziel ist es, die Werke zu ermitteln, die in besonders vielen Bundesländern verbreitet waren bzw. sind. Eine bewusste Auswahl (vgl. Merten 1995, S. 283-287) wird somit vorgenommen. Mit Hilfe des Verfahrens der Quotenauswahl (vgl. Merten 1995, S. 283-287) wird in einem aufgestellten Quotenplan festgelegt, dass diejenigen Werke ermittelt werden sollen, bei denen davon ausgegangen werden kann, dass diese von besonders vielen Schülern und Schülerinnen im Sachunterricht benutzt worden sind.

Der 30jährige Untersuchungszeitraum wird in sechs Zeitblöcke eingeteilt, die jeweils fünf Jahre umfassen (Zeitblöcke A bis F). Ebenfalls wird festgelegt, dass je zwei Werkreihen für jeden der Zeitblöcke ausgewählt werden, die jeweils aus drei Bänden für das 2., 3. und 4. Schuljahr bestehen. Werke für das 1. Schuljahr sind in der Regel nicht zulassungspflichtig, da diese als Verbrauchsmaterialien konzipiert

sind. Zudem stehen diese nicht für alle Lehrwerkreihen zur Verfügung, so dass das 1. Schuljahr in dieser Untersuchung unberücksichtigt bleibt.

Insgesamt werden zwölf Schulbuchreihen mit insgesamt 36 Einzelbänden berücksichtigt. Dabei wird jeweils die erste Auflage eines Werkes für die Untersuchung herangezogen. Dieser Umfang des Datenmaterials ist im Rahmen der Arbeit zu bewältigen und berücksichtigt zugleich eine notwendige Breite des Materials.

Folgende Werkreihen wurden für die Untersuchung ausgewählt:

Zeitblock A 1970-1974	Zeitblock B 1975-1979	Zeitblock C 1980-1984	Zeitblock D 1985-1989	Zeitblock E 1990-1994	Zeitblock F 1995-1999
Arbeitsbuch Sachunter- richt Diesterweg	Sachunter- richt Westermann	Wir entdecken unsere Welt Diesterweg	Sachbuch Oldenbourg	Mobile Westermann	Tausendfüßler Auer
Grundschul- buch Kamp	CVK-Sach- buch CVK	Sach- und Machbuch Hirschgraben	Im Blick- punkt Schroedel	Pusteblyume Schroedel	Bücherwurm Klett

Abb. 1: Übersicht über die für die Inhaltsanalyse ausgewählten Sachunterrichtslehrwerke

### 2.3 Kategorienschema

Das Kategorienschema dieser Inhaltsanalyse besteht aus zwölf Inhaltsbereichen, die sich jeweils aus einer unterschiedlichen Anzahl von Teilbereichen zusammensetzen (siehe Abb. 2). Die Inhaltsbereiche und die zugeordneten Teilbereiche sind aufgrund intensiver Pretests mit dem Datenmaterial festgelegt worden. Diese sind in einem Wechselverhältnis zwischen theoretischen Überlegungen aufgrund der Fragestellungen und dem konkreten Material entstanden.

Die Inhaltsbereiche bestehen einerseits aus Inhalten, die eine Nähe zu den Bezugswissenschaften des Sachunterrichts aufweisen (Kategorien 1-6) sowie aus ausgewählten fächerübergreifenden Bereichen (Inhaltsbereiche 7-11), die sich im Pretest als die bedeutsamsten für den Sachunterricht herausgestellt haben. In der Fächergruppe der fächerübergreifenden Erziehungsbereiche werden die Inhalte zusammengefaßt, die auch dem Sachunterricht zugeordnet werden können, dessen allgemein-übergreifende Zielsetzungen aber ebenso für die anderen Grundschulfächer resp. -lernbereiche formuliert werden (z. B. Gesundheitserziehung; Interkulturelles Lernen). Eine notwendige Restkategorie wurde eingerichtet, um die Anforderungen der Vollständigkeit zu erfüllen.



	Inhaltsbereiche	Teilbereiche
1.	biologisch orientierte Inhalte	•Tiere •Pflanzen •natürliche Lebensräume •andere biologisch orientierte Inhalte
2.	physikalisch, chemisch und technisch orientierte Inhalte	•Thermometer •Elektrizität/Stromkreis •Magnetismus •Licht und Schatten •Gewicht/Wippe/Waage •Luft •Wasser •Feuer •andere physikalisch und chemisch orientierte Inhalte •technische Gegenstände herstellen • technische Geräte kennen lernen •technische Aspekte der Versorgung •andere technisch orientierte Inhalte
3.	geografisch orientierte Inhalte	•Grundlegende Kartenarbeit •regionalgeografisch orientierte Inhalte •sozialgeografisch orientierte Inhalte • naturgeografisch orientierte Inhalte •astronomisch orientierte Inhalte •meteorologisch orientierte Inhalte •andere geografisch orientierte Inhalte
4.	historisch-chronometrisch orientierte Inhalte	•Inhalte zur Förderung der Zeitbegriffsbildung •technikgeschichtlich orientierte Inhalte •sozialgeschichtlich orientierte Inhalte •regionalgeschichtlich orientierte Inhalte •historische Gesellschaften und Kulturen •andere historisch orientierte Inhalte
5.	ökonomisch orientierte Inhalte	• öffentliche Einrichtungen/Ver- und Entsorgungsunternehmen •Einkaufen/Handel •Werbung •Warenproduktion in Handwerk und Industrie •Landwirtschaft •andere ökonomisch orientierte Inhalte
6.	sozial orientierte Inhalte	•Familie/Wohnen •Schule •Freizeit/Spielen •Leben mit anderen Menschen •andere sozial orientierte Inhalte
7.	umwelterzieherische Inhalte	•Abfall/Müll •Wasser •Luft •Lärm •ökologisch-biologisch orientierte Inhalte •andere umwelterzieherische Inhalte
8.	verkehrserzieherische Inhalte	•Sicherheitserziehung •andere verkehrserzieherische Inhalte
9.	gesundheits- und sexualerzieherische Inhalte	•Sexualerziehung •Ernährung •Körper kennen lernen •Gesundheit erhalten und fördern •bei Krankheiten helfen • andere gesundheitserzieherische Inhalte
10.	medienerzieherische Inhalte	•Bücher und Zeitungen (Printmedien) •Film und Fernsehen •andere medienorientierte Inhalte
11.	Interkulturelles, Europäisches und Eine-Welt-Lernen	• Interkulturelles Lernen •Europäisches Lernen •Eine-Welt-Lernen • andere Inhalte des Inhaltsbereichs
12.	andere Inhalte	•alle Inhalte, die nicht den 11 anderen Inhaltsbereichen zugeordnet werden können

