

Reinhard Pekrun, Rudolf vom Hofe, Werner Blum, Thomas Götz,
Sebastian Wartha, Anne Frenzel & Simone Jullien

Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA)

Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von Mathematikleistungen in der Sekundarstufe I

1 Ausgangsfragen und Zielsetzung des Projekts

National und international mangelt es an Längsschnittstudien zur Entwicklung von mathematischen Kompetenzen im Altersbereich der Sekundarstufe I. Die wenigen vorliegenden Studien sind meist auf einzelne Kompetenzbereiche, Jahrgangsstufen und Schülergruppen oder auf qualitative Analysen beschränkt, so dass verallgemeinerbare Schlussfolgerungen bisher kaum ableitbar sind (Baumert, et al., 1997; Blum et al., 1994; Köller et al., 2003; Kuechemann & Hoyles, 2003; Lehmann, Husfeldt & Peek, 2001; Rost, 2000; Watson & Kelley, 2004). Dies gilt nicht nur für die durchschnittlichen Entwicklungsverläufe von mathematischen Kompetenzen in diesem Alterszeitraum, sondern auch für die Entwicklung von Leistungsdefiziten, Geschlechterunterschieden und sozialen Disparitäten im Fach Mathematik und für die Schüler- und Kontextbedingungen der Kompetenzentwicklung. Wenig ist dementsprechend auch zu der Frage bekannt, wie die Variation von Kompetenzen und Einstellungen im Fach Mathematik zustande kommt, die sich in Vergleichsuntersuchungen wie den OECD-Erhebungen der PISA-Zyklen zeigt (Baumert et al., 2001; Blum, Neubrand et al., 2004). Für die Gestaltung von Bildungsstandards, Lehrplänen, Unterricht und Evaluierungsinstrumenten ist dieser Kenntnismangel gleichermaßen problematisch.

Ziel des Projekts ist es auf diesem Hintergrund, anhand einer Längsschnittuntersuchung die Entwicklung einer größeren Bandbreite mathematischer Kompetenzen über den gesamten Alterszeitraum der Sekundarstufe I hinweg zu verfolgen. Dabei sollen auch die Individual-, Unterrichts- und Elternhausbedingungen dieser Entwicklung analysiert werden. Dies soll gleichzeitig zu einer Aufklärung der Entwicklungsverläufe beitragen, die den Leistungen zugrunde liegen, die sich in den PISA-Erhebungen bei den 15-Jährigen bzw. den Neuntklässlern in Deutschland zeigt. Im Bereich der Schülervoraussetzungen liegt ein besonderes Augenmerk auf den – mit Ausnahme von Mathematikangst – bisher kaum untersuchten Emotionen von Schülerinnen und Schülern im Fach Mathematik. Ziel ist ferner, die gewonnenen Erkenntnisse für die Entwicklung von Produkten für die pädagogische Praxis zu nutzen (in den Bereichen Unterrichtsmaterialien, Lehrplan- und Curriculumentwicklung, Lehrerbildung, Leistungs- und Emotionsdiagnostik sowie Unterrichtsvaluation).

Die Forschungskonzeption des Projekts (Pekrun, Götz, vom Hofe et al., 2004) orientiert sich an einem Modell mathematischer Kompetenzen, das neben curricular orientierten Inhaltsbereichen (Arithmetik, Algebra, Geometrie) u.a. zwischen Modellierungskompetenzen einerseits und algorithmisch orientierten Kalkülkompetenzen andererseits differenziert. Dieses Kompetenzmodell entspricht im Wesentlichen den in den PISA-Erhebungen verwendeten Konzepten einer *mathematical literacy* und trägt der Rolle

mathematischer Grundvorstellungen für das Lösen mathematischer Probleme Rechnung (Blum & vom Hofe, 2003; Blum, vom Hofe et al., 2004; vom Hofe, 1995; vom Hofe et al., 2005). Unter Kompetenzen zu mathematischer Modellierung sind dabei Fähigkeiten zu verstehen, reale Sachverhalte in mathematische Modelle zu übersetzen, das jeweilige Problem innerhalb des Modells einer mathematischen Lösung zuzuführen und diese Lösung anschließend wieder in die Realität zu transponieren (Blum, Neubrand et al., 2004; ein einfaches Beispiel ist die Aufgabe „Schokolade“, s. Abschnitt 4.1.3).

In den Bereichen Schülermerkmale und Kontextbedingungen orientiert sich das Projekt an unterrichtstheoretischen und sozialkognitiven Modellvorstellungen, für den Bereich der Mathematikemotionen an der Kontroll-Wert-Theorie zu den Entwicklungsbedingungen und Leistungsfolgen von Lern- und Leistungsemotionen (Pekrun, 2000; in press; Pekrun, Götz, Titz et al., 2002). In dieser Theorie wird angenommen, dass Einschätzungen der Kontrollierbarkeit und Bedeutsamkeit von Lernen und Leistung zentral für die Entstehung von Lern- und Leistungsemotionen sind. Freude an Mathematik entsteht bspw. dann, wenn sich ein Schüler bzw. eine Schülerin hinreichend kompetent fühlt, den jeweiligen Stoff in Mathematik zu beherrschen (Einschätzung von Kontrolle), und gleichzeitig an diesem Stoff interessiert ist (Einschätzung des Wertes). Angst vor Misserfolg in Mathematik beruht der Theorie zufolge auf wahrgenommener mangelnder Kontrollierbarkeit von Misserfolg einerseits und seiner subjektiver Bedeutung andererseits. Ferner wird angenommen, dass (a) Unterrichts- und Kontextfaktoren, die Einfluss auf die Kontroll- und Werteinschätzungen von Schülerinnen und Schülern nehmen, auch für die Entwicklung ihrer Emotionen bedeutsam sind, (b) diese Emotionen sich ihrerseits positiv oder negativ auf Lernen und Leistung auswirken und (c) Individualbedingungen, Unterrichts- und Kontextbedingungen, Emotionen und ihre Leistungsfolgen in wechselseitigen Bedingungsbeziehungen über die Zeit hinweg stehen können (Einzelheiten in Pekrun, in press).

In zwei Voruntersuchungen des Projekts (2000–2002) wurden zunächst die Erhebungsinstrumente für die Längsschnittstudie entwickelt, überprüft und sukzessive optimiert. Im Jahr 2002 wurde der Längsschnitt in der fünften Jahrgangsstufe mit einer repräsentativen Stichprobe von bayerischen Schülerinnen und Schülern sowie deren Eltern und Mathematiklehrkräften implementiert und anschließend in jährlichen Erhebungen fortgesetzt. Bisher liegen Befunde für die ersten vier Messzeitpunkte vor (Jahrgangsstufen 5 bis 8). Auf der Basis der Befunde wurde ferner mit der Entwicklung von Produkten für die pädagogische Praxis begonnen. In der gegenwärtig laufenden, über den Zeitraum des Schwerpunktprogramms hinausreichenden Projektphase wird die Längsschnittstudie fortgesetzt (Jahrgangsstufen 9 und 10), und die Produkte für die Praxis werden weiterentwickelt und ergänzt.

In diesem Beitrag wird zunächst kurz auf die Erhebungsinstrumente eingegangen, die in diesem Projekt entwickelt worden sind, u.a. in der Längsschnittuntersuchung des Projekts zum Einsatz kommen und nach Abschluss des Längsschnitts auch anderen Anwendern zur Verfügung stehen werden. Anschließend wird die Methodik des Längsschnitts dargestellt und zentrale Befunde aus den bereits vorliegenden Erhebungen der Studie werden diskutiert. Da die Erhebungen noch nicht abgeschlossen sind, handelt es sich um einen Zwischenbericht, dem weitere Publikationen folgen werden. Die berichteten Befunde beziehen sich auf die Entwicklung von Kompetenzen, Emotionen und Schülermerkmalen in Mathematik, den Unterricht in diesem Fach und die Rolle des Elternhauses für die Entwicklung in Mathematik. Abschließend wird kurz auf die Materialien für die pädagogische Praxis eingegangen, die in diesem Projekt entwickelt werden.

2 Entwicklung von Erhebungsinstrumenten

2.1 Regensburger Mathematikleistungstest für 5. bis 10. Klassen

Zur Erfassung mathematischer Kompetenzen in der Sekundarstufe I wurde der Regensburger Mathematikleistungstest für 5. bis 10. Klassen entwickelt (vom Hofe et al., 2002; vom Hofe et al., 2005). Im Sinne der Forschungskonzeption des Projekts erfasst dieser raschskalierte Test anhand entsprechender Subskalen Modellierungskompetenzen sowie algorithmisch geprägte Kalkülkompetenzen in den Inhaltsbereichen Arithmetik, Algebra und Geometrie. Darüber hinaus können mit den Testitems auch für spezifischere Kompetenzbereiche Skalen gebildet werden (z.B. zur Bruchrechnung).

Die einzelnen Versionen des Tests sind so konstruiert und über die Jahrgangsstufen hinweg verknüpft, dass die gesamte über den Zeitraum der Sekundarstufe I hinweg anzutreffende Leistungsvariation im Fach Mathematik abgebildet werden kann. In den Erhebungen des Längsschnitts wird hierzu ein Multi-Matrix-Sampling von Testitems praktiziert. Innerhalb jeden Messzeitpunkts werden zwei zueinander parallele, jahrgangsstufenspezifische Testversionen verwendet, die über gemeinsame Ankeritems miteinander verbunden sind. Ferner findet eine Verankerung zwischen den Testversionen benachbarter Jahrgangsstufen statt. Die Testzeit betrug zum ersten Messzeitpunkt 60 Minuten (63 Testitems), zu den nachfolgenden Zeitpunkten jeweils 90 Minuten (89/117/158 Items zum zweiten, dritten und vierten Messzeitpunkt).

Analysiert man die Qualität der Skalierung anhand disjunkter Zufallsstichproben von Items (Anderson, 1973), so zeigt sich eine hohe Übereinstimmung der Itemparameter nach unabhängiger Schätzung in den Teilstichproben (s. Abbildung 1; exemplarisch dargestellt für den dritten Messzeitpunkt). Dieses Ergebnis spricht für eine gute Passung zwischen Modell und Daten. Auch der *Likelihood-Ratio-Test* doku-

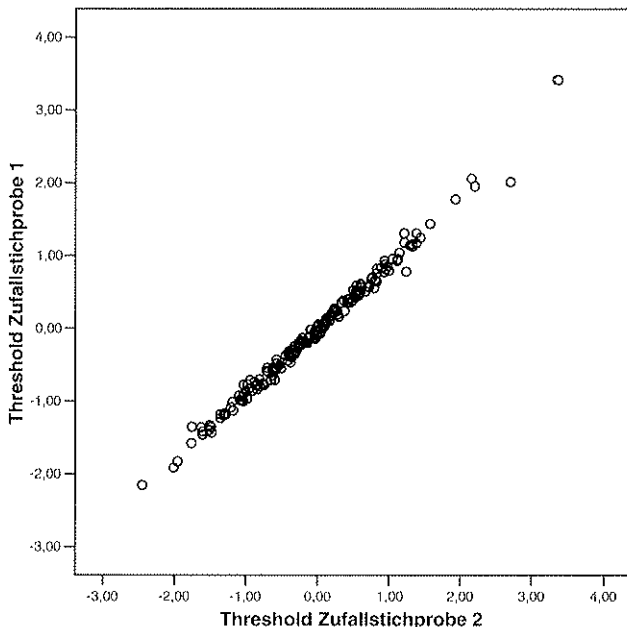


Abb. 1: Graphischer Modellgeltungstest nach Anderson für MZP 3

mentiert günstige Skalenskennwerte des Gesamttests ($\chi^2(116) = 91,92$, n.s.). Ähnlich gute Werte zeigen sich für die Subskalen (z.B. im Bereich der Algebra: Subskala „Modellierungskompetenzen“: $\chi^2(43) = 38,51$, n.s.; Subskala „Kalkülkompetenzen“: $\chi^2(20) = 14,54$, n.s.).

Neben der Quantifizierung der Daten nach dem Rasch-Modell wurden *latent-class*-Analysen durchgeführt, um die Validität der Skalen an inhaltlichen Kriterien zu überprüfen (Kleine, 2004). Dabei werden Klassen von Probanden mit ähnlichen Antwortprofilen im Test gebildet. Die Qualität der Analysen konnte anhand des *Bootstrap*-Verfahrens bestätigt werden (Langeheine & von Davier, 1996). Die Befunde dokumentieren, dass die Anordnung der Personenklassen den Anforderungsniveaus des Tests entspricht, die durch die jeweils erforderlichen Grundvorstellungen bestimmt sind (Kleine, 2004). Die Analysen liefern damit eine Bestätigung der didaktischen, an Kompetenz- und Grundvorstellungskonstrukten orientierten Konzeption des Tests.

2.2 Skalen zu Schülermerkmalen, Unterricht und Kontextbedingungen

Das in diesem Projekt eingesetzte Instrumentarium zu Schülermerkmalen, Unterrichtsvariablen und Kontextbedingungen umfasst einen Satz von neu entwickelten bzw. für das Projekt adaptierten Skalen zu den folgenden Bereichen (vgl. Pekrun, Götz, Jullien et al., 2004):

- 1) **Kognitive Grundfähigkeiten** (fluide Intelligenz: KFT-Skalen zu verbalem und nonverbalem reasoning; Heller & Perleth, 2000);
- 2) **Schülermerkmale** (schulischer Bildungsweg sowie Selbst- und Valenzüberzeugungen, Emotionen, Interesse, Motivation, Lernstrategien und Selbst- vs. Fremdregulation im Fach Mathematik; 30 Selbstberichtsskalen mit 178 Items; u.a. die Münchener Skalen zu Mathematikemotionen (englisch unter dem Titel *Achievement Emotions Questionnaire – Mathematics*, AEQ-M; Pekrun, Götz & Frenzel, 2005);
- 3) **Prozessvariablen der Aufgabenbearbeitung im Mathematiktest** (u.a. Anstrengungsbereitschaft);
- 4) **Mathematikunterricht** (Schülerfragebogen: 15 Skalen, 60 Items; Lehrerfragebogen: 16 Skalen, 81 Items);
- 5) **Elternhaus** (sozioökonomischer Status, kulturelles Kapital, Umgang des Elternhauses mit Mathematik und Leistungsanforderungen der Schule, Kooperation von Elternhaus und Schule; Schülerfragebogen mit 10 Skalen und 40 Items, Elternfragebogen mit 7 Skalen und 31 Items).

Die Verteilungs- und Reliabilitätseigenschaften dieser Skalen erwiesen sich in allen vorliegenden Erhebungen des Längsschnitts als gut (Pekrun, vom Hofe & Blum, 2003, 2006). Der gesamte Skalensatz und die Resultate von Skalen- und Itemanalysen sind in den Skalenhandbüchern zum Längsschnitt dokumentiert (Pekrun, Götz, Jullien et al., 2002, 2003, 2004; Pekrun, Jullien, Lichtenfeld et al., 2005).

