

TIMSS/III–Deutschland
Der Abschlussbericht

Zusammenfassung ausgewählter Ergebnisse
der Dritten Internationalen Mathematik- und
Naturwissenschaftsstudie zur mathematischen
und naturwissenschaftlichen Bildung am Ende
der Schullaufbahn

*Jürgen Baumert, Wilfried Bos, Jens Brockmann, Sabine Gruehn,
Eckhard Klieme, Olaf Köller, Rainer Lehmann, Manfred Lehrke,
Johanna Neubrand, Kai Uwe Schnabel, Knut Schwippert
und Rainer Watermann*

Berlin
November 2000

| | |
|--|----|
| 1. Einleitung | 5 |
| 2. Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS/III) | 8 |
| 3. Mathematische und naturwissenschaftliche Grund-bildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung von Kompetenzen | 15 |
| 4. Mathematisch-naturwissenschaftliche Grundbildung im internationalen Vergleich | 23 |
| 5. Mathematische und naturwissenschaftliche Grundbildung beim Übergang von der Schule in den Beruf | 29 |
| 6. Mathematisch-naturwissenschaftliche Grundbildung ausländischer Schulausbildungsabsolventen und schichtenspezifische Einflüsse beim Übergang auf die Sekundarstufe II | 36 |
| 7. Fachleistungen im voruniversitären Mathematik- und Physikunterricht: Theoretische Grundlagen, Kompetenzstufen und Unterrichtsschwerpunkte | 38 |
| 8. Fachleistungen im voruniversitären Mathematik- und Physikunterricht im internationalen Vergleich | 49 |
| 9. Kurswahlen, Motivation und selbstreguliertes Lernen im Mathematik- und Physikunterricht der gymnasialen Oberstufe | 54 |
| 10. Soziale Vergleichsprozesse im Kurssystem der gymnasialen Oberstufe | 62 |
| 11. Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht..... | 64 |
| 12. Unterrichtsqualität und Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe | 67 |
| 13. Institutionelle und regionale Variabilität und die Sicherung gemeinsamer Standards in der gymnasialen Oberstufe | 73 |
| 14. Studienfachwünsche und Berufsorientierungen in der gymnasialen Oberstufe | 81 |
| 15. Geschlechtsdifferenzen in den mathematisch- naturwissenschaftlichen Leistungen | 83 |
| 16. Eine Nachbemerkung zum internationalen Vergleich | 85 |

1. Einleitung

Diese Zusammenfassung gibt Auskunft über ausgewählte Ergebnisse jener Teiluntersuchungen der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS), die sich aus analytischer Perspektive mit der mathematisch-naturwissenschaftlichen Grundbildung der nachwachsenden Generation am Ende der Pflichtschulzeit und dem Mathematik- und Physikunterricht voruniversitärer Bildungsgänge beschäftigen. Abbildung 1 gibt einen Überblick über das gesamte Forschungsdesign und die Datenquellen von TIMSS-Germany. In der Waagerechten ist die Testchronologie wiedergegeben, in der Senkrechten die getesteten Klassenstufen. Die grauen Felder repräsentieren die internationalen Vorgaben, die weißen Felder die nationalen – deutschen – Ergänzungen. Insbesondere letztere erlauben eine Reihe von Analysen, die über die Darstellung der reinen Testergebnisse weit hinausgehen. An der Erhebung in der Grundschule (TIMSS/I) nahm Deutschland nicht teil. Die Ergebnisse der Mittelstufenstudie (TIMSS/II) wurden veröffentlicht in:

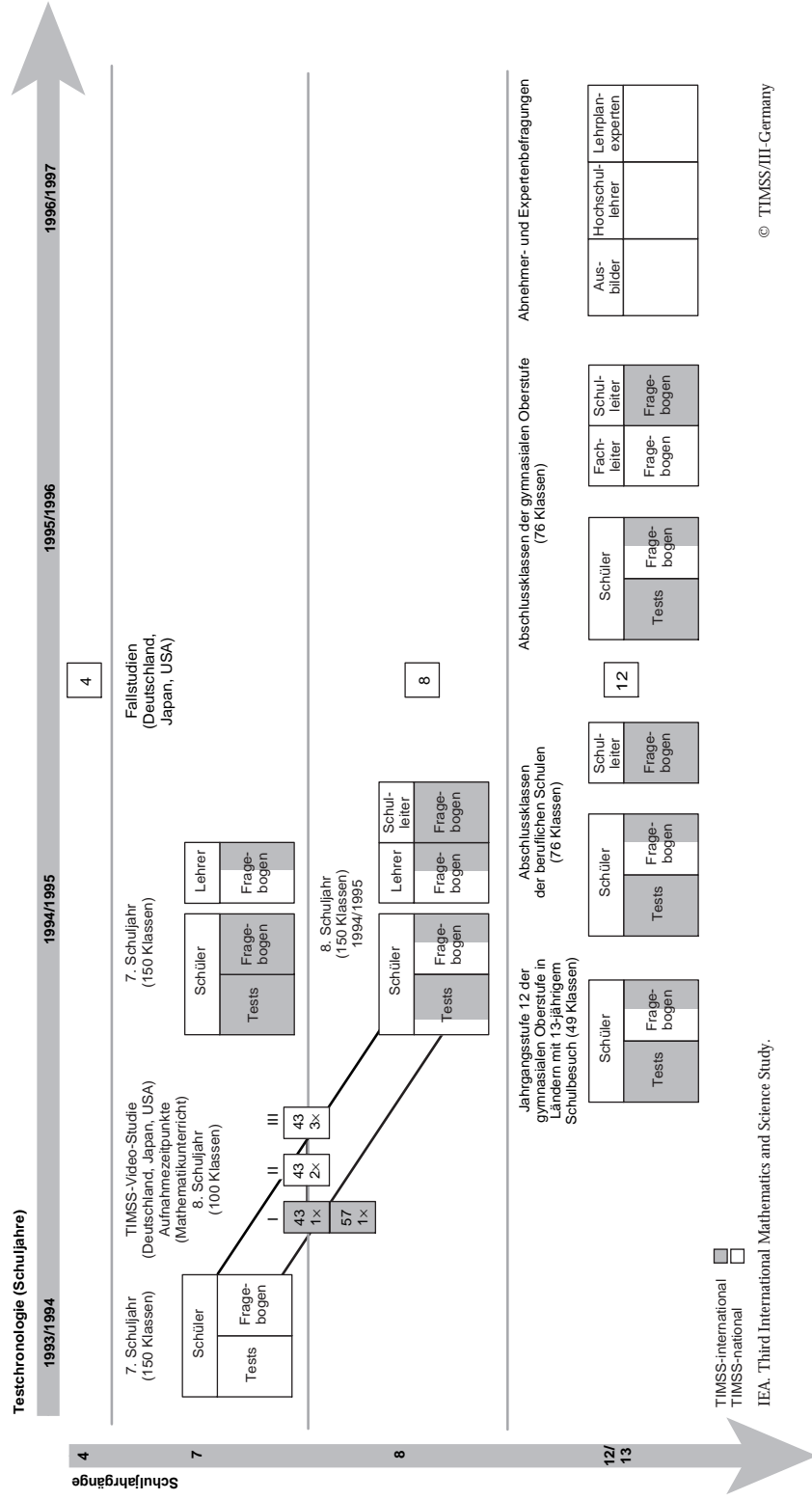
Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller O. & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske + Budrich.

Hier wird nun über die Oberstufenstudie (TIMSS/III) berichtet. Die vollständigen Ergebnisse finden sich in:

Baumert, J., Bos, W. & Lehmann, R. (Hrsg.). (2000). *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*, Bd. 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit; Bd. 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske + Budrich.

TIMSS/III ist im Unterschied zur deutschen Mittelstufenstudie, die abweichend vom internationalen Design längsschnittlich angelegt war, auch in Deutschland eine Querschnittsuntersuchung (vgl. Abb. 1). Dies schränkt die Analyse- und Aussagemöglichkeiten je nach Fragestellung in unterschiedlicher Weise ein. Während Struktur und Niveau der erworbenen mathematisch-naturwissenschaftlichen Kompetenzen sehr wohl auch in einer Survey-Studie untersucht werden können, erreichen Zusammenhangsanalysen fast immer nur explorativen Status. Sie bedürfen der Absicherung und Stützung durch Befunde aus Längsschnittstudien und experimentellen Arbeiten. Trotz dieser Einschränkungen, die wir in unseren Analysen immer

Abbildung 1: Untersuchungsplan von TIMSS-Germany: Mathematisch-naturwissenschaftliche Bildung in der Sekundarstufe I und am Ende der Schullaufbahn



wieder herausstellen, sind Untersuchungen wie TIMSS alles andere als überflüssig: Häufig genug widerlegen sie vermeintliche pädagogische oder bildungspolitische Gewissheiten und erfüllen so bereits eine wichtige diagnostische Aufgabe; darüber hinaus können sie – und TIMSS ist dafür ein gutes Beispiel – sowohl Fragen zuspitzen und auf den Punkt bringen als auch neue Perspektiven, nicht zuletzt für konstruktive pädagogische Phantasien, eröffnen.