Abb. 2: Kategorienschema der zwölf Inhaltsbereiche und der ihnen zugeordneten Teilbereiche

Dabei werden die Bezeichnungen der ersten sechs Inhaltsbereiche aus den nahestehenden Bezugswissenschaften gewählt, um die Zusammenschlüsse der einzelnen Teilbereiche auch treffend benennen zu können. Zur Abgrenzung, dass es sich nicht um biologische Inhalte der Fachdisziplin handelt, die im Sachunterricht berücksichtigt werden, sondern um Inhalte, die eine inhaltliche Nähe zur Biologie aufweisen, wurde die Begrifflichkeit entsprechend formuliert (z. B. „biologisch orientierte Inhalte“).

Die zwölf Inhaltskategorien sind wechselseitig exklusiv definiert und voneinander unabhängig. Es liegen über die Angaben in Abb. 2 zusätzlich exakte Definitionen und umfassende Beispiele (erschöpfende Aufzählung) für alle Inhaltsbereiche vor, denn nur eine einheitliche und genau ausgeführte Zuweisung der konkreten Inhalte zu den Inhaltsbereichen ermöglicht zuverlässige Trendergebnisse für den 30jährigen Untersuchungszeitraum.

#### 2.4 Inhaltsanalytische Vorgehensweise

Ohne vertiefend auf inhaltsanalytische Methodenprobleme an dieser Stelle eingehen zu können, werden im weiteren Verlauf dieses Abschnitts einzelne Aspekte der Inhaltsanalyse, die für meine Arbeit grundlegend sind, aufgezeigt.

Eine vielfach zitierte Definition der Inhaltsanalyse bildet den Ausgangspunkt: „Inhaltsanalyse ist eine Methode zur Erhebung sozialer Wirklichkeit, bei der von Merkmalen eines manifesten Textes auf Merkmale eines nichtmanifesten Kontextes geschlossen wird.“ (vgl. Merten 1995, S. 59). Gemäß dieser Definition ist es Ziel meiner inhaltsanalytischen Arbeit, die manifesten Inhalte (vgl. Merten 1995, S. 56-57) aus dem Datenmaterial (Schulbücher des Sachunterrichts) herauszufiltern, sie statistisch zu erfassen, um letztlich Antworten auf die formulierten Fragestellungen zu erhalten. Die Inhaltsanalyse ist eine schlussfolgernde Methode, die in dieser Arbeit Aussagen über Tendenzen der inhaltsbezogenen Diskussion der Sachunterrichtsdidaktik ermöglicht.

Bei dieser inhaltsanalytischen Arbeit wird bei der Datenerfassung mit der Häufigkeitsanalyse (Frequenzanalyse) gearbeitet. Ein Kategoriensystem, das alle vorfindbaren Inhalte erfassen kann, muss dafür aufgestellt und definiert werden (vgl. Abb. 2). Grundlage bildet die Fläche, die die Inhalte einer Kategorie auf einer Schulbuchseite (Analyseeinheit) einnehmen. Bei dieser Raumanalyse wird als Zählseinheit 1/10 Seite einer Schulbuchseite festgelegt. Jede Zählseinheit wird einem der 12 Inhaltsbereiche zugeordnet. Insgesamt werden dabei 3187,2 Seiten mit entsprechend 31872 Zählseinheiten in der Untersuchung berücksichtigt. Die Wiederholung der Zuordnung (Re-Test) ergab dabei eine Übereinstimmung beider von mir durchgeführten Datenerfassungen von 99,29%; d. h., 227 Zählseinheiten

sind bei der Zweiterfassung anderen Inhaltsbereichen zugeordnet worden. Die Häufigkeiten werden differenziert nach den sechs Zeitblöcken und den zwölf Inhaltsbereichen erfasst, so dass sich entsprechende Veränderungen errechnen und somit statistisch darstellen lassen. Daraus lassen sich wiederum Aussagen über inhaltliche Tendenzen des Sachunterrichts im Untersuchungszeitraum ablesen. Diese werden dabei als Spiegel der Sachunterrichts-Diskussion betrachtet, so dass an dieser Stelle Aussagen über den Status quo und Tendenzen der inhaltsbezogenen Diskussion der Sachunterrichtsdidaktik gegeben werden können. Es wird also keinesfalls versucht, hier eine allseits gültige Inhaltsverteilung der Praxis des Sachunterrichts aufzuzeigen<sup>2</sup>. Vielmehr geht es darum, eine dynamische Beschreibung der Inhaltsstruktur des Sachunterrichts zu geben, wie sie sich im Zusammenhang mit der didaktischen Diskussion im Sachunterricht nachweisen lässt.

### 3. Ergebnisse der Inhaltsanalyse

Zur besseren Übersicht werden die zwölf Inhaltsbereiche in drei Abschnitten (naturwissenschaftlich orientierte Inhaltsbereiche [1-2], sozialwissenschaftlich orientierte Inhaltsbereiche [3-6] und fächerübergreifende Erziehungsbereiche inklusive der Restkategorie [7-12]) ausgewertet.