2.3 Skalen zu Modellierungs- und Kalkülemotionen

In der letzten Projektphase neu entwickelt wurden Skalen zur Erfassung von Freude, Angst und Langeweile, die Schülerinnen und Schüler bei modellierungsorientierten Aufgaben in Mathematik einerseits und algorithmisch orientierten Kalkülaufgaben andererseits erleben („Modellierungseemotionen“ vs. „Kalkülemotionen“). Ziel dieser Entwicklung war es, Konzeptualisierungen zu Kompetenzen und Emotionen – über die domänenbezogene Konzeption von Emotionen hinaus – auch auf der Ebene von Teilkompetenzen miteinander zu verschränken. Ausgangspunkt war die Alltagserfahrung von Lehrkräften, dass Mathematikaufgaben unterschiedlichen Typs bei Schülerinnen und Schülern unterschiedlich beliebt zu sein scheinen. Vor allem Modellierungsaufgaben (z.B. Textaufgaben) stehen im Ruf, wenig beliebt zu sein.

Ähnlich wie bei Skalen zur Erfassung von aufgabenbezogenen Selbstwirksamkeitserwartungen (Pajares & Miller, 1995) verwenden die Skalen ein aufgabenbezogenes Format. Es werden Modellierungsaufgaben (Aufgaben mit Einbettung in einen Sachkontext) und formal zu lösende Kalkülaufgaben (ohne Einbettung) dargeboten. Für jede der Aufgaben ist anhand von 5-stufigen Likert-Skalen anzugeben, in welchem Ausmaß bei der Bearbeitung Freude, Angst und Langeweile erlebt werden. Um eine Konfundierung mit Aufgabenparametern jenseits der Differenzierung von Modellierungs- vs. Kalkülgehalt zu vermeiden, wurden für diese Skalen Modellierungs- und Kalkülaufgaben entwickelt, die inhaltlich und in ihrer Schwierigkeit parallelisiert sind.

Mit einer solchen Erfassung von Emotionen in der direkten Auseinandersetzung mit Mathematikaufgaben wird methodisches Neuland betreten. Anhand der Daten des dritten Messzeitpunkts des Längsschnitts (Jahrgangsstufe 7) wurden die sechs Skalen auf ihre psychometrische Qualität überprüft. Die Ergebnisse sprechen dafür, dass der Versuch erfolgreich war. Für alle sechs Skalen ergaben sich hinreichende Wertestreuungen und Konsistenzreliabilitäten ($.70 < \alpha < .72$ bei jeweils vier Items pro Skala). Für die Validität sprechen Korrelationen um $r = .40$ mit den Münchener Skalen zu Mathematikemotionen. Erste Befunde zu differentiellen emotionalen Reaktionen auf die beiden Aufgabentypen werden unten (Abschnitt 4.2.2) berichtet.

2.4 Bezüge der Instrumente zu den Erhebungen von PISA

Die PALMA-Skalen zu mathematischen Kompetenzen beziehen sich ebenso wie die PISA-Erhebungen auf Kompetenzkonzepte einer *mathematical literacy*, für die Modellierungskompetenzen zentral sind (Blum, Neubrand et al., 2004). Inhaltlich erfassen die Skalen des Projekts Kompetenzen in den Bereichen Arithmetik, Algebra und Geometrie, die den Bereichen „Quantität“, „Veränderung und Beziehungen“ sowie „Raum und Form“ in der PISA-Konzeption entsprechen (Blum, Neubrand et al., 2004). Auch skalentechnisch sind die Testinstrumente von PALMA und PISA eng miteinander verzahnt. Eine besonders gute Vergleichbarkeit wird in den Bereichen „Quantität“ und „Veränderung und Beziehung“ dadurch gewährleistet, dass Aufgaben aus der PALMA-Studie in den nationalen Erhebungen von PISA 2003 und in den mit PISA 2006 verknüpften Erhebungen zur Normierung der Bildungsstandards Mathematik für die Jahrgangsstufe 9 eingesetzt wurden.

Neben der Orientierung an Konzepten der *mathematical literacy* und der mathematischen Modellierung stellt auch die Fokussierung auf mathematische Grundvorstellungen eine konzeptuelle Gemeinsamkeit von PALMA und PISA dar. So konnte

in beiden Studien nachgewiesen werden, dass die „Grundvorstellungsintensität“ von Aufgaben eine der wichtigsten Erklärungsvariablen für die Itemschwierigkeit bei Modellierungsaufgaben darstellt (Blum, vom Hofe et al. 2004).

Im Bereich von Schülermerkmalen und Kontexten gibt es ebenfalls zahlreiche Überlappungen mit den PISA-Erhebungen. Items der PALMA-Skalen zu Freude, Angst, Interesse und Lernverhalten in Mathematik fanden Eingang in die internationalen Erhebungen von PISA 2003. Ferner wurden PALMA-Skalen zu Motivation und Selbstregulation in Mathematik und zum Elternhaus in den Schüler- und Elternfragebögen der nationalen Ergänzungserhebungen von PISA 2003 eingesetzt (Prenzel et al., 2004).

3 Methodik der Längsschnittuntersuchung

3.1 Stichproben und Erhebungen

Die Untersuchung verfolgt eine Stichprobe aus einer bayerischen Kohorte von Schülerinnen und Schülern sowie deren Eltern und Mathematiklehrkräfte in jährlichen Erhebungen von der 5. bis zur 10. Jahrgangsstufe. Um Unterricht und Klassenkontext analysieren zu können, wurde die Stichprobe so gezogen, dass es sich jeweils um intakte Klassen von Schülerinnen und Schülern handelte. Ziel der Stichprobenziehung war gleichzeitig auch, bzgl. soziodemographischer Kriterien wie Schulart, Region, Geschlecht und Schichtzugehörigkeit eine möglichst hohe Repräsentativität zu erreichen. Um diese Repräsentativität der Längsschnittstichprobe über die Zeit hinweg möglichst weitgehend zu erhalten und eine über die Messzeitpunkte hinweg zunehmende Positivselektion der Stichprobe zu vermeiden, wurden Klassenwiederholer zu jedem Messzeitpunkt weiter in die Erhebungen einbezogen. Ferner wurde Wert darauf gelegt, die anfänglich gezogenen intakten Klassen auch über die Zeit hinweg weiter zu verfolgen. Neu in die betreffenden Klassen aufgenommene Schülerinnen und Schüler wurden deshalb jeweils auch in die Längsschnittstichprobe einbezogen.

Abweichungen vom Prinzip der Weiterführung intakter Klassen mussten dort zugelassen werden, wo Schülerinnen und Schüler aufgrund von Klassen- bzw. Schulwechsel oder einer Neukonfiguration von Klassen die Klassenzuordnungen des ersten Messzeitpunkts verloren hatten. Soweit sich in der jeweiligen aufnehmenden Klasse eine hinreichende Zahl von Schülerinnen und Schülern der Längsschnittstichprobe wiederfand, wurde die Klasse als intakte Klasse in die Stichprobe aufgenommen, indem alle Schülerinnen und Schüler der Klasse in die Erhebungen einbezogen wurden. Dies führte insbesondere mit dem Übergang von der 6. zur 7. Jahrgangsstufe zu Ergänzungen der Stichprobe (hauptsächlich aufgrund einer Neukonfiguration von Klassen nach den Zweigwahlen in Gymnasien und Realschulen). Wenn es sich hingegen jeweils nur um einzelne Schülerinnen und Schüler handelte, wurde aus testökonomischen Gründen auf eine Einbeziehung der aufnehmenden Klassen verzichtet; die betreffenden Schülerinnen und Schüler wurden dann innerhalb jeder Schule in einer „Restgruppe“ zusammengefasst und getestet. Bei Schulwechsel wurden die Schülerinnen und Schüler in der Stichprobe belassen, wenn eine hinreichende Zahl von Schülerinnen und Schülern des Längsschnitts in die betreffende Schule gewechselt war (dies führte zur Neuaufnahme von Realschulen in die Schulenstichprobe beim Schülerwechsel von Haupt- in Realschulen nach der Jahrgangsstufe 6). Mit dieser Strategie konnte sichergestellt werden, dass sich die Mortalität der Längsschnittstich-

probe in Grenzen hielt (s.u.), Stichprobenausfälle von Messzeitpunkt zu Messzeitpunkt kompensiert werden konnten und die Schülerstichprobe auch zu den späteren Messzeitpunkten weiterhin überwiegend aus intakten Klassen bestand.

Die Erhebungen finden jeweils gegen Schuljahresende statt (Mai bis Juni des betreffenden Schuljahres). Dabei handelt es sich um Erhebungen an den einbezogenen Klassen bzw. den jeweiligen „Restgruppen“ von Schülerinnen und Schülern, den Mathematiklehrkräften dieser Klassen und den jeweiligen Schülereltern. Die Schülererhebungen umfassen jeweils einen Testtag mit einer Testzeit von insgesamt 180 Minuten. Im Sinne einer Qualitätssicherung des methodischen Vorgehens wurden Stichprobenziehung und Rekrutierung der Stichproben ebenso wie die Durchführung der Erhebungen vom *Data Processing Center* (DPC) der *International Association for the Evaluation of Educational Achievement* (IEA) vorgenommen. Das DPC zeichnet auch für die Durchführung von Vergleichsuntersuchungen wie PISA, IGLU und DESI in Deutschland verantwortlich.

3.1.1 Schüler-, Eltern- und Lehrerstichproben des ersten bis vierten Messzeitpunkts

Die realisierte Schülerstichprobe des ersten Messzeitpunkts (Jahrgangsstufe 5) umfasste $N = 2.070$ Schüler ($N = 1.043$, 1.027 Jungen und Mädchen; Durchschnittsalter zum Zeitpunkt der Erhebung: 11;8 Jahre) aus 83 Klassen in 42 Hauptschulen, Realschulen und Gymnasien (Ziehung von zwei Klassen pro Schule, mit Ausnahme einer einzügigen Hauptschule). Die realisierte Elternstichprobe umfasste $N = 1.977$ Eltern, die Lehrerstichprobe alle 83 Mathematiklehrer der einbezogenen Klassen. Zum zweiten Messzeitpunkt (Jahrgangsstufe 6) handelte es sich um $N = 2.059$ Schüler ($N = 1.029$ Jungen und 1.030 Mädchen, Altersdurchschnitt: 12;9 Jahre), 1.883 Schülereltern und 76 Mathematiklehrkräfte aus 81 Klassen der einbezogenen Schulen (Schwund von zwei Klassen durch Klassenzusammenlegungen). Zum dritten Messzeitpunkt waren die Hauptschüler z.T. auf Realschulen gewechselt, die neu in die Schulenstichprobe aufgenommen wurden. Es handelte sich zu diesem Zeitpunkt um $N = 2.397$ Schüler ($N = 1.195$ Jungen und 1.202 Mädchen; Durchschnittsalter: 13;7 Jahre), 2.062 Eltern und 71 Mathematiklehrer aus 74 intakten Klassen sowie 37 „Restgruppen“ von Schülerinnen und Schülern aus weiteren Klassen aus insgesamt 44 Schulen. Zum vierten Messzeitpunkt (Jahrgangsstufe 8) schließlich bestanden die Stichproben aus $N = 2.410$ Schülern aus denselben Klassen und Schulen ($N = 1.193$ Jungen und 1.217 Mädchen; Durchschnittsalter: 14;8 Jahre) sowie 1.938 Eltern und 73 Lehrkräften. Mit diesen Stichproben wurden hohe Ausschöpfungsquoten der geplanten Stichproben erreicht (Beteiligungsraten von mindestens 91% bei den Schülerinnen und Schülern, mindestens 80% bei den Eltern und mindestens 93% bei den Lehrkräften; Tabelle 1).

Tab. 1: Beteiligungsraten zu den ersten vier Messzeitpunkten (in Prozent)

	Schülerinnen und Schüler	Eltern	Lehrkräfte
5. Klasse	91.8	95.5	100.0
6. Klasse	95.4	91.4	93.8
7. Klasse	92.5	86.0	95.9
8. Klasse	96.0	80.4	98.6

3.1.2 Längsschnittstichprobe für den Zeitraum der Jahrgangsstufen 5 bis 8

Die Längsschnittstichprobe der Schülerinnen und Schüler, die an allen Messzeitpunkten von der Jahrgangsstufe 5 bis 8 teilgenommen haben, besteht aus $N = 1.421$ Schülern (davon 49,6% Mädchen). Dies entspricht 69 Prozent der Ausgangsstichprobe der Jahrgangsstufe 5. Die Mortalität der Längsschnittstichprobe betrug in den drei Zeitintervallen vom ersten zum vierten Messzeitpunkt 12% (5. bis 6. Jahrgangsstufe), 15% (6. bis 7. Jahrgangsstufe) und 8% (7. bis 8. Jahrgangsstufe). Die Entwicklung der Stichprobe liegt damit im Bereich der Schätzungen, die bei den Planungen für den Längsschnitt veranschlagt worden waren. Hinzu kommen die Eltern dieser Schülerinnen und Schüler sowie die Mathematiklehrkräfte der jeweiligen Jahrgangsstufen. Im Laufe der vier Jahre haben Schülerinnen und Schüler aus 47 Schulen (18 Hauptschulen, 15 Realschulen und 14 Gymnasien) teilgenommen.

3.1.3 Repräsentativität der Längsschnittstichprobe

Die Schülerstichproben für den Längsschnitt wurden vom DPC so gezogen, dass sie bezüglich Geschlecht, Region und Schulart Repräsentativitätskriterien genügen (nach entsprechender Gewichtung). Darüber hinaus sollten sie aber auch bezüglich des sozioökonomischen Hintergrunds der Schülerschaft möglichst repräsentativ sein. Tabelle 2 stellt die Schichtverteilung der Stichprobe bezogen auf die jeweilige Schulart dar. Diese Verteilungen sind weitgehend äquivalent zu den Verteilungen, die sich in den Erhebungen von PISA 2000 für Bayern gezeigt haben (Kunter, MPI für Bildungsforschung, persönliche Kommunikation, Oktober 2005; landes- oder bundesamtliche Statistiken sind zu diesen Schichtverteilungen nicht verfügbar).

Tab. 2: Schichtverteilung der Längsschnittstichprobe (in Prozent)

	Hauptschule	Realschule	Gymnasium	Gesamt
Angehörige der oberen Dienstklasse (I)	8.9	20.5	39.5	25.0
Angehörige der unteren Dienstklasse (II)	16.2	25.8	23.4	22.1
Routinedienstleistungen (III)	10.1	6.2	4.6	6.5
Selbständige, einschl. Landwirte (IVa-d)	9.5	9.9	8.1	9.1
Facharbeiter und leitende Arbeiter (V-VI)	31.6	25.8	15.2	22.8
Un- und angelernte Arbeiter, Landarbeiter (VII)	26.7	12.3	9.2	14.4

3.2 Ergänzende Interviewerhebungen

Zu jedem Messzeitpunkt wurden an Teilstichproben von Schülerinnen und Schülern qualitative Zusatzerhebungen in Gestalt von halbstandardisierten Interviews durchgeführt (vom Hofe et al., 2005; Wartha & vom Hofe, 2005). Die Stichproben der insgesamt fünf Interviewstudien bestanden aus 36 Gymnasiasten (Studie 1, Jahrgangsstufe 6); 41 Hauptschülern, Realschülern und Gymnasiasten (Studie 2, Jahrgangsstufe 6); 24 Hauptschülern und Gymnasiasten (Studie 3, Jahrgangsstufe 6); 36 Gymnasiasten (Studie 4, Jahrgangsstufe 7); und 36 Realschülern (Studie 5, Jahrgangsstufe 8). Wesentliches Ziel ist es, Schülerstrategien bei der Bearbeitung mathematischer Aufgaben zu analysieren. Die Studien 2 und 4 haben darüber hinaus das Emotionserleben im Fach Mathematik sowie Beziehungen zwischen Elternhaus und Schule im Fach Mathematik untersucht. Die Studien waren z.T. längsschnittlich miteinander verknüpft; so hatten von den 36 Gymnasiasten der Studie 4 (Jahrgangsstufe 7) 26 bereits an den Interviews der Studie 1 (Jahrgangsstufe 6) teilgenommen.