Der unseres Erachtens wohl wichtigste bisherige Beitrag von TIMSS liegt in der Neustrukturierung der öffentlichen und professionellen Aufmerksamkeit. Nach TIMSS finden Bildungsthemen größeres Interesse, der Unterricht selbst ist als Kernaufgabe der Schule in das Zentrum gerückt. Dieser Aufmerksamkeitswandel ist nicht folgenlos geblieben. Auf politischer Ebene ist damit begonnen worden, Gräben einzuebrennen und sich gemeinsam der Aufgabe der Qualitätsentwicklung und Qualitätssicherung zu widmen. Auf wissenschaftlicher Ebene ist anwendungsbezogene Anschlussforschung mit einem neuen Schwerpunktprogramm der Deutschen Forschungsgemeinschaft (DFG) zur Qualität von Schule und Unterricht auf den Weg gebracht. Sowohl in der Erziehungswissenschaft als auch in den Fachdidaktiken hat TIMSS der empirischen Forschung und der Entwicklung innovativer Lehr- und Lernformen Anregungen gegeben. Auf praktischer Ebene – und hier liegt der entscheidende Prüfstein – sind die Wirkungen am vielfältigsten. In der Aus- und Weiterbildung von Lehrern haben die Fragestellungen und Befunde von TIMSS starken Widerhall gefunden. Prozesse und Inhalte des kumulativen fachlichen Lernens, der intelligente Umgang mit Aufgaben und die Verbindung von motivierenden und kognitiv aktivierenden Unterrichtsformen gehören zu den Themen, die Aufmerksamkeit gefunden haben. In fast allen Ländern und länderübergreifend sind Initiativen begonnen worden, um die Rolle der Einzelschule als Ausgangspunkt und Zentrum der Qualitätsentwicklung zu stärken und die Leistungen der die Schule unterstützenden Einrichtungen zu verbessern. Mit dieser Stoßrichtung gehen die Maßnahmen, auch wenn sie zunächst dem mathematisch-naturwissenschaftlichen Unterricht gewidmet sind, weit über diesen hinaus. In der Konzentration auf den Unterricht in der einzelnen Schule versprechen die Initiativen auch Nachhaltigkeit, da ein großes Entwicklungspotential unserer Schulen in der fachlichen Qualifikation der Lehrkörper und der schulinternen Zusammenarbeit liegt.

Dennoch wollen wir nicht übersehen, dass die breite Resonanz von TIMSS auch Kosten hatte. Sie liegen vor allem in einer allzu großen Simplifizierung komplexer Befunde und der interessengeleiteten Interpretation von Ergebnissen, die sich unabhängig von politischen Orientierungen nachweisen lässt und sich keineswegs auf die öffentliche Rezeption beschränkt, sondern auch in der Pädagogik zu finden ist. Nicht selten vermischen sich berechtigte fachliche Fragen und kritische Einwände

mit der persönlichen Mission, sodass eine adäquate Behandlung der Argumente nicht einfach ist. Im vorliegenden Bericht gehen wir auf kritische Fragen, die in der bisherigen Diskussion über TIMSS aufgeworfen worden sind, sorgfältig ein und prüfen diese auf ihren fachlichen Gehalt. Zugleich hoffen wir, dass sich unsere Analysen gegenüber vorschneller Vereinfachung und interessen geleiteter Einvernahme als sperrig erweisen. In der Regel sind die Befunde so vielschichtig, dass sie sich kaum dazu eignen, vorgefasste Meinungen zu bestätigen; eher werden sie Anlass geben, Problemwahrnehmungen neu zu strukturieren. Dass dies gelingen könnte, dafür gibt es gute Anzeichen. Denn in der Schulentwicklungsforschung besteht längst Einigkeit darüber, dass sich Qualitätsentwicklung und Qualitätssicherung ergänzen – und zwar auf allen Ebenen des Schulsystems. Studien, die wie TIMSS der Beobachtung systemischer Zusammenhänge und nicht unmittelbar der Steuerung der Einzelschule dienen, können ein Element der Schulentwicklung unter anderen sein, indem sie Orientierungswissen für Bildungsplanung, Lehrplanarbeit, Lehrerfortbildung, aber auch für die didaktische Arbeit am und im Unterricht selbst liefern.

2. Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS/III)

Definition der Untersuchungspopulationen

Die Dritte Internationale Mathematik- und Naturwissenschaftsstudie untersucht Schülerinnen und Schüler aus drei Altersgruppen, die sich in jeweils unterschiedlichen Phasen ihrer Schul- und Bildungslaufbahn befinden. Population I repräsentiert die Grundschule, Population II die Sekundarstufe I und Population III die Sekundarstufe II.

Während in der Grund- und Mittelstufe aufgrund des international standardisierten Aufbaus der Schulsysteme nach Jahrgangsstufen genaue Populationsdefinitionen gefunden werden konnten, mussten für die Population III breitere Rahmendefinitionen festgelegt werden, die jeweils national angemessene Präzisierungen erlaubten. Die Schulsysteme der an TIMSS teilnehmenden Staaten sind in der Sekundarstufe II zu unterschiedlich organisiert, als dass sich eine enge Definition hätte finden lassen. Nach internationaler Übereinkunft gehören zur Population III, deren mathematisch-naturwissenschaftliche Grundbildung untersucht werden soll, alle Personen, die sich zum Zeitpunkt der Erhebung im letzten Segment der Sekundarstufe in vollzeitlicher Ausbildung befinden (*Last Segment of Fulltime Secondary Education*).

Für Deutschland wurde diese Rahmenfestlegung folgendermaßen interpretiert: *Für den Untersuchungsbereich der mathematisch-naturwissenschaftlichen Grundbildung (Mathematics and Science Literacy) bilden die Population III alle Personen, die sich zum Zeitpunkt der Erhebung im letzten Jahr der vollzeitlichen Ausbildung in der Sekundarstufe II im allgemeinbildenden oder beruflichen Schulwesen befinden und die Sekundarstufe II zum ersten Mal durchlaufen.* Auszubildende im dualen System absolvieren nach dieser Definition eine vollzeitliche Ausbildung, da schulische und betriebliche Ausbildung als Teile einer einheitlichen Konzeption betrachtet werden.

16 Prozent eines Altersjahrgangs erreichen keinen Abschlussjahrgang eines Ausbildungsgangs; sie verlassen die Sekundarstufe II als Frühabgänger und treten ohne Abschluss in das Erwerbsleben ein.

Für den Untersuchungsbereich des voruniversitären Mathematikunterrichts gehören zur Zielpopulation alle Personen, die sich im Abschlussjahr der Sekundarstufe II befinden und im Schuljahr der Erhebung einen voruniversitären Mathematikkurs besuchen bzw. im Schuljahr vor der Erhebung besucht haben (*Advanced Mathematics*). *In Deutschland sind dies alle Schüler des Abschlussjahrgangs einer gymnasialen Oberstufe. Das Abschlussjahr kann je nach Landesregelung die 12. oder 13. Jahrgangsstufe sein.* Zum Kreis der gymnasialen Oberstufe zählen Oberstufen an Gymnasien, Fachgymnasien und Gesamtschulen, unabhängig von ihrer jeweiligen Trägerschaft.

Für den Untersuchungsbereich des voruniversitären Physikunterrichts wird die Zielpopulation analog definiert. *In Deutschland gehören zu ihr alle Schülerinnen und Schüler gymnasialer Oberstufen im zuvor erläuterten Sinne, die zum Untersuchungszeitpunkt einen Grund- oder Leistungskurs im Fach Physik besuchen.*

Teilnehmende Länder, Definition der nationalen Grundgesamtheiten und Stichprobenziehung

An der Mittelstufenuntersuchung (Population II) nahmen 45 Staaten teil. Das Teilnahmeinteresse war bei der Untersuchung der Population I und III geringer. An TIMSS/III beteiligten sich 24 Länder. Die meisten Teilnehmer waren west- und osteuropäische Staaten. Darüber hinaus nahmen die englischsprachigen Länder Australien, Kanada, Neuseeland und die USA teil. Weitere Teilnehmerstaaten waren Israel und Südafrika.