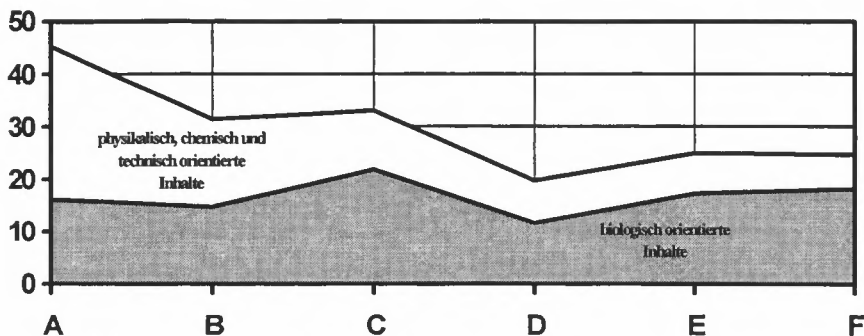
#### 3.1. Naturwissenschaftlich orientierte Inhaltsbereiche

Der Prozentanteil am Sachunterricht nimmt bei der naturwissenschaftlichen Fächergruppe im Verlauf des Untersuchungszeitraums ab (vgl. Abb. 3). Diese Fächergruppe bestreitet aktuell in den 90er Jahren ungefähr 1/4 der Inhalte des gesamten Sachunterrichts; zu Beginn des Untersuchungszeitraums waren es noch über 45% der Inhalte. Ende der 80er Jahre ist der Anteil der naturwissenschaftlich orientierten Inhaltsbereiche besonders gering und macht weniger als 20% der Sachunterrichtsinhalte aus. Im Gegenzug ist in dieser Zeit der Anteil der sozialwissenschaftlichen Inhaltsbereiche besonders umfangreich (vgl. Abb. 4).

Bildet der physikalisch, chemisch und technisch orientierte Inhaltsbereich zu Beginn der 70er Jahre mit knapp 30% noch den größten aller zwölf Inhaltsbereiche, so nimmt er in der zweiten Untersuchungshälfte nur noch ca. 1/4 dieses Anteils ein und belegt damit lediglich einen mittleren Rangplatz beim Vergleich der einzelnen Inhaltsbereiche. Bis Mitte der 80er Jahre gehört dieser Inhaltsbereich noch zu den drei umfangreichsten Inhaltsbereichen des Sachunterrichts. In den Zeitblöcken D bis F liegt der Anteil der physikalisch, chemisch und technisch orientierten Inhalte von 6,5% bis 8% der Inhalte des Sachunterrichts.

Während die Anteile des Inhaltsbereichs 2 im Untersuchungszeitraum deutlich fal-

len, schwankt der Anteil des biologisch orientierten Inhaltsbereichs ohne erkennbare Tendenzen zwischen 12% und 22% der Inhalte des Sachunterrichts. Damit gehört der Inhaltsbereich 1 während des gesamten Untersuchungszeitraums zu den umfangreicheren Inhaltsbereichen und belegt in beiden Zeitblöcken der 90er Jahre den ersten Rangplatz aller 12 Inhaltsbereiche. Biologisch orientierte Inhalte sind somit im Sachunterricht kontinuierlich umfangreich vertreten und bilden einen der inhaltlichen Grundpfeiler des Sachunterrichts während des gesamten Untersuchungszeitraums.



Inhaltsbereiche 1-2:	A	B	C	D	E	F
biologisch orientierte Inhalte	16,14	14,75	21,87	11,60	17,26	18,20
physikalisch, chemisch und technisch orientierte Inhalte	29,10	16,60	11,17	8,12	7,69	6,48
Gesamt	45,24	31,35	33,04	19,72	24,95	24,68

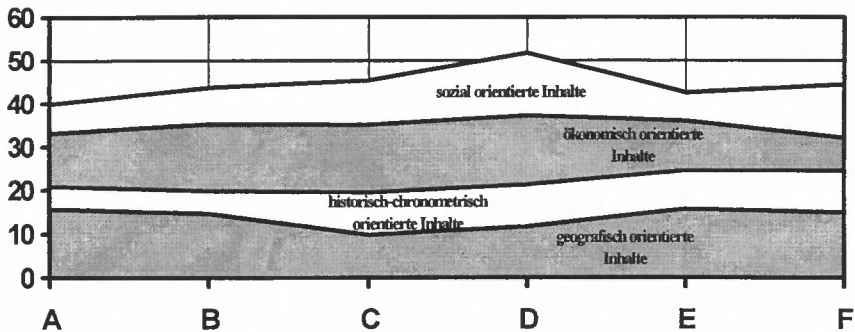
Abb. 3: Anteile der beiden naturwissenschaftlich orientierten Inhaltsbereiche am Gesamtanteil der Inhalte des Sachunterrichts in Prozent

### 3.2 Sozialwissenschaftlich orientierte Inhaltsbereiche

Die Fächer der sozialwissenschaftlich orientierten Bezugsfächer steigern ihre Anteile (vgl. Abb. 4) kontinuierlich von 40% auf 52% im Verlauf der ersten beiden Untersuchungsdekaden. Ende der 80er Jahre bestreitet diese Fächergruppe mehr als die Hälfte aller Inhalte des Sachunterrichts. In den 90er Jahren verringert sich der Anteil wieder und nimmt 43% bzw. 44% aller Sachunterrichtsinhalte ein.