In den Interviews zum Umgang mit mathematischen Aufgaben wurden die Schülerinnen und Schüler aufgefordert, schriftlich vorgelegte Aufgaben, die zu entsprechenden Aufgaben aus der Hauptuntersuchung parallelisiert waren, zu bearbeiten und ihre Lösungsprozesse anhand von lautem Denken mitzuteilen. Die Bearbeitungszeit für die Aufgaben betrug 30 Minuten. Mit einem Training der drei Interviewer und einem Interviewleitfaden wurde angestrebt, Interviewereffekte gering zu halten. Die Interviews wurden akustisch und videographisch aufgezeichnet und anhand der akustischen Aufzeichnungen transkribiert. In die Datenanalysen werden neben den Transkripten auch die Schüleraufzeichnungen auf den Aufgabenblättern und die Videoaufnahmen der Interviews einbezogen. Die Interviews dienen zur Analyse der Lösungsstrategien, Grundvorstellungen und Fehlvorstellungen von Schülerinnen und Schülern und zur Gewinnung von authentischem Beispielmaterial für die Praxis (vgl. Wartha & vom Hofe, 2005).

3.3 Aufbereitung der Daten und Umgang mit fehlenden Werten

Im Anschluss an die Erhebungen wurden die Daten jeweils quantitativ und qualitativ aufbereitet. Ein spezifisches Datenproblem bei Längsschnittuntersuchungen stellen Stichprobenausfälle und fehlende Werte aufgrund von Klassen- und Schulwechsel, Abwesenheit am Erhebungstag oder unvollständig ausgefüllten Testheften dar. Wie oben dargestellt, wurden Schülerinnen und Schüler nach Klassen- oder Schulwechsel nach Möglichkeit in der Längsschnittstichprobe belassen und weiter in die Erhebungen einbezogen. Da ferner bei der Ausgangsstichprobe des Längsschnitts (Jahrgangsstufe 5) wie bei den Stichproben der nachfolgenden Messzeitpunkte jeweils hohe Ausschöpfungsquoten erreicht wurden (s.o.) und fehlende Werte aufgrund unvollständig ausgefüllter Testhefte in nur geringem Umfang auftraten, hielten sich entsprechende Probleme bei den bisher durchgeführten Erhebungen in Grenzen. Soweit dennoch Probleme auftraten, wurden Verfahren von bedingter Mittelwertbildung zur Ermittlung von Skalenwerten sowie singuläre und multiple Imputationen von fehlenden Werten verwendet (Pekrun, vom Hofe & Blum, 2006).

4 Empirische Befunde

Wie oben dargestellt liegen zum Zeitpunkt dieses Berichts Längsschnitt- und Interviewbefunde für den Entwicklungszeitraum von der 5. bis zur 8. Jahrgangsstufen vor. Im Folgenden wird in jeweils exemplarischer Weise auf zentrale Befunde in den folgenden Bereichen eingegangen: (1) Entwicklung von mathematischen Kompetenzen, (2) Mathematikemotionen und weitere Schülermerkmale, (3) Unterricht im Fach Mathematik sowie (4) die Rolle des Elternhauses.

4.1 Entwicklung von mathematischen Kompetenzen

Mit den Längsschnittanalysen zum Erwerb mathematischer Kompetenzen soll untersucht werden, welche Entwicklung diese Kompetenzen über die Jahrgangsstufen hinweg zeigen, ob sich für verschiedene Schularten und Schülergruppen unterschiedliche Entwicklungsverläufe ergeben und in welchem Verhältnis Entwicklungen in unterschiedlichen Kompetenzklassen und Inhaltsbereichen stehen. Dabei sind insbesondere auch Lernverläufe von Interesse, die mit der Einführung neuer Formalismen ohne gleichzeitiges Aufgreifen intuitiver Vorerfahrungen oder mit Vergessenseffekten aufgrund mangelnder curricularer und unterrichtlicher Vernetzung zusammenhängen. Inhaltlich steht bei den Analysen zu den ersten drei Messzeitpunkten (Jahrgangsstufen 5 bis 7) mit der ersten grundlegenden Zahlbereichserweiterung das Gebiet der Bruchrechnung im Vordergrund. Hier ist eine wesentliche Frage des Projekts, inwieweit – so wie theoretisch erwartet – die in diesem Bereich zu beobachtenden Fehlstrategien auf Defiziten in der Ausbildung von Grundvorstellungen beruhen.

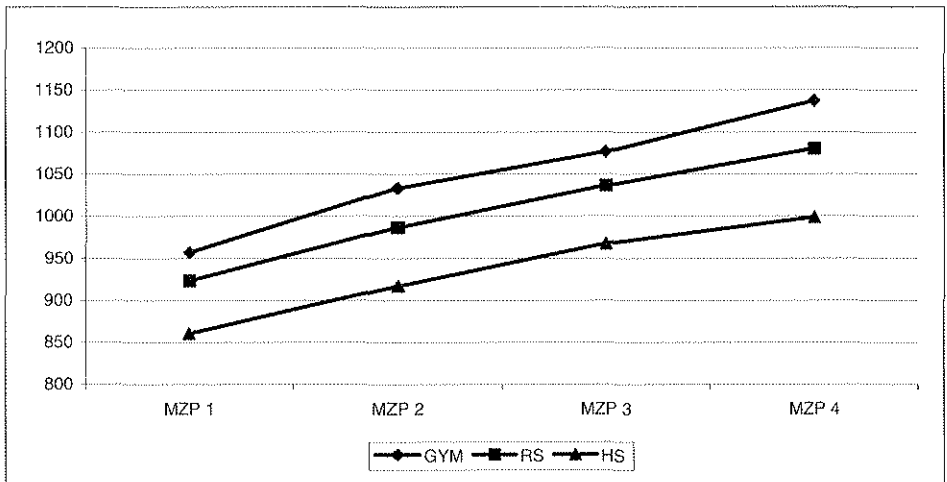


Abb. 2: Kompetenzentwicklung von der 5. bis 8. Jahrgangsstufe nach Schulform

4.1.1 Befunde der quantitativen Erhebungen zum Kompetenzerwerb

Die Befunde der quantitativen Erhebungen zeigen über die ersten vier Messzeitpunkte hinweg einen signifikanten Anstieg der Gesamleistungswerte (ANOVA für die Kompetenzwerte der vier Messzeitpunkte: $F [3, 1411] = 3122, p < .001$; erster Messzeitpunkt: $M = 919,32, SD = 72,76$; zweiter Messzeitpunkt: $M = 985,95, SD = 76,33$; dritter Messzeitpunkt: $M = 1033,16, SD = 76,49$; vierter Messzeitpunkt: $M = 1080,23, SD = 91,06$). *Post-hoc*-Tests bestätigen, dass es sich in jedem Längsschnittintervall um einen signifikanten Anstieg handelt (MZP1-2: $t (1411) = 49,15, p < .001$; MZP 2-3: $t (1411) = 36,29, p < .001$; MZP 3-4: $t (1411) = 29,25, p < .001$; vgl. vom Hofe et al., 2005). Die angegebenen Werte beziehen sich auf die Längsschnittstichprobe. Sie sind auf einen Mittelwert von 1.000 und eine Standardabweichung von 100 Punkten normiert und beruhen auf einer Skalierung der Daten der Stichprobe über alle vier Messzeitpunkte hinweg. Im Vergleich zu den Befunden aus TIMSS liegt der ermittelte Leistungsanstieg hier deutlich höher (der Leistungsanstieg von der 7. zur 8. Klasse betrug bei TIMSS in Deutschland 0.25 Standardabweichungen, vgl. Max-Planck-Institut für Bildungsforschung, 1997).

Ein kontinuierlicher Anstieg zeigt sich auch für die Leistungsentwicklung in den einzelnen Schularten (Abbildung 2). Dabei deutet sich im Sinne eines leichten Schereneffekts ähnlich wie z.B. in der Kassel-Exeter-Studie (vgl. Blum et al., 1994) an, dass im Gymnasium ein insgesamt etwas stärkerer Kompetenzzuwachs und in der Hauptschule ein etwas geringerer Zuwachs zu verzeichnen ist als im Durchschnitt der Stichprobe (die Interaktion von Messzeitpunkt und Schulart ist signifikant; $F [1418, 2] = 37, p < .01$). Soweit hier Effekte einer Regression zur Mitte über die Schuljahre hinweg eine Rolle spielen, kann nicht ausgeschlossen werden, dass dieser Schereneffekt bezüglich der tatsächlichen Schülerkompetenzen noch stärker ist, als dies in den vorliegenden, messfehlerbehafteten Kompetenzwerten deutlich wird (vgl. Marsh & Hau, 2002).

Welche Entwicklungen von Modellierungs- vs. Kalkülkompetenzen liegen diesen Veränderungen der Gesamtwerte zugrunde? Gegenwärtig liegen hierzu die Subskalenwerte für die ersten drei Messzeitpunkte vor. Die Befunde zeigen, dass die Durchschnittswerte für Kalkülkompetenzen durchweg oberhalb der Werte für Modellierungskompetenzen liegen (Abbildung 3a). Dies gilt für alle Schulformen (Abbildungen 3b-d). Allerdings sind die Wertedifferenzen in Gymnasium und Realschule stärker ausgeprägt als in der Hauptschule. Für die Hauptschüler zeigt sich, dass die Modellierungskompetenzen einen höheren Zuwachs aufweisen als die Kalkülkompetenzen und zum dritten Messzeitpunkt fast dasselbe Niveau erreichen.

Die Subskalen zu Modellierungs- und Kalkülkompetenzen wurden so konstruiert, dass die jeweiligen Items die gleiche kognitive Komplexität aufweisen. Die Befunde legen deshalb zwei Schlussfolgerungen nahe:

- 1) In den untersuchten Jahrgangsstufen scheint durchgängig ein Übergewicht der Kalkülkompetenzen gegenüber der Ausbildung von Modellierungskompetenzen zu bestehen.
- 2) Unsere anfängliche Vermutung, dass sich die Differenz zwischen diesen Kompetenzen über die Jahrgangsstufen hinweg noch vergrößert, lässt sich hingegen nicht bestätigen.

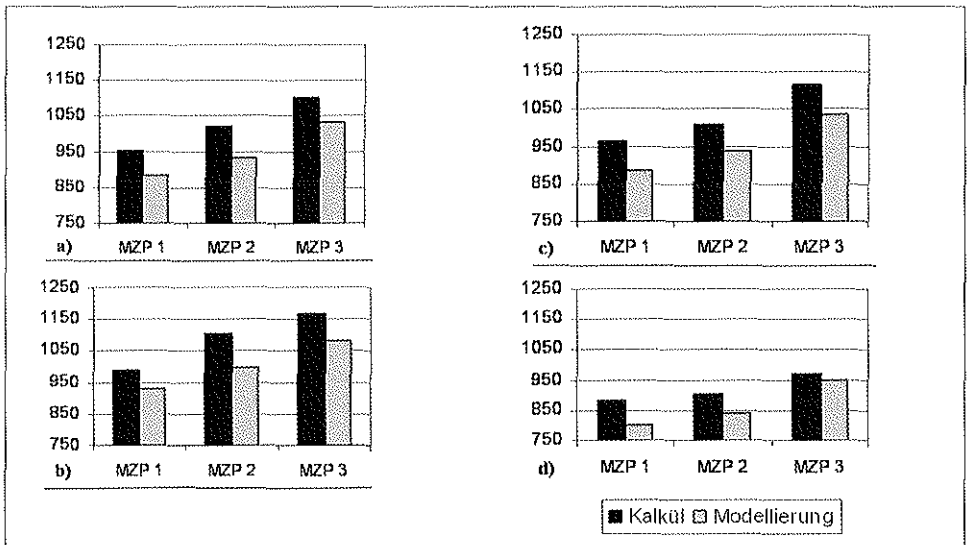


Abb. 3: Durchschnittlicher Fähigkeitswert in den untersuchten Kompetenzbereichen
 a) insgesamt, b) Gymnasium, c) Realschule, d) Hauptschule

Während dieser zweite Befund erfreulich ist, weist der erste Befund auf ein Problem hin, das von grundsätzlicher Art zu sein scheint. Die Datenlage spricht dafür, dass die in den PISA-Erhebungen bei deutschen Schülerinnen und Schülern der Jahrgangsstufe 9 festgestellten Defizite im Modellierungsbereich (Blum, Neubrand et al., 2004) keine alters- bzw. jahrgangsstufenspezifischen Phänomene sind, sondern auf einer lang-jährigen Entwicklung beruhen, die sich vermutlich durch die gesamte Sekundarstufe I zieht und insofern auf ein grundsätzliches Problem des deutschen Mathematikunterrichts aller Jahrgangsstufen der Sekundarstufe I schließen lässt. Die Befunde des fünften und sechsten Messzeitpunkts des Längsschnitts (Jahrgangsstufen 9 und 10) werden zeigen, inwieweit sich diese Entwicklung tatsächlich bis zum Abschluss der Sekundarstufe I verfolgen lässt.

Dabei gibt es allerdings beträchtliche Unterschiede zwischen einzelnen Schulen und Schulklassen. In Abbildung 4 wird dies am Beispiel von zwei Gymnasialklassen mit günstigen vs. problematischen Verläufen im Bereich der Bruchrechnung verdeutlicht. Während bei Klasse 56 ein erheblicher Kompetenzzuwachs von der fünften zur sechsten Jahrgangsstufe und ein weiterer Zuwachs von der sechsten zur siebten Jahrgangsstufe festzustellen ist, stagnieren die durchschnittlichen Kompetenzwerte in Klasse 31 auf relativ niedrigem Niveau. Wir arbeiten gegenwärtig daran, Art und Ursachen solcher Defizite näher zu untersuchen. Erste Analysen zeigen deutliche Hinweise auf unterrichtliche Ursachen für problematische Verläufe wie bei Klasse 31: Im Vergleich zu Klasse 56 ist der Unterricht aus Schülersicht in Klasse 31 durch ein signifikant geringeres Maß an Modellierung und Vernetzung gekennzeichnet. Dies weist auf einen vergleichsweise schematischen und rezeptiven Unterricht hin, der für die Defizite in der Kompetenzentwicklung mitverantwortlich sein könnte.