Nicht alle der 24 Länder nahmen an allen drei Teilkomponenten von TIMSS/III teil. Tabelle 1 gibt einen Überblick über die teilnehmenden Länder, differenziert nach

Tabelle 1: Teilnehmende Länder nach Untersuchungsbereichen

| Mathematisch-naturwissenschaftliche Grundbildung | Voruniversitärer Mathematikunterricht | Voruniversitärer Physikunterricht |
|--|---------------------------------------|-----------------------------------|
| Australien | Australien | Australien |
| Dänemark | Dänemark | Dänemark |
| Deutschland | Deutschland | Deutschland |
| Frankreich | Frankreich | Frankreich |
| Island | Griechenland | Griechenland |
| Israel | Israel | Israel |
| Italien | Italien | Italien |
| Kanada | Kanada | Kanada |
| Litauen | Litauen | Lettland |
| Neuseeland | Österreich | Norwegen |
| Niederlande | Russland | Österreich |
| Norwegen | Schweden | Russland |
| Österreich | Schweiz | Schweden |
| Russland | Slowenien | Schweiz |
| Schweden | Tschechien | Slowenien |
| Schweiz | USA | Tschechien |
| Slowenien | Zypern | USA |
| Südafrika | | Zypern |
| Tschechien | | |
| Ungarn | | |
| USA | | |
| Zypern | | |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Untersuchungsbereichen. Größte Aufmerksamkeit fand die Untersuchung der mathematisch-naturwissenschaftlichen Grundbildung am Ende der Sekundarstufe II.

Die international gewünschte Zielpopulation konnte nicht in allen Teilnehmerstaaten vollständig erreicht werden. Aus unterschiedlichen Gründen mussten jeweils Teile der gewünschten Population aus der Untersuchung ausgeschlossen werden. Dies war auch für Deutschland der Fall. Mit einem Ausschöpfungsgrad der national definierten Grundgesamtheit für den Untersuchungsbereich der mathematisch-naturwissenschaftlichen Grundbildung von 78,4 Prozent des einschlägigen Altersjahrgangs liegt die Bundesrepublik Deutschland im Rahmen der von einer Reihe anderer Teilnehmerländer getroffenen Definition ihrer jeweiligen nationalen Untersuchungspopulation. Tabelle 2 gibt differenziert nach Teilnehmerstaaten einen Überblick, welchen Anteil am Altersjahrgang die international gewünschte und national definierte Grundgesamtheit eines jeweiligen Landes abdeckt. Die Kennziffer, die den Anteil der international gewünschten Zielpopulation am relevanten Al-

Tabelle 2: Ausschöpfungsgrad der international gewünschten Zielpopulation und national definierten Untersuchungspopulation nach Ländern in Prozent des einschlägigen Altersjahrgangs¹ (*Age Cohort Coverage Efficiency*)

| Land | Ausschöpfungsgrad (<i>Coverage Efficiency</i>) | | Nationale Ausschlüsse |
|--------------------------|--|---|--|
| | International gewünschte Population (<i>Target Population Coverage Index</i>) | National definierte Population (<i>TIMSS Coverage Index</i>) | |
| Australien | 72,1 | 68,1 | Schüler in vollzeitlichen Berufsvorbereitungsklassen |
| Dänemark | 59,0 | 57,7 | |
| Deutschland ² | 83,6 | 78,4 | Schüler an Waldorfschulen, Kollegschulen, Berufs- und Fachakademien sowie an Schulen des Gesundheitswesens |
| Frankreich | 84,7 | 83,9 | |
| Griechenland | 66,8 | 10,0 | Schüler ohne voruniversitären Mathematik- und Physikunterricht |
| Island | 54,6 | 54,6 | |
| Israel ³ | • | • | |
| Italien ⁴ | 74,0 | 74,0 | |
| Kanada | 89,8 | 70,3 | Schüler in Ontario mit Abschlussexamen im Dezember 1995 |
| Litauen | 42,5 | 42,5 | Nicht litauisch sprechende Schüler und Schüler an Privatschulen |
| Neuseeland | 70,5 | 70,5 | |
| Niederlande ² | 92,0 | 78,0 | Schüler in beruflichen Kurzausbildungsgängen und in betrieblicher Ausbildung |
| Norwegen | 85,2 | 84,0 | |
| Österreich | 92,7 | 75,9 | Schüler in Bildungsgängen mit weniger als drei Jahren Dauer |
| Russland | 84,3 | 48,1 | Schüler an beruflichen Schulen und nicht russisch sprechende Schüler |
| Schweden | 70,6 | 70,6 | |
| Schweiz | 84,0 | 81,9 | Ausbildungsrichtungen, die mit weniger als fünf Schülern an einer Schule vertreten waren |
| Slowenien | 87,8 | 87,8 | |
| Südafrika | 48,9 | 48,9 | |
| Tschechien | 77,6 | 77,6 | |
| USA | 65,6 | 63,1 | Klassenwiederholer und nicht englischsprachige Schüler |
| Ungarn | 65,4 | 65,3 | |
| Zypern | 61,4 | 47,9 | Schüler an privaten und beruflichen Schulen |