Die Anteile des geografisch orientierten Inhaltsbereichs am Gesamt der Inhalte des Sachunterrichts schwanken während des gesamten Untersuchungszeitraums zwischen 10% und 15%. In den 70er Jahren liegt der Anteil bei 15% der Inhalte des Sachunterrichts. Die 80er Jahre sind durch einen Rückgang dieser Inhalte auf nunmehr 10% gekennzeichnet, obwohl die sozialwissenschaftliche Fächergruppe gera-

de in dieser Dekade ihre höchsten Anteile verbuchen kann. In den 90er Jahren ist



Inhaltsbereiche 3-6:	A	B	C	D	E	F
geografisch orientierte Inhalte	15,58	14,61	9,79	11,70	15,61	14,76
historisch-chronometrisch orientierte Inhalte	5,23	5,22	9,70	9,56	8,89	9,67
ökonomisch orientierte Inhalte	12,21	15,36	15,55	15,91	11,43	7,50
sozial orientierte Inhalte	6,85	8,53	10,29	14,53	6,57	12,36
Gesamt	39,87	43,72	45,33	51,70	42,50	44,29

Abb.4: Anteile der vier sozialwissenschaftlich orientierten Inhaltsbereiche am Gesamtanteil der Inhalte des Sachunterrichts in Prozent

wieder das Niveau der 70er Jahre von 15% der Inhalte des Sachunterrichts erreicht. Damit nimmt dieser Inhaltsbereich nach dem biologisch orientierten Inhaltsbereich in den 90er Jahren den zweiten Rangplatz aller zwölf Inhaltsbereiche ein. Während der Anteil der historisch-chronometrisch orientierten Inhalte in den 70er Jahren 5% der gesamten Unterrichtsinhalte des Sachunterrichts ausmacht, steigt dieser mit Beginn der 80er Jahre und nimmt in den beiden letzten Dekaden des Untersuchungszeitraums kontinuierlich 9% der Inhalte des Sachunterrichts ein. In den 90er Jahren belegt der Inhaltsbereich 4 in beiden Zeitblöcken dieser Dekade den 5. Rangplatz aller zwölf Inhaltsbereiche.

Zu Beginn des Untersuchungszeitraums können 12% der Sachunterrichtsinhalte dem ökonomisch orientierten Inhaltsbereich zugeordnet werden. Im Zeitblock B und in den 80er Jahren nimmt dieser Bereich kontinuierlich jeweils 15% der Inhalte ein und belegt damit in dieser Zeit den 1. bzw. 2. Rangplatz aller ausgewiesenen Inhaltsbereiche. Die 90er Jahre sind durch einen Rückgang der ökonomisch orientierten Inhalte gekennzeichnet. Im Block E fällt dieser Bereich auf zunächst 11% und verringert seine Anteile am Ende des Untersuchungszeitraums weiter auf nunmehr 7,5%, so dass dieser Inhaltsbereich nur noch den 4. bis 6. Rangplatz aller Inhaltsbereiche belegt.

Die Inhalte des sozial orientierten Inhaltsbereichs steigen in den 70er und 80er Jahren kontinuierlich von 7% auf 15% der Inhalte und sind damit im besonderen Maße für den Anstieg der gesamten sozialwissenschaftlichen Fächergruppe in diesem Zeitraum verantwortlich. In den 90er Jahren zeigt sich kein einheitliches Bild, hier liegen die Anteile bei 7% und 12% der Inhalte.

### 3.3 Inhaltsbereiche der fächerübergreifenden Erziehungsbereiche

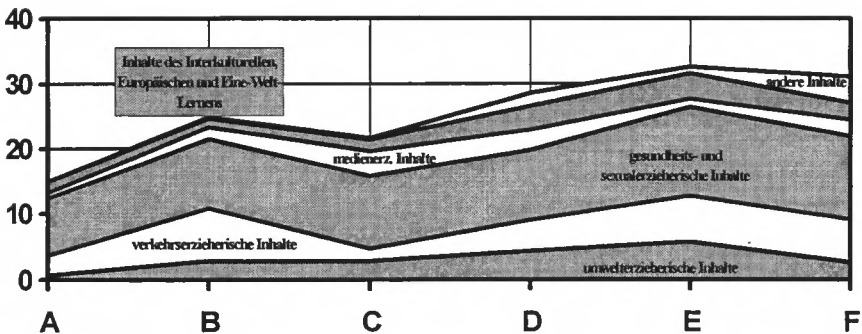
Die Inhaltsbereiche der fächerübergreifenden Erziehungsbereiche und der Inhaltsbereich, der die „anderen Inhalte“ erfasst, können zusammengefasst ihre Anteile im Verlauf des Untersuchungszeitraums deutlich ausbauen. Nehmen sie in der ersten Hälfte der 70er Jahre weniger als 15% der Sachunterrichtsinhalte ein, so sind es in den 90er Jahren über 30%. Nahezu 1/3 der Inhalte des Sachunterrichts werden in der letzten Untersuchungsdekade aus diesen Inhaltsbereichen bestritten und weisen keine Bezugsfachorientierung auf.

Umwelterzieherische Inhalte steigern ihre Anteile kontinuierlich im Untersuchungszeitraum von weniger als 1% im Zeitblock A auf knapp 6% im Zeitblock E. Danach stellt sich ein Rückgang der Anteile auf 3% der Inhalte am Ende des Untersuchungszeitraums ein.

Die verkehrserzieherischen Inhalte schwanken in den 70er und 80er Jahren zwischen 2% und 8% der Inhalte des Sachunterrichts. In den 90er Jahren nehmen sie zwischen 6% und 7% ein.

Gesundheits- und sexualerzieherische Inhalte bauen ihren Anteil im Verlauf des Untersuchungszeitraums kontinuierlich von 9% auf 13% aus. Sie bilden damit während des gesamten Untersuchungszeitraums den umfangreichsten Inhaltsbereich innerhalb der Fächergruppe der fächerübergreifenden Erziehungsbereiche. Der medienerzieherische Inhaltsbereich konnte seine Anteile von den 70er Jahren bis Ende der 80er Jahre von 1% auf 3% erweitern. In den 90er Jahren liegt der Anteil wiederum unterhalb des Anteils der 80er Jahre.