Kompetenzwerte Bruchrechnung

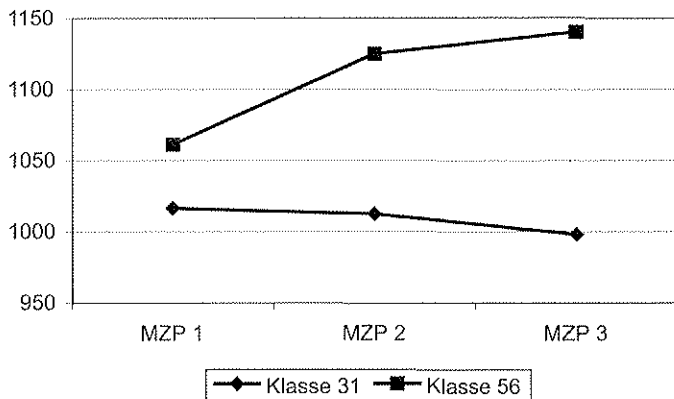


Abb. 4: Kompetenzentwicklung Subskala.
Bruchrechnung in Klasse 31 (N = 25) und 56 (N = 26)

4.1.2 Schülerinnen und Schüler mit Leistungsdefiziten („Risikogruppe“)

Für Teilstichproben von Schülerinnen und Schülern, die aufgrund niedriger Kompetenzwerte Anlass zur Sorge über den weiteren schulischen und beruflichen Werdegang geben, wurden vertiefende Analysen durchgeführt. In Anlehnung an die bei den PISA-Erhebungen übliche Terminologie (Baumert et al., 2001) lassen sich diese Schülerinnen und Schüler als „Risikogruppe“ bezeichnen. Bei PISA war die Zugehörigkeit zur „Risikogruppe“ anhand der individuellen Kompetenzwerte definiert worden (PISA-Skalenwert für mathematische Kompetenz gleich oder kleiner 420; Blum, Neubrand et al., 2004). Angesichts der relativen Beliebigkeit eines solchen Kriteriums ist dies nicht unproblematisch. Hier wurde die Zuordnung deshalb – in jeweils jahrgangsstufenspezifischer Weise – anhand inhaltlicher Kriterien bestimmt. Hierfür wurde zunächst ein qualitatives Rating aller Items des Mathematiktests durchgeführt, wobei ein Item dann als risikorelevantes Item eingestuft wurde, wenn die folgenden Kriterien erfüllt waren:

- 1) Zur Lösung des Items sind höchstens elementare Kenntnisse zu Inhalten des Lehrplans der jeweiligen Jahrgangsstufe auf *Hauptschulniveau* erforderlich.
- 2) Die durch das Item repräsentierten Teilkompetenzen werden in den nachfolgenden Jahrgangsstufen nicht mehr vermittelt, sondern vorausgesetzt.
- 3) Die zur Lösung des Items benötigten Kompetenzen werden als relevant für Beruf bzw. alltägliches Leben angesehen. Als Richtlinie für die Berufsrelevanz wurde das Anforderungsprofil derzeit üblicher Berufseingangstests der Industrie- und Handelskammern verwendet.

Für die empirische Definition der Risikogruppe auf der Basis dieser Kriterien wurde der *Threshold* des schwierigsten risikorelevanten Items als oberer Grenzwert des Fähigkeitsparameters bestimmt. Schülerinnen und Schüler mit einem niedrigeren Fähigkeitsparameter werden als „Risikoschüler“ klassifiziert. Dies beinhaltet, dass „Risikoschüler“ das schwierigste unter den risikorelevanten Items mit nur 50-prozentiger Wahrscheinlichkeit lösen können.

Unter Zugrundelegung dieser Definition ergeben sich die in Abbildung 5 dargestellten Verteilungen über Jahrgangsstufen und Schularten hinweg (der vierte Messzeitpunkt konnte hier noch nicht einbezogen werden). Zunächst fällt auf, dass die „Risikogruppe“ mit jeder Jahrgangsstufe anwächst (von 15% in der Jahrgangsstufe 5 bis zu knapp 19% in der Jahrgangsstufe 7). Deutlich wird ferner, dass es sich bei den Schülerinnen und Schülern der „Risikogruppe“ jeweils vor allem um Hauptschüler handelt. So sind etwa in der Jahrgangsstufe 5 nur ca. zwei von 100 Schülerinnen und Schülern des Gymnasiums und etwa sechs von 100 Realschülern der „Risikogruppe“ zuzuordnen, hingegen mehr als ein Drittel aller Hauptschüler. Bei 85 Prozent aller Schüler der „Risikogruppe“ in dieser Jahrgangsstufe handelt es sich um Hauptschüler. Dieses Bild setzt sich in den Jahrgangsstufen 6 und 7 fort, wobei im Bereich der Hauptschule ein jährlicher Zuwachs von etwa 10% zu verzeichnen ist. Vor dem Übertritt in die Jahrgangsstufe 8 konnte die Hälfte der teilnehmenden Hauptschüler die laut Lehrplan für diese Jahrgangsstufe relevanten Aufgaben nicht hinreichend lösen.

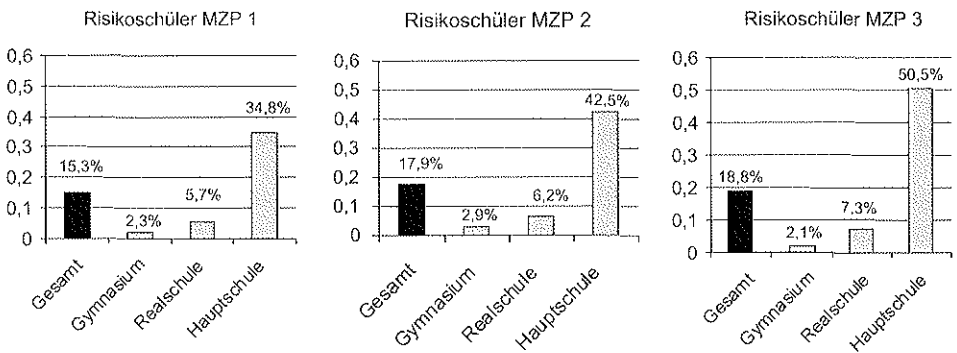


Abb. 5: „Risikoschüler“ nach Schularten

Die Daten für diese „Risikogruppen“ werden gegenwärtig näher analysiert. Der Zusammenhang mit der Mathematik-Jahreszeugnisnote ist wie erwartet: Alle Schülerinnen und Schüler mit der Zensur 6 und über die Hälfte der Schüler mit der Zensur 5 wurden als „Risikoschüler“ identifiziert. Kongruent hierzu haben die „Risikoschüler“ deutlich ungünstigere Fähigkeitsselbstkonzepte und höhere Angstwerte in Mathematik als der Schülerschnitt. Ferner ergeben sich erste Hinweise, dass die Klassenzugehörigkeit von Jahrgangsstufe zu Jahrgangsstufe zunehmend bedeutsamer dafür wird, ob ein Schüler bzw. eine Schülerin der „Risikogruppe“ des betreffenden Jahrgangs zuzuordnen ist oder nicht – Schülerinnen und Schüler der „Risikogruppe“ treten in den höheren Jahrgangsstufen in verstärktem Maße „geballt“ in einzelnen Klassen auf. Dieser Befund lässt vermuten, dass es Schulklassen gibt, in denen ein Abgleiten in die „Risikogruppe“ aufgrund günstiger unterrichtlicher oder sozioökonomischer Bedingungen verhindert oder reduziert wird, während andere Klassen solche Risikoentwicklungen begünstigen.

In weiterführenden Analysen wird zu prüfen sein, wie sich diese Entwicklung über die nachfolgenden Jahrgangsstufen fortsetzt und welche Schüler-, Unterrichts- und Kontextbedingungen ihr jeweils zugrunde liegen. Aus diesen Analysen sind Erkenntnisse zu erhoffen, anhand welcher Unterrichts- und Stützmaßnahmen defizitäre Kompetenzentwicklungen im Fach Mathematik verhindert oder reduziert werden können.

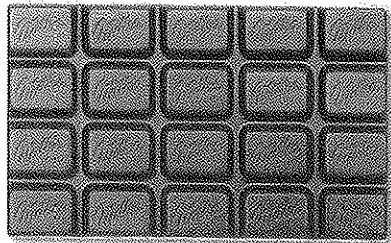
4.1.3 Befunde aus den Interviewstudien

In den vertiefenden Interviewstudien haben wir Lösungsstrategien bei der Aufgabenbearbeitung im Detail analysiert. Untersucht werden sollte insbesondere, welche Grundvorstellungen jeweils aktiviert werden bzw. welche Fehlvorstellungen eine Lösung gegebenenfalls verhindern. Die Interviewserien konzentrierten sich auf zentrale Gebiete der mathematischen Kompetenzentwicklung in den jeweiligen Jahrgangsstufen (Jahrgangsstufe 6: Bruchrechnung und Proportionalität; Jahrgangsstufe 7: Bruchrechnung und negative Zahlen; Jahrgangsstufe 8: Gleichungen und Funktionen).

Ein zentraler Befund für Probanden, die sowohl in der 6. wie in der 7. Jahrgangsstufe an den Interviews teilgenommen hatten, liegt im Nachweis spezifischer mentaler Fehlvorstellungen zur Bruchrechnung, die vermutlich aus Übergeneralisierungen robuster älterer Grundvorstellungen entstanden waren, und in Hinweisen auf eine hohe zeitliche Stabilität dieser Vorstellungen. Im Folgenden wird ein Beispiel zur multiplikativen Anteilbildung vorgestellt, das den Charakter solcher mentalen Fehler demonstriert.

Aufgabe Schokolade

Lily nimmt sich die Hälfte der dargestellten Tafel Schokolade. Davon isst sie $\frac{3}{5}$ auf. Wie viel Stücke hat sie gegessen?



Entgegen der erwarteten Anteilbildung über einen Proportionalitätsschluss oder durch die multiplikative Anteiloperation $\frac{3}{5} \times 10$ führt die hier ausgewählte Schülerin eine Division durch. Auf Nachfragen des Interviewers wird ihre Fehlstrategie deutlich:

- 44 I: Wieso jetzt geteilt durch null Koma sechs?
- 45 S: Ja ehm, ah.
- 46 Weil also zehn ist ja null Komma fünf und das ist ja die Hälfte von dem Ganzen.
- 47 Und die null Komma sechs sind die drei Fünftel.
- 48 Und das braucht man ja von der Hälfte, weil's ja nur die Hälfte ist, und nicht das
- 49 Ganze.
- 50 I: Meine Frage bezieht sich jetzt mehr auf das Rechenzeichen. Wieso geteilt durch?
- 51 Und nicht mal oder plus oder minus?
- 52 S: Ja weil, ehm, dann wär's ja mehr und das muss ja immer weniger werden, weil sie
- 53 isst ja nicht mehr, als Tafel da ist, sondern weniger als die Tafel da ist.
- 54 *S führt die Division $10 : 0,6$ korrekt durch. (26 Sekunden)*

Wie in diesem Beispiel wurde die fehlerhafte Wahl des Rechenzeichens in den Interviews in erster Linie damit begründet, dass nur auf diese Weise eine Verkleinerung des Ausgangswertes zu erreichen sei. Diese typische Fehlstrategie beruht auf einer nicht

adäquat erweiterten Grundvorstellung zur Multiplikation: Während eine Multiplikation bei natürlichen Zahlen stets vergrößert, kann diese Rechenoperation bei Bruchzahlen auch verkleinern. Die fehlerverursachende Wirkung von nicht erweiterten Grundvorstellungen wurde durch Detailanalysen von Daten aus den quantitativen Erhebungen bestätigt (Wartha, 2005).

In der Jahrgangsstufe 7 wurde die zitierte Schülerin mit einer weiteren parallelierten Fassung dieser Aufgabe konfrontiert. Interessanterweise wählt die Schülerin auch ein Jahr nach dem oben dargestellten Interview eine (unpassende) Rechenoperation, die ihrer Ansicht nach die Verkleinerung der Ausgangsgröße bewirkt. Dies spricht für eine hohe Stabilität ihrer Fehlvorstellungen.

Zahlreiche weitere Befunde aus den Interviewerhebungen bestätigen die Wirksamkeit von Grundvorstellungen im Lösungsprozess. Die Befunde belegen die positiven Wirkungen vorhandener Grundvorstellungen ebenso wie die Probleme, die sich ergeben, wenn Grundvorstellungen nicht hinreichend ausgebildet oder nicht angemessen erweitert wurden (vgl. auch Fischbein et al., 1990; vom Hofe & Wartha, 2005). Insgesamt lässt sich ca. die Hälfte aller in den qualitativen Studien analysierten Fehler beim Aufgab lösen auf nicht erfolgreich ausgebildete Grundvorstellungen zurückführen.

4.1.4 Schlussfolgerungen

Die Befunde zeigen, dass die Durchschnittswerte für mathematische Schülerkompetenzen über den Zeitraum von der 5. zur 8. Jahrgangsstufe hinweg zunehmen, die Entwicklung aber je nach Kompetenzbereich und Schülergruppe recht unterschiedlich verläuft. Bereits von der Jahrgangsstufe 5 an sind modellierungsorientierte Kompetenzen schwächer ausgeprägt als Kalkülkompetenzen. Eine über die Jahrgangsstufen hinweg zunehmend größere Zahl von Schülerinnen und Schülern ist einer „Risikogruppe“ zuzuordnen, die auch einfache Aufgabenstellungen der jeweiligen Jahrgangsstufe nicht beherrscht. Dabei belegen die Befunde der qualitativen Analysen, dass unzureichend ausgebildeten Grundvorstellungen eine zentrale Rolle für die Schwierigkeiten von Schülerinnen und Schülern mit mathematischen Aufgaben zukommt; eine stärkere Förderung dieser Vorstellungen könnte ungünstigen Lernverläufen entgegenwirken. Ferner ergeben sich erste Hinweise zur Bedeutung von Schulart, Mathematikunterricht und Klassenzugehörigkeit für problematische Entwicklungen, aus denen sich in den weiteren Analysen ebenfalls Ansatzpunkte für pädagogische Konsequenzen erwarten lassen.

4.2 Emotionen und Schülermerkmale in Mathematik

4.2.1 Entwicklungsverläufe

Welche Veränderungen von Emotionen, Motivation und Lernverhalten im Fach Mathematik stehen den Kompetenzentwicklungen in diesem Fach gegenüber? Für die ersten vier Messzeitpunkte (Jahrgangsstufen 5 bis 8) werden hier exemplarisch die Veränderungen der durchschnittlichen Ausprägungen einiger zentraler Merkmale dargestellt (Tabelle 3). Wie man sieht, sind die Durchschnittswerte für das Erleben von leistungsbezogener Angst und Scham im Fach Mathematik relativ stabil, während die

Werte für positive Emotionen abnehmen und für Langeweile ansteigen. Allerdings fallen diese Veränderungen über die Jahrgangsstufen hinweg zunehmend geringer aus; es scheint sich um eher asymptotische Verläufe zu handeln. Die Entwicklung ist kongruent zur Entwicklung der durchschnittlichen Werte für Selbstwirksamkeit in Mathematik, die sich ebenfalls zu stabilisieren scheinen. Auch die Abnahme der Werte für selbstberichtete Anstrengung und wahrgenommene Selbstregulation des Lernens wird im Laufe der Zeit geringer.

Insgesamt aber ergeben sich über die vier Jahrgangsstufen hinweg deutliche Veränderungen von bis zu zwei Dritteln einer Standardabweichung. Dies spricht dafür, dass im bisher abgebildeten Zeitraum ähnlich wie bei der Kompetenzentwicklung eine erhebliche Entwicklungsdynamik zu verzeichnen ist. Entwicklungsdynamik ist im Übrigen auch aus den Koeffizienten interindividueller Merkmalsstabilität über die Jahrgangsstufen hinweg zu folgern (Stabilität der Emotionswerte von $.46 < r < .63$ für die drei Ein-Jahres-Intervalle und $.31 < r < .40$ für das Drei-Jahres-Intervall von der 5. zur 8. Jahrgangsstufe).