¹ Durchschnittlicher Jahrgang der 15- bis 19-Jährigen 1995.

² Abweichungen vom internationalen Report (Mullis u.a., 1998) aufgrund korrigierter Populationsdaten in diesem Bericht.

³ Wegen unbefriedigender Populations-/Stichprobenstatistiken wurde für Israel kein Ausschöpfungsgrad berechnet.

⁴ In Italien haben sich vier Regionen nicht an TIMSS beteiligt, sodass in Italien die Grundgesamtheit (TPCI) eingeschränkt ist.

tersjahrgang angibt, bezeichnen wir als *Target Population Coverage Index* (TPCI). Die Kennziffer, die den entsprechenden Anteil der national definierten Grundgesamtheit anzeigt, wird *TIMSS Coverage Index* (TCI) genannt. Dieser Index ist die zentrale Referenzgröße der internationalen Publikationen. Bei deutlich abweichenden *TIMSS Coverage Indices* ist ein Vergleich der Befunde zwischen Staaten entweder überhaupt nicht, nur unter Vorbehalten und nach differenzierten Analysen der jeweils ausgeschlossenen Jahrgangsteile oder nach Normierung der Jahrgangsteile möglich. Dieser Vorbehalt ist für den gesamten internationalen Vergleich essentiell. *Im Unterschied zu den TIMSS-Untersuchungen in der Sekundarstufe I sind für die Population III keine direkten Ländervergleiche zulässig.* Man kann diesen Sachverhalt nicht nachdrücklich genug betonen, um Missinterpretationen vorzubeugen.

Besondere Bedeutung misst TIMSS der Untersuchung jener Teilpopulation von Schülerinnen und Schülern bei, die einen voruniversitären Mathematik- oder Physikkurs besuchen. In Deutschland zählen zu diesen „Mathematikspezialisten“ alle Schüler einer gymnasialen Oberstufe, da der durchgehende Besuch des Mathematikunterrichts – zumindest bis zum Ende der 12. Jahrgangsstufe – zur Erlangung der allgemeinen Hochschulreife obligatorisch ist. Zu den „Physikspezialisten“ zählen wir in Deutschland alle Schülerinnen und Schüler einer gymnasialen Oberstufe, die zum Erhebungszeitpunkt einen Grund- und/oder Leistungskurs im Fach Physik belegt hatten.

Im Schuljahr 1995/96 betrug der Anteil der Mathematikspezialisten, die eine Abschlussklasse an der Oberstufe eines Gymnasiums oder einer Gesamtschule besuchten, 25,3 Prozent der einschlägigen Alterskohorte. Der Anteil der Physikspezialisten am Altersjahrgang fällt infolge der Wahlmöglichkeit mit 9,3 Prozent erheblich niedriger aus.

Hinsichtlich des internationalen Leistungsvergleichs dieser Oberstufenschüler gelten die für die Gesamtpopulation beschriebenen Vorbehalte analog. Die Vergleichbarkeit der Befunde hängt maßgeblich davon ab, dass der voruniversitäre Mathematikunterricht ähnliche Anteile am Altersjahrgang erfasst, der Ausschöpfungsgrad also ähnlich ist. Je geringer der Anteil des einschlägigen Altersjahrgangs ist, der einen voruniversitären Mathematik- oder Physikunterricht besucht, desto positiver dürfte die jeweilige Teilpopulation ausgelesen sein. Tabelle 3 veranschaulicht die von Land zu Land sehr unterschiedliche Reichweite des erweiterten Mathematik- und Physikunterrichts.