Interkulturelles, Europäisches und Eine-Welt-Lernen nimmt in der ersten Untersuchungshälfte einen deutlich geringeren Anteil an den Inhalten des Sachunterrichts ein als in der zweiten Hälfte. In den Zeitblöcken D bis F sind es 3% bis 4% der Sachunterrichtsinhalte, die dem Inhaltsbereich 11 zugeordnet werden können. Die Inhalte des Inhaltsbereichs 12 (andere Inhalte) sind in der ersten Hälfte des Untersuchungszeitraums mit weniger als 0,3% der Inhalte des Sachunterrichts kaum vertreten. In der zweiten Untersuchungshälfte hat sich dieser Anteil deutlich erhöht, so dass am Ende des Untersuchungszeitraums 4% der Inhalte des Sachunterrichts keinem der 11 inhaltlich ausgerichteten Inhaltsbereiche zugeordnet werden.



Inhaltsbereiche 7-12:	A	B	C	D	E	F
umwelterzieherische Inhalte	0,60	2,84	2,83	4,33	5,70	2,59
verkehrserzieherische Inhalte	3,10	7,97	1,85	4,69	6,99	6,51
gesundheits- und sexualerzieherische Inhalte	8,60	10,66	11,11	10,73	13,73	12,77
medienerzieherische Inhalte	0,92	1,87	3,82	3,14	1,28	2,59
Inhalte des Interkulturellen, Europäischen und Eine-Welt-Lernens	1,67	1,32	1,75	3,69	3,85	2,59
andere Inhalte	0,00	0,27	0,27	2,00	1,00	3,98
Gesamt	14,89	24,93	21,63	28,58	32,55	31,03

Abb. 5: Anteile der fünf fächerübergreifenden Erziehungsbereiche am Gesamtanteil der Inhalte des Sachunterrichts in Prozent

### 3.4 Anteile und Veränderungen der drei Fächergruppen

Der direkte Vergleich der drei Fächergruppen und der zwölf Inhaltsbereiche (vgl. Abb. 6) veranschaulicht die Gesamttendenzen, die sich für die Inhalte des Sachunterrichts im 30jährigen Untersuchungszeitraum ergeben. Folgende Tendenzen können für die Inhalte des Sachunterrichts in den drei Dekaden benannt werden:

- Die Stellung naturwissenschaftlich orientierter Inhalte des Sachunterrichts wird durch den Rückgang der physikalisch, chemisch und technisch orientierten Inhalte im Untersuchungszeitraum geschwächt.
- Sozialwissenschaftlich orientierte Inhalte stellen kontinuierlich ab Mitte der 70er Jahre die umfangreichste Fächergruppe des Sachunterrichts mit über 40% der Inhalte des Sachunterrichts dar.
- Die Anteile der fächerübergreifenden Erziehungsbereiche und der anderen Inhalte bauen zusammengefasst ihre Anteile im Sachunterricht während des Untersuchungszeitraums deutlich aus und bestimmen Ende der 90er Jahre knapp 1/3 der Inhalte des Sachunterrichts. Somit ist diese Fächergruppe im Sachunterricht der 90er Jahre umfangreicher vertreten als die beiden naturwissenschaftlichen Inhaltsbereiche. Der Sachunterricht bestreitet aktuell weniger

als 70% der Inhalte aus den bezugswissenschaftlich orientierten Inhaltsbereichen.

Zusammengefasst kann folgende inhaltliche Tendenz für die Inhalte des Sachunterrichts in den 30 Jahren beschrieben werden:

Der Ausbau der Anteile der fächerübergreifenden Inhalte während des Untersuchungszeitraums korrespondiert mit einer Reduzierung der bezugsfachorientierten Anteile des Sachunterrichts. Während die sozialwissenschaftlich orientierten Inhalte ihre Anteile weitgehend behaupten können, verringern sich die Anteile der naturwissenschaftlich orientierten Fächergruppe.

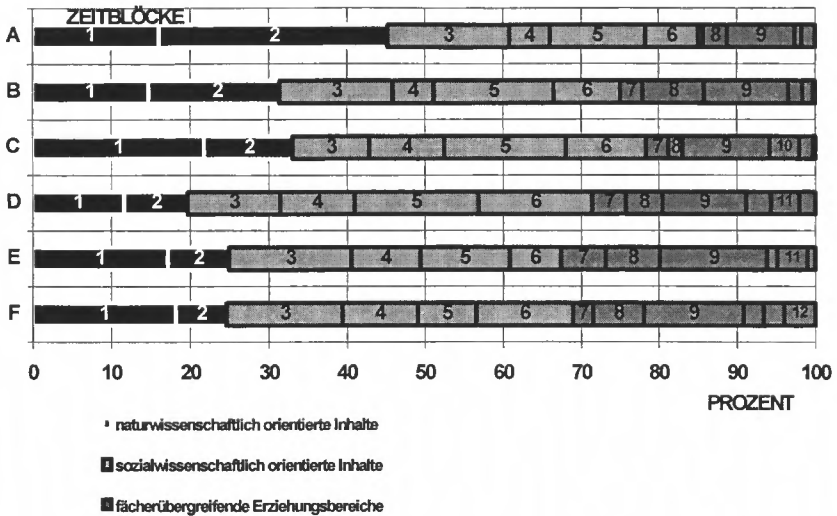


Abb. 6: Prozentuale Verteilung auf die zwölf Inhaltsbereiche geordnet nach den drei Fächergruppen