Mögliche Gründe für die eher ungünstige Entwicklung von positiven Mathematikemotionen, Selbstwirksamkeit, Anstrengung und wahrgenommener Selbstregulation in Mathematik könnten in der Abnahme der von den Schülerinnen und Schülern berichteten kognitiven Aktivierung im Mathematikunterricht und dem Mangel an (wahrgenommener) Autonomieunterstützung durch Lehrkräfte und Eltern über die Jahrgangsstufen hinweg liegen (s. Abschnitt 4.3 sowie Pekrun, vom Hofe & Blum, 2003).

Tab. 3: Mathematikemotionen und Schülermerkmale: Entwicklungen von der 5. bis zur 8. Jahrgangsstufe

	Mittelwert				Standardabweichung				F-Werte
	5	6	7	8	5	6	7	8	
<i>Emotionen</i>									
Freude	29.73	26.72	24.69	24.14	7.58	7.55	7.27	7.03	244.68**
Stolz	25.53	23.83	22.13	22.03	7.52	7.41	7.14	7.17	103.28**
Ärger	16.14	17.53	19.34	19.11	7.42	7.60	7.53	7.58	77.89**
Angst	34.12	34.46	34.99	33.41	12.36	12.56	12.48	12.42	10.16**
Scham	14.97	15.08	15.08	14.32	6.92	6.91	6.81	6.68	7.30**
Hoffnungslosigkeit	12.11	12.63	13.40	13.08	5.75	5.93	6.08	6.22	18.66**
Langeweile	11.74	13.50	15.15	15.31	5.75	6.21	6.32	6.20	144.76**
<i>Motivation, Regulation</i>									
Selbstwirksamkeit	14.10	13.44	12.58	12.79	3.55	3.89	3.78	3.92	76.21**
Interesse	18.66	15.91	14.06	13.67	5.95	5.98	5.54	5.47	333.25**
Anstrengung	26.70	24.69	24.03	23.73	5.04	5.75	5.91	5.90	100.69**
Selbstreg. Lernen	21.56	20.55	19.67	19.56	4.66	4.50	4.52	4.48	73.59**
Fremdreg. Lernen	17.21	16.13	15.38	15.13	4.71	4.35	4.38	4.37	67.64**

Anmerkungen: N = 1.424. *p < .05. **p < .01.

4.2.2 Bedingungen der Emotionsentwicklung:

I. Emotionen bei Modellierungs- vs. Kalkülaufgaben

Rufen Modellierungsaufgaben andere Emotionen hervor als Kalkülaufgaben? Diese Frage ist unter mathematikdidaktischer Perspektive wie auch für die Diagnostik von Mathematikemotionen von erheblicher Bedeutung. Modellierungsaufgaben sind aus didaktischer Sicht entscheidend für die mathematische Bildung. Lehrkräfte aber berichten aus ihrem Unterrichtsalltag, dass Aufgaben dieser Art (z.B. Textaufgaben) bei Schülerinnen und Schülern eher unbeliebt seien. Allerdings könnte es sein, dass im Unterricht eingesetzte Modellierungsaufgaben relativ schwierig und deshalb wenig beliebt sind. Notwendig ist es deshalb, die von Modellierungs- vs. Kalkülaufgaben hervorgerufenen Emotionen unter Kontrolle der jeweiligen Aufgabenschwierigkeit zu untersuchen.

Möglich wird eine solche Untersuchung mit den hier entwickelten Skalen zu Freude, Angst und Langeweile bei Modellierungs- und Kalkülaufgaben (s. Abschnitt 2.3). Die gegenwärtig vorliegenden Analysen zu Daten des dritten Messzeitpunkts (Frenzel, Jullien & Pekrun, 2006; Zirngibl et al., 2005) zeigen zunächst, dass Modellierungs- und Kalkülemotionen deutlich miteinander korrelieren ($r = .59, .65$ und $.71$ für Freude, Angst und Langeweile bei Modellierungs- vs. Kalkülaufgaben; disattenuiert entspräche dies – bei Konsistenzreliabilitäten um $\alpha = .70$ – Korrelationen nahe Eins). Die geringe interindividuelle Differenzierung der beiden Gruppen von Emotionen belegt, dass es gerechtfertigt ist, bei einer Erhebung individueller Ausprägungen von Mathematikemotionen auf aufgabenspezifische Differenzierungen zu verzichten. Für Zwecke interindividuell angelegter Analysen scheint es nicht notwendig zu sein, über die in den Münchener Skalen zu Mathematikemotionen vorgesehene fachbezogene Spezifizierung von Schüleremotionen hinauszugehen.

Unabhängig hiervon aber zeigt sich im Mittelwertvergleich der Emotionen zu den beiden Aufgabentypen, dass Schülerinnen und Schüler durchaus unterschiedlich auf die beiden Aufgabentypen reagieren: Sie berichteten signifikant mehr Angst und weniger Freude für die Kalkülaufgaben als für die Modellierungsaufgaben. Dies gilt für leistungsschwache Schülerinnen und Schüler in noch deutlicherem Maße als für leistungsstarke. Zwar werden die Unterschiede für Langeweile nicht signifikant, und auch für Freude und Angst sind die Effektstärken gering. Dennoch ist dieser Befund von erheblicher unterrichtspraktischer Bedeutung, denn er zeigt, dass Modellierungsaufgaben von Schülerinnen und Schülern – entgegen den erwähnten Vorannahmen vieler Lehrkräfte – emotional zumindest nicht ungünstiger bewertet werden als gleich schwierige Kalkülaufgaben. Einem Einsatz solcher Aufgaben steht damit aus einer emotionspädagogischen Perspektive nichts im Wege (solange darauf geachtet wird, dass der Schwierigkeitsgrad der Aufgaben nicht als zu hoch erlebt wird).

4.2.3 Bedingungen der Emotionsentwicklung: II. Mehr-Ebenen-Analysen zum Einfluss der Mathematikleistung

Aus den Annahmen des Kontroll-Wert-Modells der Leistungseemotionen (Pekrun, 2000; in press) folgt, dass Leistungen und Leistungsrückmeldungen als eine zentrale ontogenetische Determinante der Entwicklung leistungsbezogener Emotionen anzusehen sind. Empirisch mangelt es hierzu national wie international an Evidenz (mit Ausnahme von Befunden für Prüfungsangst; vgl. Pekrun, 1992; Zeidner, 1998). Dabei

fehlen insbesondere auch Untersuchungen, die der Mehr-Ebenen-Struktur schulischer Leistungen Rechnung tragen.

In Mehr-Ebenen-Analysen haben wir deshalb untersucht, welchen Einfluss die individuelle Mathematikleistung einerseits und das Leistungsniveau der Schulklasse andererseits auf die Entwicklung von Emotionen im Fach Mathematik nehmen (Götz et al., 2004; Pekrun & Götz, 2005). Ausgangsannahme war, dass der *big-fish-little-pond*-Effekt (BFLP-Effekt; Marsh, 1987) von Leistungen auf die Selbstkonzeptentwicklung in den Bereich der Emotionsentwicklung übertragbar ist: Erfolg und Misserfolg schlagen sich in den individuellen Fähigkeitsselbstkonzepten nieder, diese aber sind ihrerseits als wesentliche Bedingungen der leistungsbezogenen Emotionsentwicklung anzusehen. Der BFLP-Effekt beinhaltet, dass die Fähigkeitsselbstkonzepte von Schülerinnen und Schülern umso höher sind, je *höher* das individuelle Leistungsniveau und je *niedriger* das Leistungsniveau der Bezugsgruppe ist (da Erfolge sich günstig auf Selbstkonzepte auswirken, bei niedrigerem Leistungsniveau der jeweiligen Bezugsgruppe aber wahrscheinlicher sind als bei höherem Leistungsniveau der Gruppe).

Für die Emotionsentwicklung dürfte dabei auch eine Rolle spielen, dass nicht nur die individuellen Erfolge und Misserfolge, sondern auch die erfolgs- und misserfolgskontingenten, emotionserzeugenden Reaktionen von Bezugspersonen durch die Verortung eines Schülers bzw. einer Schülerin im Leistungskontinuum der Bezugsgruppe beeinflusst sein können (Götz et al., 2004). Erwartet wurden deshalb (a) emotionsgünstige Wirkungen der individuellen Leistung, d.h. positive Effekte der Individualleistung auf positive Emotionen und negative Effekte auf negative Emotionen, sowie (b) emotionsungünstige Effekte des mittleren Leistungsniveaus der Schulklasse (nach Ausparialisierung der Individualleistung), also negative Effekte auf positive Emotionen und positive Effekte auf negative Emotionen.

Überprüft wurden diese Annahmen anhand längsschnittlicher Mehr-Ebenen-Analysen, in denen Mathematikfreude und Mathematikangst in der Jahrgangsstufe 6 vorhergesagt wurden. Prädiktoren waren die individuellen und die klassenweise aggregierten Kompetenzwerte in der Jahrgangsstufe 5. Die Befunde entsprechen den Erwartungen (Tabelle 5). Wie man sieht, hat die Individualleistung einen positiven Effekt auf Freude und einen negativen Effekt auf Angst in Mathematik, während sich für die Klassenleistung gegenläufige Effekte zeigten. Dieses Befundmuster bleibt erhalten, wenn Freude und Angst in der Jahrgangsstufe 5 als zusätzliche Prädiktoren eingeführt werden (jeweils Modell 2 in Tabelle 4). Nach einer solchen Kontrolle von autoregressiven Effekten lassen sich die Koeffizienten für die Individual- und Klassenleistung als Veränderungseffekte interpretieren. Anhand der individuellen und aggregierten Mathematikleistung können hier also auch die *Veränderungen* von Freude und Angst im Fach Mathematik von der 5. zur 6. Jahrgangsstufe vorhergesagt werden. In zusätzlichen Analysen zeigten sich für die Entwicklung von Stolz, Scham und Hoffnungslosigkeit ähnliche Effekte der Mathematikleistung wie für Entwicklungen von Freude bzw. Angst (Pekrun & Götz, 2005).

Tab. 4: Mehrebenenanalysen; Wirkungen von individueller und aggregierter Leistung (5. Jahrgangsstufe) auf Freude und Angst (6. Jahrgangsstufe; aus Götz et al., 2004)

5. Jahrgangsstufe	Freude in der 6. Jahrgangsstufe (Individualebene)				Angst in der 6. Jahrgangsstufe (Individualebene)			
	Modell 1		Modell 2		Modell 1		Modell 2	
	β	p	β	p	β	p	β	p
<i>Individualebene</i>								
Leistung	.32	<.001	.17	<.001	-.44	<.001	-.22	<.001
Freude			.52	<.001				
Angst							.53	<.001
<i>Klassenebene</i>								
Leistung	-.40	<.001	-.20	<.001	.41	<.001	.21	<.001
R ²	.07		.36		.10		.37	

Die Befunde bestätigen die Annahme, dass die Leistungen von Schülerinnen und Schülern Einfluss auf die Entwicklung ihrer Emotionen nehmen. Im Sinne eines *big-fish-little-pond*-Effekts auf die Emotionsentwicklung zeigen sich dabei gegenläufige Effekte von Individualleistung einerseits und Leistungsniveau der Schulklasse andererseits. Allgemein ist zu folgern, dass eine Modellierung von Merkmalsentwicklungen bei Schülerinnen und Schülern voraussetzt, dass den Mehr-Ebenen-Strukturen von Entwicklungsbedingungen und Kontexten Rechnung getragen wird (zu möglichen pädagogischen Konsequenzen Götz et al., 2004).

4.2.4 Folgen von Emotionen für die Leistungsentwicklung

Aus den Annahmen der Kontroll-Wert-Theorie ist auch zu folgern, dass zwar Leistungen eine entscheidende Quelle der leistungsbezogenen Emotionsentwicklung darstellen, dass aber Emotionen ihrerseits auch Einfluss auf Kompetenzerwerb und Leistungsentwicklung nehmen (Pekrun, in press). Diese Annahme einer reziproken Kausalität von Leistung und Emotionen wurde für die ersten drei Messzeitpunkt (Jahrgangsstufen 5 bis 7) anhand von ersten *cross-lagged*-Strukturgleichungsanalysen für Freude, Angst und Hoffnungslosigkeit einerseits und die Mathematikleistung andererseits untersucht (mit zunächst getrennten Modellbildungen für die drei Emotionen; Pekrun, Jullien et al., 2004). Die Modelle erwiesen sich als gut an die Daten angepasst. Für jedes der beiden Zeitintervalle der Analysen (Jahrgangsstufen 5 und 6 sowie 6 und 7) zeigten sich sowohl signifikante diachrone Effekte der Leistung auf die drei Emotionen wie auch signifikante diachrone Effekte der Emotionen auf die nachfolgende Leistung. Zwar fielen die Effekte der Emotionen auf die Leistung teils geringer aus als Effekte der umgekehrten Wirkrichtung, insgesamt aber bestätigten die Befunde Annahmen zu einer wechselseitigen Beeinflussung von Emotions- und Leistungsentwicklung. Aus Längsschnittbefunden dieser Art ist zu folgern, dass unidirektionale Produktionsmodelle schulischer Bildungseffekte zu kurz greifen, wenn der Komplexität und Multidirektionalität von Kompetenz- und Persönlichkeitsentwicklungen im schulischen Kontext Rechnung getragen werden soll.

4.2.5 Folgen von Emotionen für Bildungsentscheidungen: Mathematikemotionen und die Zweigwahl an Realschulen

Welche Bedeutung kommt den fachbezogenen Emotionen von Schülerinnen und Schülern für ihre Bildungsentscheidungen zu? Aus den Annahmen der Kontroll-Wert-Theorie folgt, dass Emotionen über Leistungen hinaus auch für zukunftsorientierte motivationale Prozesse und resultierende Bildungs- und Berufswahlen wesentlich sind. Die Brauchbarkeit dieser Annahme wurde exemplarisch für die Wahl zwischen alternativen Schulzweigen geprüft, welche die Realschüler unserer Stichprobe mit dem Übergang in die Jahrgangsstufe 7 vorzunehmen hatten (Wahl zwischen mathematisch-naturwissenschaftlichem, betriebswissenschaftlichem und sozial-künstlerischem Zweig). Anhand logistischer Regressionen wurde die Prädiktionskraft von Kompetenzen und Emotionen (Freude und Langeweile) im Fach Mathematik in der Jahrgangsstufe 6 für diese Entscheidung geprüft. Die Befunde zeigen, dass die Wahl des Realschulzweiges wesentlich von der Freude an diesem Fach bestimmt wird. Während die Wahl des mathematischen Zweiges anstelle des sozialen Zweiges sowohl vom Freudeerleben wie von der mathematischen Kompetenz positiv beeinflusst wird ($B_{\text{Freude}} = 1,10$; $B_{\text{Kompetenz}} = 0,39$; jeweils $p < .001$), wird die Entscheidung dafür, den betriebswirtschaftlichen Zweig anstelle des sozialen Zweiges zu wählen, nur durch die Freude am Fach Mathematik signifikant beeinflusst ($B_{\text{Freude}} = 0,40$; $p < .001$). Dies bestätigt Annahmen zur Bedeutsamkeit fachbezogener Emotionen für zukunftsorientierte Entscheidungsprozesse. In den weiteren Analysen wird zu prüfen sein, inwieweit die Befunde auf analoge Entscheidungen in den anderen Schularten übertragbar sind und welche Folgen diese Entscheidungen ihrerseits für die weitere Kompetenz- und Emotionsentwicklung haben.

