Zuerst ersch. in: Bildungsqualität von Schule : Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen / hrsg. von Manfred Prenzel ... (Hrsg.)(Zeitschrift für Pädagogik, Beiheft ; 45). Weinheim : Beltz, 2002, S. 83-100

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

Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Konstruktion des Regensburger Mathematikleistungstests für 5. bis 10. Klassen[1]

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

1. Zielsetzung und Konzeption der Längsschnittstudie PALMA

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

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

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

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

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

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

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

2. Mathematische Grundbildung und die Ziele des Regensburger Mathematikleistungstests

PISA folgt einer Konzeption mathematischer Grundbildung, die sich im Sinne von „mathematical literacy" vor allem auf Kompetenzen zu einem realitätsorientierten, flexibelmodellierenden Umgang mit mathematischen Problemen konzentriert. Für solche Kompetenzen sind aus mathematikdidaktischer Perspektive problemlösebezogene mentale Repräsentationen mathematischer Inhalte zentral, die häufig als mathematische „Grundvorstellungen“ bezeichnet werden. Es mangelt bisher an Instrumenten zur Erfassung dieser Kompetenzen und an Untersuchungen zu ihrer Entwicklung im Laufe der Schulzeit.

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

2.1 Mathematische Grundbildung

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

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

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

2.2 Der Prozess des Modellierens

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

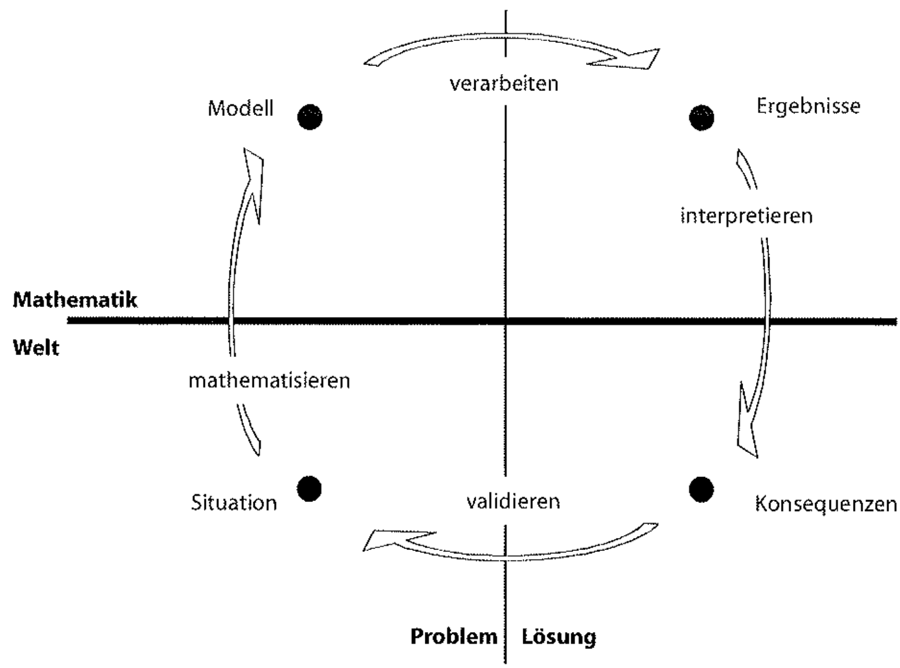

Abb. 1: ***Prozess der Modellierung***

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

Zentrale kognitive Tätigkeiten beziehen sich hierbei auf das *Übersetzen zwischen Realität und Mathematik*, wenn beispielsweise zu einer Sachsituation eine angemessene Mathematisierung gefunden werden muss oder wenn ein mathematisches Ergebnis wieder im Hinblick auf die Sachsituation interpretiert werden soll. Hierfür braucht man Vorstellungen davon, welche mathematischen Inhalte oder Verfahren zu einer bestimmten Sachsituation passen könnten bzw. umgekehrt, welche Situationen sich mit bestimmten mathematischen Inhalten modellieren lassen.

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

2.3 Mathematische Grundvorstellungen

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

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

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

2.4 Ziele der Konstruktion des Regensburger Mathematikleistungstests

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

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

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

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

3. Methode

3.1 Methodische Konzeption der Längsschnittstudie

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

Untersuchungsanlage und Stichproben

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

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

Variablen und Instrumente

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4. Ergebnisse und Diskussion

4.1 Raschskalierung des Mathematikleistungstests

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

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

Tab. 1: ***Mittlere Fähigkeit $\overline{\theta_v}$ der Probanden nach Klassenstufe***				
Jahrgangsstufe	$\overline{\theta_v}$	N	s	D(−1)[s]
5. Klasse	−1.48	176	1.25	
6. Klasse	−.77	103	.81	.56
7. Klasse	−.29	140	.89	.59
8. Klasse	.26	150	1.15	.63
9. Klasse	.99	109	.93	.63
10. Klasse	1.47	102	1.09	.51
Anmerkung: D(−1)[s]: Fähigkeitsdifferenz der Schulformen in Einheiten der Standardabweichung gegenüber dem Vorjahr.				