#### 4. Diskussion der Ergebnisse

Die Ergebnisse der Untersuchung verdeutlichen, dass die Inhalte des Sachunterrichts in den 30 Jahren der Existenz des Faches nicht konstant bleiben, sondern Veränderungen der Inhaltsberücksichtigung zu erkennen sind. Auch wenn einige Inhalte durchaus kontinuierlich in allen Zeitblöcken der Untersuchung berücksichtigt werden, sind deutliche Veränderungen der Inhaltsbereiche sichtbar und beschreibbar (vgl. Abschnitt 3). Dabei spiegeln die Ergebnisse die didaktischen Trends der jeweiligen Zeit wider. So sind in der Phase der Wissenschaftsorientierung zu Beginn der 70er Jahre die Anteile der bezugsfachorientierten Inhaltsbereiche besonders umfangreich (u.a. physikalisch, chemisch und technisch orientierte Inhalte), während Ende der 80er Jahre die sozialwissenschaftlich orientierten



tierten Inhalte ausgeprägter vertreten sind (Phase der sozialen Wende des Sachunterrichts). In den 90er Jahren werden die fächerübergreifenden Erziehungsbereiche verstärkt diskutiert und entsprechend bei den Inhalten des Sachunterrichts stärker berücksichtigt.

Die Empfänglichkeit der Inhalte des Sachunterrichts, auf Trends zu reagieren (vgl. Abschnitt 3.4), liegt u.a. an der starken Vernetzung des Faches mit anderen Bereichen (vgl. Abschnitt 1). Durch das fehlende fest umrissene Curriculum des Sachunterrichts können aktuelle Inhaltstrends relativ kurzfristig in den Sachunterricht aufgenommen werden. Diese Flexibilität der Inhalte des Sachunterrichts ist auf der einen Seite als Gefahr anzusehen, auf der anderen Seite aber durchaus auch als Chance zu betrachten.

Diese Chance bietet u.a. die Möglichkeit, dass die Inhalte des Sachunterrichts aktuell und zeitauthentisch bleiben können, sowie dass dabei insgesamt weniger gegen verfestigte Strukturen angekämpft werden muss, als dieses in anderen Fächern die Situation ist.

Andererseits kann ein fehlendes festes Curriculum, das letztlich die Beliebigkeit der Inhalte bedeutet, das gesamte Fach in eine Legitimationskrise führen, worauf in der Literatur (u. a. Löffler u.a. 2000, S. 219; Feige 2000, S. 75; Schreier 2000, S. 13-14; Breitschuh 1997; Marquardt-Mau 1998, S. 10) bereits mehrfach hingewiesen wurde. Dadurch ist aber auch eine Anfälligkeit für nicht immer wünschenswerte Trends deutlicher gegeben. An dieser Stelle sei auf die sich abzeichnende Reduzierung des Umfangs des Sachunterrichts an der Studentafel des Grundschulunterrichts verwiesen, bei der durchaus ein Zusammenhang zur ungefestigten Inhaltsstruktur des Faches hergestellt werden kann. Einen Hinweis auf die aktuelle Situation des Sachunterrichts im Fächerkanon der Grundschule erhält man etwa dadurch, dass in einem aktuellen Werk zur Grundschulpädagogik und -didaktik (Haarmann & Kalb 1999) dem Sachunterricht keiner der 34 dort enthaltenen Beiträge gewidmet ist, sondern dieser Lernbereich unter der Überschrift „Fächerübergreifendes Lernen in der Grundschule“ lediglich mitberücksichtigt wird.

Die Anatomie der Inhalte des Sachunterrichts der vergangenen 30 Jahre zeigt, dass sowohl kontinuierliche Inhalte vorhanden sind, als auch die Inhalte des Sachunterrichts auf aktuelle Trends reagieren. Dabei sind die Inhalte des Sachunterrichts beschreibbar als eine Addition fester und variabler Inhalte, die sich zwischen den Polen von Kontinuität und Wandel einordnen lassen. Neben den sich im Wandel befindenden Inhalten des Sachunterrichts sind durch meine Untersuchung ebenso Inhalte identifizierbar, die als kontinuierlich-traditionelle Inhalte des Sachunterrichts bezeichnet werden können und z. T. auch schon in der Heimatkunde berücksichtigt wurden. Dieses gilt u.a. für einzelne Teilbereiche zweier traditioneller Inhaltsbereiche der Heimatkunde (biologisch orientierte Inhalte und geografisch orientierte Inhalte), die beide in den 90er Jahren die ersten beiden Rangplätze mit

18% bzw. 15% der Inhalte belegen und zusammen ca. 1/3 der Inhalte des Sachunterrichts einnehmen.

Die Untersuchung der Inhalte des Sachunterrichts über den genannten Zeitraum zeigt, dass der Sachunterricht aus einem Konglomerat von bezugsfachorientierten Inhalten und Inhalten aus fächerübergreifenden Bereichen besteht. Dabei belegt die Untersuchung empirisch, was bereits in der Literatur angemerkt wurde (u.a. Fölling-Albers 1993, S. 10; Köhnlein 1998, S. 30, Daum 2000, S. 55): Es besteht der aktuelle Trend, sachunterrichtsübergreifende Inhaltsbereiche verstärkt im Sachunterricht zu berücksichtigen. Damit ist der Sachunterricht zu einem Sammel- und Auffangbereich der fächerübergreifenden Inhalte in der Grundschule geworden. Dieses hat zur Folge, dass sich der Inhaltsrahmen der Didaktik des Sachunterrichts erweitert hat. Die Verteilung der Inhalte des Sachunterrichts zeigt, dass sie heute aus ca. 2/3 der Inhalte der Bezugsfächer bestehen und ca. 1/3 aus den fächerübergreifenden Erziehungsbereichen. Dabei sind im Sachunterricht diese beiden Ebenen miteinander verzahnt, so dass für den aktuellen Sachunterricht keine klare „2/3 zu 1/3-Formel“ aufgestellt werden kann. Der Sachunterricht besteht nicht aus einer Addition dieser beiden Inhaltsgruppen, bei denen beide Summanden klar zu benennen wären und damit einigermaßen deutliche Strukturen für den Sachunterricht beschreibbar sind. Bei der Erfassung des Datenmaterials wurde deutlich, dass umfangreiche und vorerst noch nicht vollständig aufzeigbare Verbindungen untereinander existieren, so dass diese Anteile im konkreten Sachunterricht nur schwer in isolierter Form zu erkennen sind, wie diese inhaltsanalytische Untersuchung es zunächst offenzulegen scheint. Bezogen auf Diskussionen in der Didaktik des Sachunterrichts zeigt sich, dass sich keiner der beiden aufgestellten Pole (Fachorientierung auf der einen Seite und fächerübergreifende Orientierung auf der anderen Seite) in reiner Form durchgesetzt hat, sondern ein „Sowohl-als-auch“ im konkreten Sachunterricht vorhanden ist.