4.2.6 Schlussfolgerungen

Ähnlich wie beim Erwerb mathematischer Kompetenzen findet sich im Untersuchungszeitraum auch für Emotionen, Motivation und Lernverhalten im Fach Mathematik eine erhebliche Entwicklungsdynamik. Im Schülerdurchschnitt ist in den ersten Jahrgangsstufen der Sekundarstufe zunächst eine Abnahme von positiven Emotionen, Interesse und wahrgenommener Selbstregulation zu verzeichnen, die aber im Laufe der Jahre asymptotisch auszulaufen scheint. Die Befunde zu Mathematikemotionen belegen, dass Kompetenzen und schulische Leistungen für die individuelle Entwicklung zentral sind, wobei individuelle Leistung und das Leistungsniveau der Schulklasse im Sinne eines *big-fish-little-pond*-Effekts auf die Emotionsentwicklung gegenläufige Effekte ausüben. Darüber hinaus zeigt sich in längsschnittlichen Regressions- und Strukturgleichungsanalysen, dass Emotionen ihrerseits Einfluss auf die Leistungsentwicklung und darüber hinaus auch auf Bildungsentscheidungen innerhalb der Schullaufbahn nehmen. Die Befunde liefern damit – im Kontext der Gesamtuntersuchung zunächst vorläufige – Bestätigungen von Annahmen zur Bedeutsamkeit von Schüleremotionen und ihren Verflechtungen mit der Leistungsentwicklung.

4.3 Unterricht im Fach Mathematik

4.3.1 Entwicklungen von der 5. bis zur 8. Jahrgangsstufe

In Tabelle 5 wird die Entwicklung des Mathematikunterrichts von der 5. bis 8. Jahrgangsstufen anhand exemplarischer, schülerperzipierter Unterrichtsmerkmale dargestellt. Wie man sieht, gibt es bis zur Jahrgangsstufe 7 hin eine Abnahme der Durchschnittswerte für einen modellierungsorientierten Unterricht. Mit einer Werteabnahme von insgesamt knapp einer Standardabweichung handelt es sich um eine recht deutliche Veränderung. Dagegen bleiben die Werte für kalkülorientierten Unterricht von der Jahrgangsstufe 6 an weitgehend stabil. Die Werte für Autonomiegewährung, Lehrerenthusiasmus und Unterstützung nach Misserfolg zeigen einen Abfall von der Jahrgangsstufe 5 zur Jahrgangsstufe 6, der sich im darauffolgenden Schuljahr leicht fortsetzt; anschließend scheinen sich aber auch diese Werte zu stabilisieren. Für Variablen wie Leistungsdruck von Lehrerseite oder Zeitverschwendung im Unterricht werden die Werteveränderungen – bedingt durch die Größe der Stichproben – zwar ebenfalls signifikant, von bedeutsamen Veränderungen kann aber nicht die Rede sein.

Insgesamt spricht dieses Befundmuster dafür, dass der Mathematikunterricht bis zur Jahrgangsstufe 8 hin Ausprägungen zeigt, die stabil bleiben oder sich über die Jahrgangsstufen hinweg zunehmend stabilisieren. Bedeutsam ist die Abnahme der Werte für den wahrgenommenen kognitiven Aktivierungsgehalt des Unterrichts, hier durch die – für die Entwicklung von Modellierungskompetenzen mutmaßlich besonders wichtige – Dimension des modellierungsorientierten Unterrichts indiziert. Dieser Befund spricht für eine zunehmend defizitäre Situation im Bereich kognitiver Aktivierung, kongruent zu den oben diskutierten Problemen deutscher Schülerinnen und Schüler bei mathematischen Modellierungsaufgaben.

Tab. 5: Veränderung von Unterrichtsmerkmalen aus Schülersicht (5. bis 8. Klasse)

	Mittelwert				Standardabweichung				F-Werte
	5	6	7	8	5	6	7	8	
Modellierung/ Vernetzung	37.92	35.71	32.62	32.33	8.95	9.50	9.03	9.11	120.32**
Kalkülorientierung	22.02	20.04	19.69	19.70	4.57	4.39	4.67	4.44	96.37**
Autonomiegewährung	21.15	19.48	18.58	18.45	5.05	4.91	5.13	5.19	94.17**
Enthusiasmus	18.21	16.85	15.34	15.59	4.88	5.02	4.89	5.10	101.16**
Unterst. n. Misserfolg	7.41	6.95	6.64	6.56	2.64	2.51	2.43	2.53	37.51**
Bestraf. n. Misserfolg	4.52	4.88	5.05	4.98	2.06	2.21	2.27	2.17	21.87**
Leistungsdruck	9.99	10.10	10.18	9.91	3.40	3.43	3.30	3.30	2.98*
Zeitverschwendung	10.67	10.69	11.15	11.01	3.72	3.62	3.72	3.98	5.67**

Anmerkungen: * $p < .05$. ** $p < .01$.

4.3.2 Dimensionen der Unterrichtswahrnehmung von Mathematiklehrkräften

Ein wesentliches Ziel des Projekts liegt in der Analyse von kumulativen Einflüssen des Mathematikunterrichts auf die Schülerentwicklung im Laufe der Sekundarstufe I. Bereits auf der Ebene querschnittlicher Beziehungen zeigen sich z.T. deutliche Zusammenhänge zwischen Unterrichtsvariablen und den Kompetenzen und Merkmalen von Schülerinnen und Schülern im Fach Mathematik (Pekrun, Götz, vom Hofe et al., 2004). Entscheidender aber wird die Prüfung der Frage sein, welcher Einfluss sich für die Seriation der jeweiligen Unterrichtsverhältnisse über die Jahrgangsstufen hinweg nachweisen lässt. Eine Prüfung solcher Effekte soll nach Abschluss der Längsschnitterhebungen vorgenommen werden.

Ein erster Arbeitsschritt zur Vorbereitung dieser Analysen wurde durchgeführt, indem die dimensional Strukturen von Unterrichtswahrnehmungen analysiert wurden, und zwar zunächst für die Wahrnehmungen von Mathematiklehrkräften. Die Analysen orientieren sich am Vorgehen von Baumert et al. (2004), um Vergleichbarkeit mit entsprechenden Analysen für die Erhebungen von PISA 2003 möglich zu machen. Es sollte festgestellt werden, inwieweit die in den PALMA-Erhebungen erfassten Unterrichtswahrnehmungen theoretisch anzunehmenden Dimensionen von Unterrichtsqualität entsprechen, welche Zusammenhänge sich zwischen den Wahrnehmungsdimensionen nachweisen lassen und welche Wahrnehmungsunterschiede es zwischen Lehrkräften aus Hauptschule, Realschule und Gymnasium gibt.

Da die Daten des vierten Messzeitpunkts noch nicht hinreichend aufbereitet waren, wurden in die Analysen zunächst nur die Antworten von Lehrkräften aus den ersten drei Erhebungszeitpunkten aufgenommen (Klasse 5 bis 7). Falls eine Lehrkraft mehrmals bei den PALMA-Erhebungen teilgenommen hatte, wurde nur der jeweils zeitlich jüngste Datensatz dieser Lehrkraft einbezogen. Es konnten insgesamt Daten von 168 Lehrkräften analysiert werden, wobei 26 Datensätze aus den Erhebungen des ersten Messzeitpunkts und jeweils 71 aus dem zweiten und dritten Messzeitpunkt stammten. Von den 168 Lehrkräften waren 56 am Gymnasium tätig (davon 20% weiblich), 65 an der Realschule (43% weiblich) und 44 an der Hauptschule (41% weiblich).

Zur Analyse der zwölf einbezogenen Unterrichtsskalen wurden zunächst explorative Faktorenanalysen durchgeführt. Die so gewonnene Dimensionierung wurde anschließend anhand einer konfirmatorischen Analyse überprüft. Es konnten fünf Dimensionen der Unterrichtswahrnehmung identifiziert werden, die folgendermaßen benannt werden können (in Klammern jeweils die zugeordneten Skalen): (1) *kognitiv aktivierender Unterricht* (Modellierung und Vernetzung; schüleradaptive Instruktion und Verständlichkeit; Autonomieunterstützung); (2) *Individualisierung* (individuelle Bezugsnormorientierung; positive Verstärkung; Unterstützung nach Misserfolg); (3) *Engagement* (Unterrichtsengagement; Enthusiasmus); (4) *regelerorientiertes Unterrichten* (Kalkülorientierung; Bestrafung nach Misserfolg); und (5) *ineffektive Klassenführung* (Zeitverschwendung; Unterrichtsstörungen). Die konfirmatorisch-faktorenanalytische Prüfung (LISREL 8.72; Jöreskog & Sörbom, 2002) bestätigte diese Struktur ($\chi^2 = 66,82$; $df = 44$, $p = .015$; CFI = .96; NNFI = .93; RMSEA = .057). Inhaltlich entsprechen die fünf Dimensionen theoretisch erwartbaren Basisdimensionen der Unterrichtsqualität (Helmke, 2003; Brophy, 1999; Ditton, 2002; Gruehn, 2000). Gleichzeitig sind sie auch weitgehend kongruent zu den von Baumert et al. (2004) auf der Basis der PISA 2003-Daten identifizierten Faktoren. Dies ist insofern bemerkenswert, als bei den PISA-Erhebungen andere Skalen zum Einsatz gekommen waren als im vorliegenden Projekt. Die Befunde sprechen damit nicht nur für theoretische

Brauchbarkeit des PALMA-Datensatzes zu den Unterrichtswahrnehmungen von Mathematiklehrkräften, sondern gleichzeitig auch für die Äquivalenz zu den PISA-Erhebungen in diesem Bereich.

Auch die Zusammenhänge zwischen den fünf Dimensionen (Abbildung 6) entsprechen weitgehend unseren theoretischen Erwartungen. Im Einklang mit den Ergebnissen von Baumert et al. (2004) zeigt sich, dass Lehrkräfte, die ihren eigenen Unterricht als kognitiv aktivierend beschreiben, auch von einer Individualisierung ihres Unterrichts und einem geringen Ausmaß an ineffektiver Klassenführung berichten. Anders jedoch als bei Baumert et al. (2004), wo sich kein signifikanter Zusammenhang zwischen Individualisierung und ineffektiver Klassenführung zeigt, ist diese Korrelation hier signifikant negativ. Dieser Zusammenhang ist plausibel, da Individualisierung erschwert wird, wenn eine effektive Klassenführung nicht möglich ist. Jenseits der mit Baumert et al. (2004) direkt vergleichbaren Beziehungen zeigt sich, dass Engagement deutlich positiv mit kognitiv aktivierendem Unterricht und negativ mit ineffektiver Klassenführung korreliert. Zusammenhängen dieser Art könnten bidirektionale Bedingungsbeziehungen zugrunde liegen: Engagement begünstigt einen anspruchsvolleren Unterricht, dieser wiederum kann affektiv befriedigender sein und damit das Lehrerengagement fördern. Bemerkenswert ist ferner der positive Zusammenhang zwischen regelorientiertem Unterrichten und ineffektiver Klassenführung. Auch hier sind bidirektionale Beziehungen denkbar (z.B. könnten Unterrichtsstörungen ein stärker regelorientiertes Unterrichten provozieren, das aber seinerseits wiederum Demotivation und Unterrichtsstörungen auf Seiten der Schülerinnen und Schüler begünstigt).

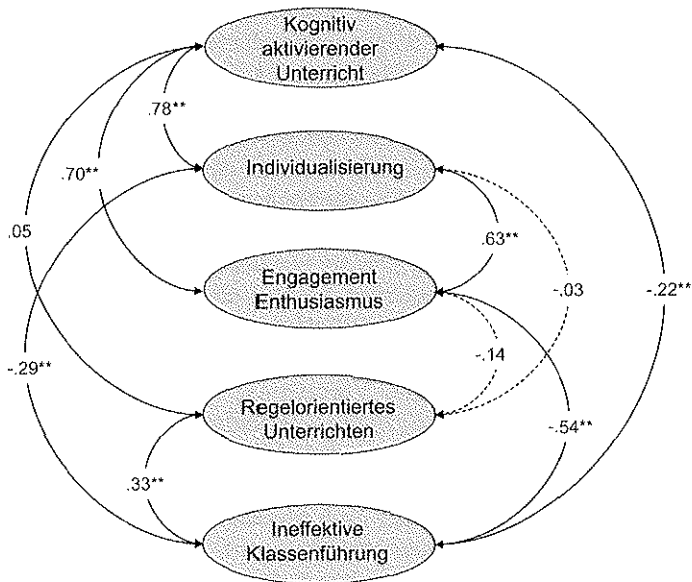


Abb. 6: 5-Faktorenmodell der Unterrichtswahrnehmung durch Mathematiklehrkräfte

In einem weiteren Analyseschritt wurden die Wahrnehmungen von Lehrkräften unterschiedlicher Schulformen verglichen. Hierzu wurden für jede der fünf Dimensionen z-standardisierte Faktorwerte berechnet (Abbildung 7). Im Einklang mit den Befunden von Baumert et al. (2004) berichten Hauptschullehrkräfte deutlich mehr Individualisie-

ung als Realschul- und Gymnasiallehrkräfte. Gleichzeitig aber erhalten sie auch höhere Werte für regelorientiertes Unterrichten und ineffektive Klassenführung und niedrigere Werte für Engagement und Enthusiasmus. Dies spricht für ein Unterrichtsklima, das durch Störungen belastet ist, das eigene Engagement reduziert und zu regelorientiertem Unterricht provoziert. Die Werte für kognitive Aktivierung hingegen sind für Lehrkräfte aller drei Schularten nahezu identisch (in den PISA-Erhebungen waren die Werte für kognitiv aktivierenden Unterricht bei Gymnasiallehrkräften höher als bei Lehrkräften der Haupt- und Realschule, Baumert et al., 2004; allerdings war der Bezugspunkt dort der Mathematikunterricht der Jahrgangsstufe 9).

Es ist geplant, auf analoge Weise die Dimensionalität der Unterrichtswahrnehmungen von Schülerinnen und Schülern zu analysieren (vgl. Ditton, 2002, Gruehn, 2000). Da die Lehrer- und Schülerskalen von PALMA weitgehend parallelisiert sind, wird es dann möglich sein, Lehrer- und Schülerwahrnehmungen des Mathematikunterrichts direkt miteinander zu vergleichen. Für den nächsten Projektzeitraum sind entsprechende Analysen auch für die Daten der Jahrgangsstufen 9 und 10 geplant. Schließlich wird der gesamte Satz dimensionaler Analysen die Variablenbasis liefern, um Veränderungen des Mathematikunterrichts aus Lehrer- und Schülerperspektive über die Jahrgangsstufen hinweg vertieft zu analysieren und die Einflüsse von kumulativ gleichbleibendem oder sich änderndem Unterricht auf die Kompetenz- und Merkmalsentwicklung von Schülerinnen und Schülern im Fach Mathematik zu untersuchen.