Tab. 2: ***Mittlere Fähigkeit $\overline{\theta_v}$ der Probanden nach Schulform***				
Schulform	$\overline{\theta_{rv}}$	N	s	D(−1)[s]
Hauptschule	−.87	204	1.32	
Realschule	−.14	270	1.48	.49
Gymnasium	.43	306	1.28	.29
Anmerkung: D(−1)[s]: Fähigkeitsdifferenz der Schulformen in Einheiten der Standardabweichung gegenüber dem Vorjahr				

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

4.2 Itemanalysen und Testrevision

Eine genauere Analyse der Schwierigkeitsstaffelung innerhalb der Testversionen zeigt, dass die vorläufigen Testversionen eine differenzierte Erfassung des Fähigkeitsspektrums der unteren Klassenstufen erlauben, während die Erfassung in den höheren Klassen- und Leistungsstufen vergleichsweise weniger differenziert ausfällt. Hieraus ergeben sich Konsequenzen für die weitere Testentwicklung: Die Differenzierung der vorläufigen Testversionen scheint noch nicht optimal zu sein. Dies gilt insbesondere für die Testversionen für die 6.–10. Klassenstufe, da diese fünf Klassenstufen ein erhebliches Leistungsspektrum umfassen. Allerdings scheint es nicht notwendig zu sein, für jede Klassenstufe getrennte Testversionen zu entwickeln, da sich die Leistungsbandbreiten aufeinander folgender Klassenstufen stark überschneiden (siehe Tabelle 1); zudem würde dies die Analyse von Entwicklungsverläufen eher erschweren.