Die Ergebnisse verweisen darauf, dass das inhaltliche Profil des Sachunterrichts insgesamt als nur schwer überschaubar bezeichnet werden kann. Welche Wege könnten beschritten werden, um überschaubare inhaltliche Strukturen für den Sachunterricht zu schaffen? Dieses kann nur durch einen klaren inhaltlichen Rahmen erreicht werden. Dieser Inhaltsrahmen muss aber flexibel auf aktuelle Zeitveränderungen reagieren können, so dass ein Konzept von notwendiger Offenheit, aber auch Geschlossenheit aufzustellen ist.

## Anmerkungen

- <sup>1</sup> Dieser Beitrag stellt einen Ausschnitt einer umfangreicher angelegten Forschungsarbeit dar. Nicht berücksichtigt werden können an dieser Stelle die detaillierten Ergebnisse der Tendenzen, die sich innerhalb der 12 Inhaltsbereiche für den Untersuchungszeitraum abzeichnen. Zusammengefügt ergeben diese Einzelanalysen der Teilbereiche ein konkretes Bild der inhaltlichen

Entwicklungslinien des gesamten Lernbereichs. Die statistische Auswertung wird dabei ergänzt durch eine Beschreibung der konkreten inhaltlichen Entwicklungen auf der Grundlage des Datenmaterials.

- <sup>2</sup> Der Sachunterricht in der Schulpraxis ist nach meinen Beobachtungen weit davon entfernt, eine einheitliche Berücksichtigung der Inhalte vorweisen zu können. Jede Lehrkraft setzt individuelle Prämissen bei der Inhaltsauswahl und gelangt zu anderen Inhaltsentscheidungen. Die Inhalte des Sachunterrichts sind dabei eingebettet in die regionale und soziale Wirklichkeit vor Ort und unterliegen somit verschiedenen Einflußfaktoren. Ein streng umrissener, verbindlicher Rahmen ist daher problematischer zu konstruieren als in anderen Unterrichtsfächern der Grundschule (beispielsweise Mathematik oder Deutsch).

## Literatur

- Bolscho, D. (1978). *Lehrpläne zum Sachunterricht*. Köln: Aulis Verlag Deubner.
- Breitschuh, G. (1997). *Inhalte des Sachunterrichts im 4. Schuljahr*. Kiel: GDSU-Tagung.
- Daum, E. (2000). Die Fächer lassen einen im Stich - Plädoyer für mehr Wirklichkeitsbewusstsein im Sachunterricht. In G. Löffler, V. Möhle, D. von Reeken & V. Schwier (Hrsg.), *Sachunterricht - Zwischen Fachbezug und Integration* (S. 50-62). Bad Heilbrunn: Klinkhardt.
- Einsiedler, W. & Schirmer, G. (1986). Sachunterrichtsreform und Unterrichtsgestaltung. Eine Analyse von Schülerarbeitsmappen 1968-1981. *Die Deutsche Schule* 78, H.3, 1986, S. 316-326.
- Einsiedler, W. (1998). *The Curricula of Elementary Science Education in Germany*. Erlangen-Nürnberg: IfG (=Berichte und Arbeiten aus dem Institut für Grundschulforschung; Nr. 88).
- Feige, B. (2000). Integrativer und fächerübergreifender Sachunterricht – historische, fachdidaktische und allgemeindidaktische Orientierungen. In G. Löffler, V. Möhle, D. von Reeken & V. Schwier (Hrsg.), *Sachunterricht – Zwischen Fachbezug und Integration* (S. 63-79). Bad Heilbrunn: Klinkhardt.
- Fölling-Albers, M. (1993). Der Sachunterricht in der Grundschule - Auf der Suche nach einem Profil. In D. Richter (Hrsg.). *Grundlagen des Sachunterrichts* (S. 9-19). Oldenburg: Zentrum für Pädagogische Berufspraxis.
- Haarmann, D. & Kallb, P. E. (Hrsg.). (1999). *Grundschule 2000*. Weinheim & Basel: Beltz.
- Haug, J. (1969). *Heimatkunde und Volkskunde*. Tübingen: Tübinger Vereinigung für Volkskunde E.V. Tübinger Schloss.
- Höcker, G. (1968). Inhalte des Sachunterrichts im 4. Schuljahr. *Die Grundschule*, 1, H.3, S. 10-14.
- Kaiser, A. (1997). *Einführung in die Didaktik des Sachunterrichts*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Katzenberger, L. F. (1972). *Der Sachunterricht der Grundschule in Theorie und Praxis Teil 1*. Ansbach: Prögel.
- Klafki, W. (1992). Allgemeinbildung in der Grundschule und der Bildungsauftrag des Sachunterrichts. In R. Lauterbach, W. Köhnlein, K. Spreckelsen & E. Klewitz (Hrsg.), *Brennpunkte des Sachunterrichts* (S. 11-31). Kiel: IPN.
- Köhnlein, W. (1998). Grundlegende Bildung - Gestaltung und Ertrag des Sachunterrichts. In B. Marquardt-Mau, H. Schreier (Hrsg.), *Grundlegende Bildung im Sachunterricht* (S. 27-46). Bad Heilbrunn: Klinkhardt.
- Marquardt-Mau, B. (1998). Grundlegende Bildung im Sachunterricht. In B. Marquardt-Mau, H. Schreier (Hrsg.), *Grundlegende Bildung im Sachunterricht* (S. 7-15). Bad Heilbrunn: Klinkhardt.
- Meiers, K. (1994). *Sachunterricht*. Zug/CH: Klett und Balmer & Co.
- Merten, K. (1995). *Inhaltsanalyse*. Opladen: Westdeutscher Verlag.