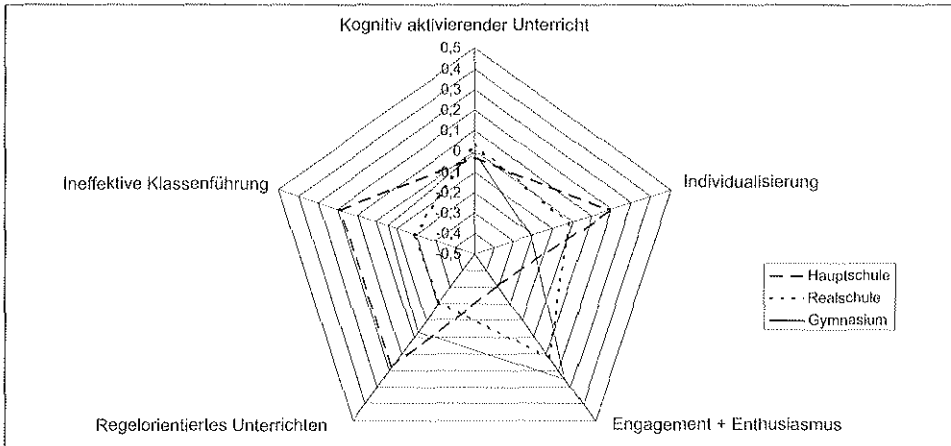


Abb. 7: Basisdimensionen der Unterrichtswahrnehmung von Mathematiklehrkräften nach Schulform (z-standardisierte Faktorwerte)

4.4 Elternhaus, Emotionen und Mathematikleistung

Neben Unterricht und Schulklasse ist auch das Elternhaus als wesentlicher Kontext der schulbezogenen Kompetenzentwicklung anzusehen. Zu den sozialen Disparitäten von Kompetenzentwicklung und Bildungsbeteiligung, zur Schichtabhängigkeit von elterlichem Engagement und zu den wechselseitigen Beziehungen zwischen Elternengagement und Schülerentwicklung wurden erste Analysen durchgeführt. Im Hintergrund stehen die sozialkognitiven Annahmen der Kontroll-Wert-Theorie zur Rolle elterlichen Verhaltens bei der Mediation von Schichteinflüssen und dem Aufbau von Kompetenzen und affektiven Merkmalen auf der Schülerseite (Jullien, 2006; Pekrun, in press).

4.4.1 Soziale Disparitäten von Bildungsbeteiligung und Kompetenzen

Die Befunde von PISA 2000 und PISA 2003 zeigen, dass Deutschland zu den Ländern mit auffallend hohen sozialen Disparitäten von Indikatoren der schulischen Bildung zählt. Die PISA-Befunde beziehen sich auf Kompetenzen und Bildungsbeteiligung bei den 15-Jährigen. Die Längsschnittdaten von PALMA machen es möglich, das Ausmaß sozialer Disparitäten auch für die zeitlich vorgelagerten Jahrgangsstufen der Sekundarstufe I zu analysieren und die Entwicklung der Disparitäten über die Schuljahre hinweg zu untersuchen. Die Befunde zeigen, dass Schichtzugehörigkeit und Schülerkompetenzen in Mathematik in signifikant negativen Zusammenhängen stehen. Ferner zeigt sich, dass die Chancen eines Gymnasialbesuchs für Kinder aus Elternhäusern der oberen Dienstklasse auch nach Kontrolle von kognitiven Grundfähigkeiten und mathematischer Kompetenz bereits in der Jahrgangsstufe 5 der vorliegenden bayerischen Stichprobe um das 2.6-fache höher liegen als bei Kindern aus Facharbeiterfamilien (Tabelle 6). Darüber hinaus ergeben sich Hinweise, dass das Ausmaß der sozialen Disparitäten von der fünften zur siebten Jahrgangsstufe zunimmt. Die Analysen des nächsten Projektzeitraums (Jahrgangsstufen 9 und 10) werden es möglich machen, diese Entwicklung über den gesamten Zeitraum der Sekundarstufe I hinweg zu verfolgen. Dabei wird es auch möglich sein, die Daten zu schulischen Bildungslaufbahnen und sozialen Kontexten zu verwenden, um Gründe für eine Zunahme sozialer Disparitäten (z.B. schichtkorreliertes Verlassen des Gymnasiums aufgrund schichtabhängig unzureichender familiärer Stützung bei Leistungsproblemen) zu analysieren. Ferner wird es möglich sein, die Entwicklung der Disparitäten zu entsprechenden Befunden von PISA 2006 für 15-Jährige bzw. die Jahrgangsstufe 9 in Beziehung zu setzen.

Tab. 6: Relative Chancen des Gymnasialbesuchs in Abhängigkeit von der Sozialschichtzugehörigkeit (Verhältnisse der Beteiligungschancen [odd ratios]; Referenz: Jugendliche aus Facharbeiterfamilien)

Sozialschicht der Bezugsperson im Haushalt		5. Jahrgangsstufe			7. Jahrgangsstufe		
		Modell I	Modell II	Modell III	Modell I	Modell II	Modell III
I	Obere Dienstklasse	3.3	2.8	2.6	4.3	3.3	3.1
II	Untere Dienstklasse	1.6	1.4	1.4	2.1	1.5	1.4
III	Routinedienstleistungen	1.0	1.0	1.1	0.9	0.9	1.0
IVa-d	Selbständige aus manuellen Berufen einschl. Landwirte	1.2	1.0	1.0	1.5	1.3	1.1
V-VI	Facharbeiter und leitende Arbeiter	1.0	1.0	1.0	1.0	1.0	1.0
VII	Un- und angelernte Arbeiter, Landarbeiter	0.6	0.7	0.8	0.6	0.7	0.7

Anmerkungen: Modell I entspricht dem unkontrollierten Modell, Modell II kontrolliert kognitive Grundfähigkeit, Modell III kontrolliert kognitive Grundfähigkeit und Mathematikkompetenz. $N_{5, Klasse} = 2.009$, $N_{7, Klasse} = 2.317$.

4.4.2 Schichtabhängigkeit von elterlichem Engagement

Zu den möglichen Mediatoren von Einflüssen der familiären Schichtzugehörigkeit auf die Kompetenzentwicklung von Schülerinnen und Schülern zählen der Umgang von Eltern mit den Leistungsanforderungen der Schule und die Kooperation des Elternhauses mit der Schule (z.B. Hanafin & Lynch, 2002). Erste Analysen zeigen, dass der elterliche Umgang mit Lernverhalten und Leistungen ihrer Kinder im Fach Mathematik tatsächlich partiell schichtabhängig ist. So berichten Kinder von Facharbeitern sowie un- und angelernten Arbeitern bzw. Landarbeitern im Vergleich zu Kindern aus Familien der Dienstklassen von höherem Leistungsdruck und weniger elterlicher Unterstützung in Mathematik. Kongruent hierzu fallen auch die Werte für – mutmaßlich durch elterlichen Leistungsdruck beeinflusste – Angst im Fach Mathematik höher aus (Jullien, 2006). Dabei zeigte sich in Strukturgleichungsanalysen mit Mehr-Gruppen-Vergleichen, dass zwar die Variablenausprägungen schichtabhängig variieren, die strukturellen Beziehungen zwischen elterlichem Verhalten und Schüleremotionen aber über Schichtgruppen hinweg gleichförmig sind. Dies bestätigt Annahmen zur relativen Universalität der Bedingungsbeziehungen von Lern- und Leistungsemotionen (Pekrun, in press).

4.4.3 Elterliches Engagement und Mathematikemotionen von Schülerinnen und Schülern: Cross-lagged-Analysen zu Wechselwirkungen

In der Kontroll-Wert-Theorie zu Leistungsemotionen (Pekrun, 2000; in press) wird angenommen, dass die Sozialumwelten von Schülerinnen und Schülern Einfluss auf ihre Emotionen nehmen, diese aber ihrerseits auf die Interaktionen in sozialen Umwelten zurückwirken. In ersten Analysen wurde diese Annahme für schülerperzipierte Variablen des Umgangs der Familie mit Mathematik einerseits und Schüleremotionen andererseits geprüft, wobei der Zeitraum von der 5. bis zur 6. Jahrgangsstufe einbezogen werden konnte (Jullien, 2006; Jullien & Pekrun, 2005). Hierzu dienten kompetitive Strukturgleichungsanalysen alternativer Bedingungsmodelle. Die Befunde belegen deutliche synchrone Zusammenhänge von Variablen des perzipierten Elternengagements mit den Mathematikemotionen von Schülerinnen und Schülern in der Jahrgangsstufe 5, zu denen über die Jahrgangsstufen hinweg diachrone Effekte treten, die für bidirektionale Bedingungsbeziehungen sprechen.

Auffällig ist bei diesen Analysen, dass die Effekte familiärer Variablen auf Schüleremotionen für die Zeitintervalle von der 5. bis 6. Jahrgangsstufe und von der 6. zur 7. Jahrgangsstufe jeweils etwa gleich stark sind, während deutlichere Effekte der umgekehrten Wirkrichtung erst im zweiten Zeitintervall nachweisbar sind. Möglicherweise spricht dies dafür, dass Elternwirkungen auf Schüleremotionen zunächst im Vordergrund stehen und Rückwirkungen erst mit zunehmendem Alter eine Rolle spielen. Zu prüfen wird sein, ob sich das Befundmuster über die nachfolgenden Messzeitpunkte hinweg fortsetzt und welche Rolle dem schüler- und elternperzipierten Elternverhalten – über die fachbezogene Emotionsentwicklung hinaus – für die Entwicklung mathematischer Kompetenzen zukommt.

5 Produkte für die Praxis

Die Methoden, die in diesem Projekt entwickelt worden sind, und die Befunde der Projekterhebungen können für die pädagogische Praxis im Fach Mathematik nutzbar gemacht werden. Wir haben deshalb damit begonnen, auf der Basis der Skalenerwicklungen und Längsschnittbefunde des Projekts Produkte für die Praxis zu entwickeln, die für eine solche Nutzung verwendet werden können. Dabei handelt es sich im Wesentlichen um die folgenden vier Bereiche (Überblick in Tabelle 7).

Tab. 7: Stand der Entwicklung von Produkten für die Praxis

Produkte	Stand der Fertigstellung
(1) <i>Unterrichtsmaterial: Umsetzung von PALMA-Befunden in Lehrbücher und Aufgabensammlungen</i>	
Mathematik heute, Band 5 (Hrsg.: H. Griesel, H. Postel, R. vom Hofe)	Braunschweig: Schroedel, 2004
Mathematik heute, Band 6 (Hrsg.: H. Griesel, H. Postel, R. vom Hofe)	Braunschweig: Schroedel, 2005
Grundbildung durch Grundvorstellungen – Aufgaben zur mathematischen Kompetenzentwicklung, Teil I, Klasse 5–7 (R. vom Hofe, S. Wartha, T. Hafner, M. Kleine)	Braunschweig: Schroedel, in Druck (erscheint Frühjahr 2007)
Grundbildung durch Grundvorstellungen – Aufgaben zur mathematischen Kompetenzentwicklung, Teil II, Klasse 8–10 (R. vom Hofe, S. Wartha, T. Hafner, M. Kleine)	Fertigstellung nach Abschluss der Längsschnitterhebungen
Spiele und Aufgaben zur Bruchrechnung; in: <i>Mathematik Lehren</i> , Heft 128 (S. Wartha & R. vom Hofe)	Velber: Friedrich, 2005
(2) <i>Expertisen für die Lehrplanentwicklung</i>	
Empfehlungen für die Lehrplanentwicklung 1. Teil (Klasse 5–7)	Fertigstellung 2007
Empfehlungen für die Lehrplanentwicklung 2. Teil (Klasse 8–10)	Fertigstellung nach Abschluss der Längsschnitterhebungen
(3) <i>Materialien für die Lehrerbildung</i>	
Freude wecken – Ängste nehmen (Hrsg.: T. Götz & M. Kleine); Themenheft der Zeitschrift <i>Mathematik Lehren</i>	Velber: Friedrich, 2006
Materialien zur Analyse von Schülerstrategien, Teil 1 Bruchrechnung (Hrsg.: R. vom Hofe)	Erscheinen 2006, zunächst als Hochschulschrift
Materialien zur Analyse von Schülerstrategien, Teil 2 und 3 (Rationale Zahlen; Funktionen und Sachrechnen)	Werden derzeit entwickelt
(4) <i>Test- und Evaluationsinstrumente</i>	
(a) Regensburger Mathematikleistungstest für 5.–10. Klassen	Fertigstellung nach Abschluss der Längsschnitterhebungen
(b) Münchener Skalen zu Mathematikemotionen (<i>Achievement Emotions Questionnaire – Mathematics</i> , AEQ-M; Pekrun, Götz & Frenzel, 2005)	
(c) Skalen zur Unterrichtsevaluation in Mathematik	

- 1) **Unterrichtsmaterial:** Das zentrale inhaltliche Thema im Fach Mathematik der Jahrgangsstufen 5 bis 7 ist die Erweiterung der Zahlbereiche von den natürlichen Zahlen (5. Klasse) zu den Bruchzahlen (6. Klasse) und schließlich zu den negativen bzw. rationalen Zahlen (7. Klasse). In den Längsschnitterhebungen des Projekts wurden Lösungsstrategien, Grundvorstellungen und Fehlvorstellungen zu diesen Inhalten identifiziert. Die Befunde zu diesen Strategien und Vorstellungen haben Eingang in das bundesweit verwendbare Lehrbuch *Mathematik heute* (Band 5 und 6) gefunden. Darüber hinaus wurde eine Aufgabensammlung für die Jahrgangsstufen 5 bis 7 mit didaktisch kommentierten, grundbildungsrelevanten Aufgaben erstellt, die gezielt die Ausbildung von Grundvorstellungen fördern (diese Sammlung wird im Frühjahr 2007 erscheinen).
- 2) **Lehrplan- und Curriculumentwicklung:** Es ist geplant, auf der Basis der Längsschnittbefunde eine zweiteilige Expertise für die Lehrplan- und Curriculumentwicklung zu erstellen. Die längsschnittlichen Analysen für den ersten Teil der Expertise (Jahrgangsstufen 5 bis 7) sind weitgehend fertig gestellt, sodass dieser Teil der Expertise in Kürze verfasst werden kann. Der zweite Teil der Expertise (Jahrgangsstufen 8 bis 10) ist für den Zeitraum nach Beendigung des Längsschnitts geplant.
- 3) **Materialien für die Lehrerbildung:** Parallel zu der Erstellung von Unterrichtsmaterialien haben wir mit der Entwicklung von Materialien für die Lehrerbildung begonnen. Dabei handelt es sich zum einen um Materialien zur mathematischen Modellierung von Sachsituationen (in den Themengebieten „Bruchrechnen“ und „rationale Zahlen“). Der erste Teil dieser Materialien wurde bereits mehrfach bei Lehrerfortbildungen (z.B. im Rahmen von SINUS-Weiterbildungen) erprobt und erscheint 2007, der zweite Teil ist in Vorbereitung. Zum anderen handelt es sich um Materialien zur Gestaltung eines emotions- und kompetenzfördernden Mathematikunterrichts. Hierzu wurde ein Themenheft der Zeitschrift *Mathematik lehren* gestaltet, das im Frühjahr 2006 erschienen ist. Auch zu diesem Themenbereich sind weitere Materialien in Vorbereitung.
- 4) **Leistungs- und Emotionsdiagnostik, Unterrichtsvaluation:** Es wurde mit der Analyse und Validierung des Regensburger Mathematikleistungstests für 5. bis 10. Klassen begonnen, der in den Voruntersuchungen des Projekts entwickelt wurde (s. Abschnitt 2.1 sowie vom Hofe et al., 2002). Ferner wurde auch mit der Validierung der im Projekt entwickelten Skalen zu Mathematikemotionen (Pekrun, Götz & Frenzel, 2005) und zum Mathematikunterricht begonnen. Diese Arbeiten sollen so fortgesetzt werden, dass die Instrumente anschließend in der schulbezogenen Praxis eingesetzt werden können. Beendet können diese Arbeiten allerdings erst dann werden, wenn die Daten aus allen Erhebungen des Längsschnitts vorliegen und der Abschluss des Längsschnitts die Freigabe von Testitems zulässt. Publierte Fassungen der Instrumente werden daher erst nach Beendigung des Längsschnitts zur Verfügung stehen.