Tabelle 3: Ausschöpfungsgrad der Teilpopulationen mit voruniversitärem Mathematik- bzw. Physikunterricht nach Ländern in Prozent des einschlägigen Altersjahrgangs¹ (*Mathematics and Physics TIMSS Coverage Indices* [MTCI, PTCI])

| Land | Ausschöpfungsgrad (<i>Coverage Efficiency</i>) | |
|--------------------------|--|--|
| | Personen in voruniversitären Mathematikprogrammen (MTCI) | Personen in voruniversitären Physikprogrammen (PTCI) |
| Australien | 15,7 | 12,6 |
| Dänemark | 20,6 | 3,2 |
| Deutschland ² | 25,3 | 9,3 |
| Frankreich | 19,9 | 19,9 |
| Griechenland | 10,0 | 10,0 |
| Italien | 14,1 | – |
| Kanada | 15,6 | 13,7 |
| Lettland | – | 3,0 |
| Litauen | 2,6 | – |
| Norwegen | – | 8,4 |
| Österreich | 33,3 | 33,1 |
| Russland | 2,1 | 1,8 |
| Schweden | 16,2 | 16,3 |
| Schweiz | 14,3 | 14,2 |
| Slowenien | 75,4 | 38,6 |
| Tschechien | 11,0 | 11,0 |
| USA | 13,7 | 14,5 |
| Zypern | 8,8 | 8,8 |

¹ Durchschnittlicher Jahrgang der 15- bis 19-Jährigen 1995.

² Abweichungen vom internationalen Report (Mullis u.a., 1998) aufgrund der Verwendung von Populations- anstelle von Stichprobendaten in diesem Bericht.

Stichprobe

Die realisierte Hauptstichprobe umfasst in Deutschland insgesamt 5.345 Schüler (Tab. 4 und 5). Diese Stichprobe bildet die Basis des internationalen Vergleichs. Die Stichprobengröße reicht jedoch nicht aus, um systematische Ländervergleiche in der Bundesrepublik Deutschland durchzuführen, da die Länder mit Ausnahme der neuen Länder nur proportional zu ihrer Größe in der Stichprobe vertreten sind. Dieser Stichprobenplan entspricht der internationalen Anlage der Studie, in der *nationale* Kennwerte ermittelt werden sollen. Die Untersuchung wurde nicht mit dem

Ziel systematischer Bundesländervergleiche geplant; wohl aber sollte der Vergleich größerer regionaler Einheiten – zum Beispiel neue und alte Bundesländer oder Länder mit Abiturprüfungen nach 12 oder 13 Schuljahren – möglich sein.

Tabelle 4: Realisierte Stichprobe nach Schulart und Erhebungsjahr in Deutschland (ungewichtet)

| Erhebungsjahr | Allgemeinbildende Schulen | | Berufliche Schulen | | | | Insgesamt |
|-----------------|---------------------------|---------------------------|--------------------|-----------------|-------------------|---------------|-----------|
| | Gymnasien | Integrierte Gesamtschulen | Fachgymnasien | Fachoberschulen | Berufsfachschulen | Berufsschulen | |
| 1995 | 2.582* | 73* | 49 | 111 | 375 | 931 | 4.121 |
| 1996 | 3.786 | 93 | – | – | – | – | 3.879 |
| Summe | 6.368 | 166 | 49 | 111 | 375 | 931 | 8.000 |
| Hauptstichprobe | 3.786 | 93 | 49 | 111 | 375 | 931 | 5.345 |

* Nur 12. Jahrgang in Bundesländern mit 13 Schuljahren.