- Mitzlaff, H. (1985). *Heimatkunde und Sachunterricht*. Dortmund: Diss.
- Löffler, G., Möhle, V., von Reeken, D. & Schwier, V. (2000). *Sachunterricht – Zwischen Fachbezug und Integration*. Bad Heilbrunn: Klinkhardt.
- Schreier, H. (1979). *Sachunterricht – Themen und Tendenzen*. Paderborn: Schöningh.
- Schreier, H. (2000). Einführung. In G. Löffler, V. Möhle, D. von Reeken & V. Schwier (Hrsg.), *Sachunterricht – Zwischen Fachbezug und Integration* (S. 13-19). Bad Heilbrunn: Klinkhardt.
- Strunck, U., Lück, G. & Demuth, R. (1998). Der naturwissenschaftliche Sachunterricht in Lehrplänen, Unterrichtsmaterialien und Schulpraxis. *Zeitschrift für Didaktik der Naturwissenschaften* 4, H. 1, S. 69-80.

# Autorenspiegel

Beate Blaseio

Grundschullehrerin in Husum und Lehrbeauftragte am Institut für Biologie und Sachunterricht und ihre Didaktik an der Universität Flensburg

Professor Dr. Wolfgang Einsiedler

Friedrich-Alexander-Universität Erlangen, Institut für Grundschulforschung, Lehrstuhl für Grundschulpädagogik und -didaktik

Dipl.-Soz. Achim Engelen

Westfälische Wilhelms-Universität Münster, Seminar für Didaktik des Sachunterrichts

Professor Dr. Richard Gunstone

Monash University Centre, Faculty of Education, Clayton, Australia

Professor Dr. Hartmut Giest

Universität Potsdam, Institut für Grundschulpädagogik

Dr. Ilonca Hardy

Max-Planck-Institut für Bildungsforschung Berlin, Forschungsbereich Erziehungswissenschaft und Bildungssysteme

Dr. Andreas Hartinger

Universität Regensburg, Philosophische Fakultät II, Institut für Pädagogik, Lehrstuhl für Grundschulpädagogik und -didaktik

Angela Jonen

Westfälische Wilhelms-Universität Münster, Seminar für Didaktik des Sachunterrichts

Dr. Susanne Körber

Ludwig-Maximilians-Universität München, Fakultät für Psychologie und Pädagogik

Dr. Hermann Laux

Universität Koblenz-Landau, Institut für Grundschulpädagogik

Professor Dr. Philipp Mayring

Pädagogische Hochschule Ludwigsburg, Institut für Pädagogische Psychologie und Soziologie

Professorin Dr. Kornelia Möller

Westfälische Wilhelms-Universität Münster, Seminar für Didaktik des Sachunterrichts

Professor Dr. Hanns Petillon

Universität Koblenz-Landau, Institut für Grundschulpädagogik

Professor Dr. Hans-Günter Roßbach

Universität Lüneburg, Institut für Schul- und Hochschulforschung

Professor Dr. Jürgen Rost  
Christian-Albrechts-Universität Kiel, Institut für die Pädagogik der Naturwissenschaften (IPN)

Professorin Dr. Beate Sodian  
Ludwig-Maximilians-Universität München, Lehrstuhl für Entwicklungspsychologie

Professor Dr. Kay Spreckelsen  
Gesamthochschule Universität Kassel, Fachbereich Physik, Arbeitsgruppe Didaktik der Physik

Professorin Dr. Elsbeth Stern  
Max-Planck-Institut für Bildungsforschung Berlin, Forschungsbereich Erziehungswissenschaft und Bildungssysteme

Dipl.-Psych. Claudia Thoermer  
Ludwig-Maximilians-Universität München, Lehrstuhl für Entwicklungspsychologie

Professor D. Richard White  
Monash University Centre London



Wissenschaftliche Aussagen bedürfen der Verlässlichkeit der Wege, auf denen sie gewonnen wurden. So stehen Methoden der wissenschaftlichen Erkenntnisgewinnung in der Didaktik des Sachunterrichts im Zentrum des vorliegenden Bandes.

Der spezifische Beitrag dieses Bandes zur Diskussion von Forschungsmethoden besteht dabei in der Darstellung sachunterrichtsdidaktisch relevanter empirischer Forschung.

Dies geschieht in drei Teilen: Im Anschluß an eine Bestandsaufnahme über die aktuelle Forschungslage im Sachunterricht werden unterschiedliche empirische Ansätze und ihr Verhältnis zueinander thematisiert. Der dritte Teil schließlich beinhaltet inhaltliche Beiträge; hier werden aktuelle empirische Arbeiten zum Sachunterricht dargestellt.

Die Beiträge dieses Bandes sind aber nicht nur für Forschergruppen interessant, sondern auch für Lehrkräfte, da sich fast alle dargestellten Verfahren nicht nur für wissenschaftliche Studien, sondern zugleich auch für informelle schulische Evaluationen eignen und darüber hinaus durchaus auch Anstöße für die Gestaltung des eigenen Unterrichts geben können.

**KLINKHARDT**

3-7815-1187-1



9 783781 511873