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

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

4.3 Subskalen und differenziertere Analysen

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

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

4.4 Entwicklungsverläufe und didaktische Handlungsperspektiven

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

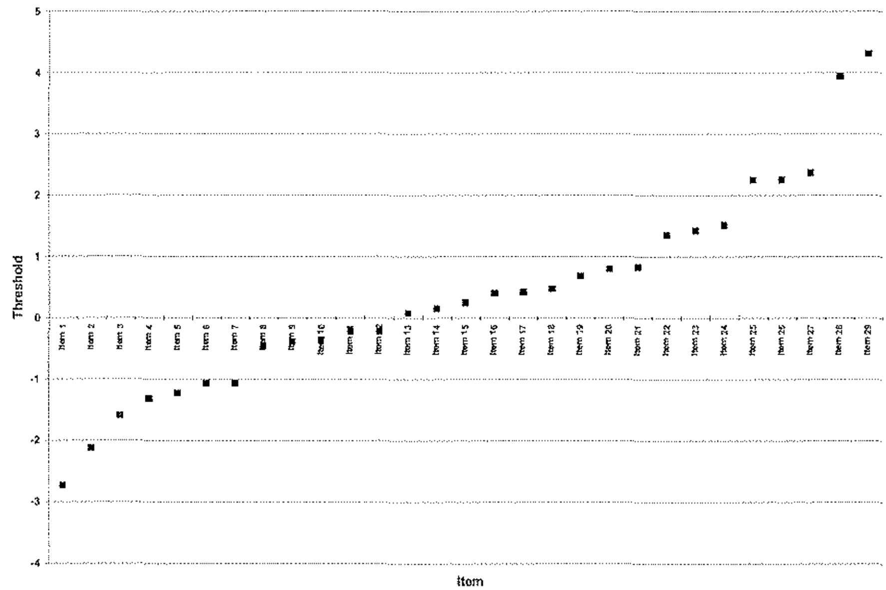

Abb. 2: ***Subskala Prozentrechnung***

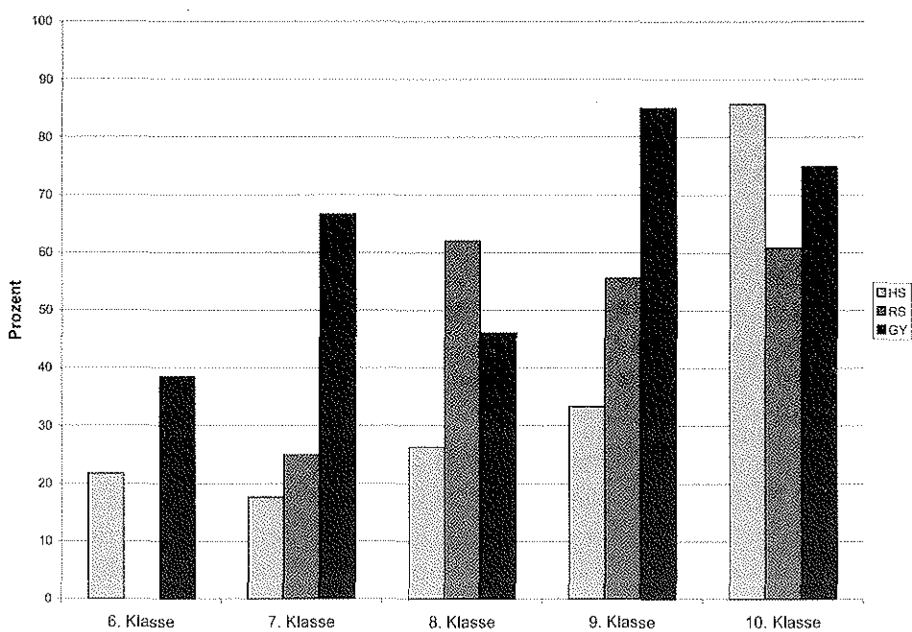

Abb. 3: *Relative Lösungshäufigkeiten bei der Aufgabe „Fahrradkontrolle"*

(1) Stagnierende Verläufe beim Verständnis des Anteilsbegriffs

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

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

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

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

(2) Einbruchseffekte beim Verständnis von Proportionalität

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

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

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

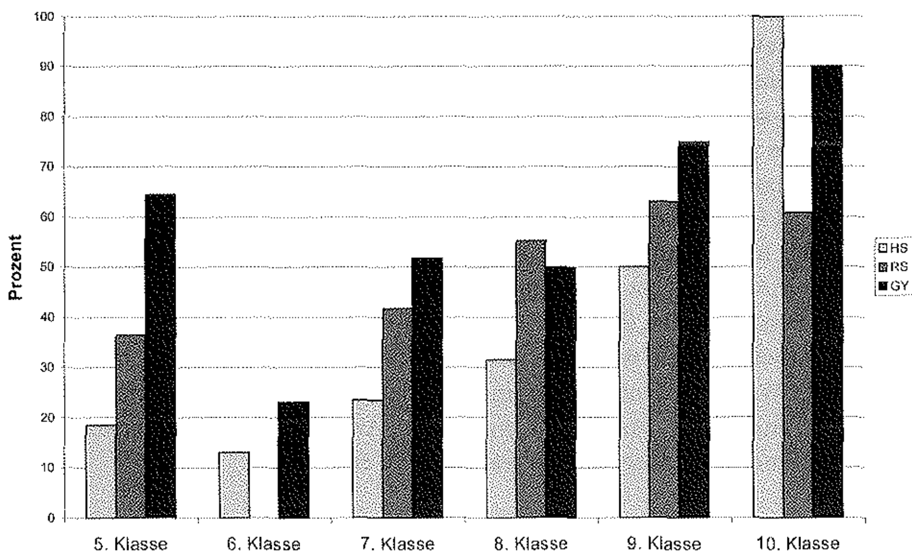

Abb. 4: ***Relative Lösungshäufigkeiten bei der Aufgabe „Lotto"***

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

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

4.5 Schlussfolgerungen

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

nahme von Sonderschulen). Die Kompetenzzunahme von einer Klassenstufe zur nächsten entspricht etwa einer halben Standardabweichung der Kompetenzvariation innerhalb der Klassenstufen; dies spricht für die Validität des Instruments.
Dabei ergeben sich überraschende Befunde, wenn man Kompetenzverläufe unterhalb der Ebene eines globalen Summenscores für die Mathematikleistung analysiert. In einzelnen curricularen Bereichen scheinen Phasen mit positivem Lernzuwachs solchen mit stagnierender Leistungsentwicklung oder sogar negativem Lernzuwachs gegenüber zu stehen. Bei der Interpretation ist allerdings zu berücksichtigen, dass die hier berichteten Befunde dieser Art zunächst querschnittlich angelegt sind und auf relativ kleinen Stichproben basieren. Zentrales Ziel der Längsschnittstudie wird es sein, sowohl gelungene Kompetenzentwicklungen wie auch problematische Leistungsverläufe in methodisch angemessener Weise zu identifizieren und ihre Bedingungen in Unterricht und sozialen Kontexten zu analysieren, um die sich ergebenden Ansatzpunkte für unterrichtliche und didaktische Handlungskonsequenzen realistisch zu erhärten.

Literatur

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

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

Schubring, G. (1978): Das genetische Prinzip in der Mathematikdidaktik. Stuttgart: Klett-Cotta.
Schupp, H. (1988): Anwendungsorientierter Mathematikunterricht in der Sekundarstufe zwischen Theorie und neuen Impulsen. In: Der Mathematikunterricht, 34 (6), S. 5–16.

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

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