6 Schlussfolgerungen

Ziel dieses Projekts ist es, Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von Mathematikleistungen bei Schülerinnen und Schülern der Sekundarstufe I zu untersuchen. Damit wird gleichzeitig eine Aufklärung der Entwicklungen und Bedingungen angestrebt, die entsprechenden Schülerleistungen in den

OECD-Studien PISA zugrunde liegen. Zu diesem Zweck wird eine Längsschnittuntersuchung durchgeführt, die Stichproben aus einer Kohorte von Schülerinnen und Schülern der Sekundarstufe I von den Jahrgangsstufen 5 bis 10 untersucht. Daten liegen bisher aus den Erhebungen der ersten vier Messzeitpunkte (Jahrgangsstufen 5 bis 8) vor.

Die Befunde zeigen eine Zunahme mathematischer Kompetenzen von der 5. zur 8. Jahrgangsstufe, wobei kalkülorientierte Kompetenzen über alle Jahrgangsstufen hinweg deutlicher ausgeprägt zu sein scheinen als – für mathematische Bildung im Sinne einer *mathematical literacy* besonders wesentliche – Modellierungskompetenzen. Dahinter aber verbirgt sich je nach Kompetenzbereich, Schülergruppe und Einzelschule eine Fülle von spezifischen Entwicklungsverläufen. Die Analysen belegen, dass die Schwierigkeiten vieler Schülerinnen und Schüler mit dem Stoff der bisher untersuchten Jahrgangsstufen auf Probleme im Bereich mathematischer Grundvorstellungen zurückgeführt werden können. Eine über die Jahre hinweg zunehmend größere Zahl von Schülerinnen und Schülern ist einer „Risikogruppe“ zuzuordnen, der die Beherrschung auch einfacher Aufgabenstellungen des Lehrplans der jeweiligen Jahrgangsstufe nicht gelingt. Kongruent hierzu findet sich eine im Schülerdurchschnitt eher ungünstige Entwicklung von Emotionen, Interesse und Lernverhalten im Fach Mathematik. Mögliche Gründe für diese Probleme liegen in der Abnahme von wahrgenommenem kognitivem Anregungsgehalt und Motivierungsqualität des Mathematikunterrichts im Untersuchungszeitraum.

Die Resultate von ersten längsschnittlichen Strukturgleichungs- und Mehr-Ebenen-Analysen sprechen dafür, dass Entwicklungen von Kompetenzen, affektiven Schülermerkmalen und Kontextvariablen in Mathematik über die Jahre hinweg in wechselseitigen Bedingungsbeziehungen stehen. Dabei konnte im Berichtszeitraum international erstmalig gezeigt werden, dass individuelle Schülerleistungen und das Leistungsniveau der Schulklasse gegenläufige Effekte auf die Emotionsentwicklung ausüben (günstige Wirkungen hoher Individualleistung, ungünstige Wirkungen eines hohen Leistungsniveaus der Klasse). Ferner belegen die Befunde einen Trend zu über die Jahrgangsstufen hinweg zunehmenden sozialen Disparitäten von Kompetenzen und Bildungsbeteiligung. In einem nächsten Analyseschritt sollen mögliche Gründe für diese Entwicklung untersucht werden.

Auf der Basis der Befunde wurde mit der Entwicklung von Produkten für die pädagogische Praxis begonnen. Hierzu zählen Materialien für den Mathematikunterricht, Expertisen zur Lehrplan- und Curriculumentwicklung, Kurse und Materialien für die Lehrerbildung sowie Instrumente zur Leistungs- und Emotionsdiagnostik und zur Evaluation von Mathematikunterricht.

Literatur

- Anderson, E.B. (1973). A goodness to fit test for the rasch-model. *Psychometrika*, 38, 123-140.
- 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., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). *Mathematikunterricht aus der Sicht der PISA-Schülerinnen und -Schüler und ihrer*

- Lehrkräfte. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 314-354). Münster: Waxmann.
- Blum, W. & Hofe, R. vom (2003). Welche Grundvorstellungen stecken in der Aufgabe? *Mathematik lehren*, 118, 14-18.
- Blum, W., Hofe, R. vom, Jordan, A. & Kleine, M. (2004). Grundvorstellungen als aufgabenanalytisches und diagnostisches Instrument bei PISA. In M. Neubrand (Hrsg.), *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland. Vertiefende Analysen im Rahmen von PISA 2000* (S. 145-157). Wiesbaden: VS-Verlag.
- Blum, W., Kaiser, G., Burghes, D. & Green, N. (1994). Entwicklung und Erprobung eines Tests zur „mathematischen Leistungsfähigkeit“ deutscher und englischer Lernender in der Sekundarstufe I. *Journal für Mathematik-Didaktik*, 15, 149-168.
- Blum, W., Neubrand, M., Ehmke, T., Senkbeil, M., Jordan, A., Ullfig, F. & Carstensen, C.H. (2004). Mathematische Kompetenz. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. xavier-Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 47-92). Münster: Waxmann.
- Brophy, J. (1999). *Teaching*. Brussels: International Academy of Education.
- Ditton, H. (2002). Lehrkräfte und Unterricht aus Schülersicht: Ergebnisse einer Untersuchung im Fach Mathematik. *Zeitschrift für Pädagogik*, 48, 262-286.
- Fischbein, E., Tirosh, D., Stavy, R. & Oster, A. (1990). The autonomy of mental models. *For the Learning of Mathematics*, 10, 23-30.
- Frenzel, A.C., Jullien, S. & Pekrun, R. (2006). „Thomas hat 60 Euro gespart ...“ oder „ $1/4x + 60 = x$ “: Freude und Angst beim Bearbeiten von Text- und Rechenaufgaben. *Mathematik lehren*, 52, 57-59.
- Götz, T., Pekrun, R., Zirngibl, A., Jullien, S., Kleine, M., Hofe, R. vom & Blum, W. (2004). Leistung und emotionales Erleben im Fach Mathematik: Längsschnittliche Mehrebenenanalysen. *Zeitschrift für Pädagogische Psychologie*, 18, 201-212.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen: Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Hanafin, J. & Lynch, A. (2002). Peripheral voices: Parental involvement, social class, and educational disadvantage. *British Journal of Sociology of Education*, 23, 35-49.
- Heller, K. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+ R)*. Göttingen: Beltz Test GmbH.
- Helmke, A. (2003). *Unterrichtsqualität erfassen, bewerten, verbessern*. Seelze: Kallmeyer.
- Hofe, R. vom (1995). *Grundvorstellungen mathematischer Inhalte*. Heidelberg: Spektrum.
- Hofe, R. vom, & Wartha, S. (2005). Grundvorstellungen als Fehlerquelle bei der Bruchrechnung. In H.-W. Henn & G. Kaiser (Hrsg.), *Mathematikunterricht im Spannungsfeld von Evolution und Evaluation – Festschrift für Werner Blum* (S. 202-211). Hildesheim: Franzbecker.
- Hofe, R. vom, Kleine, M., Pekrun, R. & Blum, W. (2005). Zur Entwicklung mathematischer Grundbildung in der Sekundarstufe I – theoretische, empirische und diagnostische Aspekte. In M. Hasselhorn (Hrsg.), *Jahrbuch für pädagogisch-psychologische Diagnostik. Tests und Trends* (S. 263-292). Göttingen: Hogrefe.
- Hofe, R. vom, Pekrun, R., Kleine, M. & Götz, T. (2002). Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Konstruktion des Regensburger Mathematikleistungstests für 5.-10. Klassen. In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 83-100). Weinheim: Beltz. (= *Zeitschrift für Pädagogik*, 45. Beiheft)
- Jöreskog, K. & Sörbom, D. (2002). *LISREL 8.53* (Update December 2003). Chicago: Scientific Software International.

- Jullien, S. (2006). *Elterliches Engagement und Lern- und Leistungsemotionen*. München: Utz.
- Jullien, S. & Pekrun, R. (2005, August). *Parents' influence on students' academic emotions*. Paper presented at the 11th Biennial Conference of the European Association for Research on Learning and Instruction, Nicosia, Cyprus.
- Kleine, M. (2004). *Quantitative Erfassung von mathematischen Leistungsverläufen in der Sekundarstufe I – methodische Grundlagen, Testkonstruktion und Testentwicklung*. Hildesheim: Franzbecker.
- Köller, O., Watermann, R., Trautwein, U. & Lüdtke, O. (Hrsg.). (2003). *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemeinbildenden und beruflichen Gymnasien*. Opladen: Leske und Budrich.
- Kuechemann, D. & Hoyles, C. (2003, February). *The quantity of students' explanations on a non-standard geometry item*. Paper presented at the Third Conference of the European Society for Research in Mathematics, Belaria, Italy.
- Langeheine, R. & Davier, M. von (1996). Bootstrap-Verfahren: Die „Kunst“, sich am Schopf aus dem Sumpf zu ziehen. *Newsletter der Fachgruppe Methoden*, 4, 5-7.
- Lehmann, R.H., Husfeldt, V. & Peek, R. (2001). Lernstände und Lernentwicklungen im Fach Mathematik – Ergebnisse der Hamburger Untersuchung (LAU) in den Jahrgangsstufe 5 und 6. In G. Kaiser, N. Knoche, D. Lind & W. Zillmer (Hrsg.), *Leistungsvergleiche im Mathematikunterricht* (S. 29-50). Hildesheim: Franzbecker.
- Marsh, H.W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280-295.
- Marsh, H.W. & Hau, K.T. (2002). Multilevel modeling of longitudinal growth and change: Substantive effects of regression toward the mean artifacts? *Multivariate Behavioral Research*, 37, 245-282.
- Max-Planck-Institut für Bildungsforschung (MPIB). (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Pajares, F. & Miller, M.D. (1995). Mathematics self-efficacy and mathematics performance: The need for specificity of assessment. *Journal of Counselling Psychology*, 42, 190-198.
- Pekrun, R. (1992). Kognitionen und Emotionen in studienbezogenen Lern- und Leistungssituationen: Explorative Analysen. *Unterrichtswissenschaft*, 20, 308-324.
- Pekrun, R. (2000). A social cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 143-163). Oxford, UK: Elsevier.
- Pekrun, R. (in press). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*.
- Pekrun, R. & Götz, T. (2005, August). *Classroom environment, academic achievement, and students' emotions: Multi-level implications of control-value theory*. Paper presented at the 11th Biennial Conference of the European Association for Research on Learning and Instruction, Nicosia, Cyprus.
- Pekrun, R., Götz, T. & Frenzel, A. C. (2005). *Achievement Emotions Questionnaire – Mathematics (AEQ – M) – User's Manual*. Department of Psychology, University of Munich.
- Pekrun, R., Götz, T., Hofe, R. vom, Blum, W., Jullien, S., Zirngibl, A., Kleine, M., Wartha, S. & Jordan, A. (2004). Emotionen und Leistung im Fach Mathematik: Ziele und erste Befunde aus dem „Projekt zur Analyse der Leistungsentwicklung in Mathematik“ (PALMA). In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 345-363). Münster: Waxmann.

- Pekrun, R., Götz, T., Titz, W. & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative und qualitative research. *Educational Psychologist*, 37, 91-106.
- Pekrun, R., Götz, T., Jullien, S., Zirngibl, A., Hofe, R. vom & Blum, W. (2002). *Skalenhandbuch PALMA 1. Messzeitpunkt (5. Jahrgangsstufe)*. Department Psychologie, Universität München.
- Pekrun, R., Götz, T., Jullien, S., Zirngibl, A., Hofe, R. vom & Blum, W. (2003). *Skalenhandbuch PALMA: 2. Messzeitpunkt (6. Jahrgangsstufe)*. Department Psychologie, Universität München.
- Pekrun, R., Götz, T., Jullien, S., Zirngibl, A., Hofe, R. vom & Blum, W. (2004). *Skalenhandbuch PALMA: 3. Messzeitpunkt (7. Jahrgangsstufe)*. Department Psychologie, Universität München.
- Pekrun, R., Hofe, R. vom & Blum, W. (2003). *Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Arbeitsbericht zur zweiten Projektphase (4/2002–3/2004)*. Department für Psychologie, Universität München.
- Pekrun, R., Hofe, R. vom & Blum, W. (2006). *Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA): Arbeitsbericht zur dritten Projektphase (4/2004–3/2006)*. Department für Psychologie, Universität München.
- Pekrun, R., Jullien, S., Lichtenfeld, S., Frenzel, A., Götz, T., Hofe, R. vom & Blum, W. (2005). *Skalenhandbuch PALMA: 4. Messzeitpunkt (8. Jahrgangsstufe)*. Department Psychologie, Universität München.
- Pekrun, R., Jullien, S., Zirngibl, A., Hofe, R. vom & Perry, R.P. (2004, April). *Emotions, self-regulated learning, and academic achievement: Testing a model of reciprocal causation*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. & Schiefele, U. (Hrsg.). (2004). *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Rost, D.H. (Hrsg.) (2000). *Hochbegabte und hochleistende Jugendliche*. Münster: Waxmann.
- Wartha, S. (2005). Fehler in der Bruchrechnung durch Grundvorstellungsumbrüche. In G. Graumann (Hrsg.), *Beiträge zum Mathematikunterricht 2005* (S. 593-596). Hildesheim: Franzbecker.
- Wartha, S. & Hofe, R. vom (2005). Probleme bei Anwendungsaufgaben in der Bruchrechnung. *Mathematik lehren*, 128, 10-17.
- Watson, J. & Kelly, B. (2004). Statistical variation in a chance setting. A two-year-study. *Educational Studies in Mathematics*, 57, 121-144.
- Zeidner, M. (1998). *Test Anxiety. The State of the Art*. New York: Plenum Press.
- Zirngibl, A., Götz, T., Pekrun, R., Hofe, R. vom & Kleine, M. (2005, März). *Aufgabenspezifische Erfassung von Mathematikemotionen*. Vortrag auf der 66. Tagung der Arbeitsgemeinschaft für Empirische Pädagogische Forschung (AEPF). Berlin.



BIQUA

Manfred Prenzel & Lars Allolio-Näcke (Hrsg.)

Untersuchungen zur Bildungs- qualität von Schule

Abschlussbericht des DFG-Schwerpunktprogramms

WAXMANN