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Tabelle 5: Realisierte Stichprobe der Abschlussklassen in der gymnasialen Oberstufe an Gymnasien, integrierten Gesamtschulen und Fachgymnasien nach Kurswahl in Mathematik und Physik in Deutschland (ungewichtet)

| Fach | Fach nach Einführungsphase oder früher abgegeben | Grundkurs nicht durchgehend belegt | Grundkurs durchgehend belegt | Leistungskurs | Keine Angaben | Insgesamt |
|------------|--|------------------------------------|------------------------------|---------------|---------------|-----------|
| Mathematik | – | 324 | 2.049 | 1.436 | 119 | 3.928 |
| Physik | 1.966 | 338 | 864 | 518 | 242 | 3.928 |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

3. Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung von Kompetenzen

Wichtigste Grundlage der TIMS-Studie sind die mathematischen und naturwissenschaftlichen Tests, die aufgrund umfangreicher theoretischer Vorarbeiten entwickelt wurden. Die Interpretation der Untersuchungsergebnisse beruht auf der Aussagekraft (Validität) der verwendeten Tests. Im ersten Band unseres vollständigen Berichts wird der Aufbau der beiden Grundbildungstests (mathematische Grundbildung und naturwissenschaftliche Grundbildung) beschrieben und ausführlich, in theoretischen Aussagen und empirischen Analysen, deren Validität belegt, während parallel im zweiten Band die Validität der beiden Fachleistungstests für den voruniversitären Mathematik- bzw. Physikunterricht beleuchtet wird. In dieser Zusammenfassung ausgewählter Ergebnisse gehen wir auf folgende Fragestellungen ein:

- Zunächst wird die inhaltliche Konzeption der TIMSS-Grundbildungstests erläutert. Hintergrund ist zum einen eine Sichtung von zentralen Stoffgebieten des Fachunterrichts, zum anderen die Diskussion um Grundbildung bzw. *Mathematics and Science Literacy*, die in der Erziehungswissenschaft, in den Fachdidaktiken und auch in der Bildungspolitik derzeit intensiv geführt wird. Tatsächlich liegt ein Verdienst der TIMS-Studie darin, der Diskussion um Ziele und Kriterien der mathematisch-naturwissenschaftlichen Grundbildung neue Impulse gegeben zu haben. Danach gehen wir kurz auf empirische Strukturanalysen ein, mit denen unter anderem geprüft wird, ob man tatsächlich die Anwendung von Kenntnissen und Fertigkeiten in Alltagssituationen als integralen Bestandteil der Grundbildung, wie sie TIMSS erfasst, ansehen kann.
- In einem zweiten Schritt wird die Aussagekraft der Aufgaben zur mathematischen Grundbildung näher untersucht. Zum einen prüfen wir, welche Anforderungsmerkmale verantwortlich sind für den Schwierigkeitsgrad der Testaufgaben. Zum zweiten legen wir bestimmte Kompetenzniveaus (*Proficiency Levels*) auf der TIMSS-Leistungsskala fest und prüfen, welche Aufgaben typischerweise auf den unterschiedlichen Kompetenzniveaus gelöst werden können. Beide Vorgehensweisen erlauben es, inhaltlich zu beschreiben, worin sich erfolgreichere Bearbeiter von weniger erfolgreichen unterscheiden. Nicht nur die Leistungsmenge, sondern die Qualität der Leistungen und der bewältigten Anforderungen wird in vier Kompetenzstufen beschrieben und durch Beispielaufgaben belegt.

- Abschließend behandeln wir analog Anforderungsmerkmale und Kompetenzstufen in der naturwissenschaftlichen Grundbildung.

Konzeption und Aufbau der Grundbildungstests

Die TIMSS-Tests zur mathematisch-naturwissenschaftlichen Grundbildung markieren einen Wendepunkt in der Konzipierung von Tests für internationale Vergleichsstudien, indem sie einen Kompromiss darstellen zwischen Lehrplanbindung einerseits und Orientierung am praktischen Alltagshandeln andererseits. Frühere Studien dieser Art waren primär an den Stoffen und Lernzielen der Mathematik bzw. der naturwissenschaftlichen Fächer orientiert. Dieser Ansatz kam für die Entwicklung der Grundbildungstests in TIMSS/III schon aufgrund der großen Heterogenität der Untersuchungspopulationen, die teilweise gar keinen regulären Mathematik- und Naturwissenschaftsunterricht mehr besuchen, nicht mehr in Frage. Er wurde daher verknüpft mit einem normativen Verständnis von *Mathematics and Science Literacy*. Leitend ist eine funktionale Sicht auf die Bewährung des in der Schule erworbenen Wissens in vielfältigen Anforderungssituationen. Dieser Wechsel von einer curricularen zu einer funktionalen Begründung der Testkonzepte ist inzwischen im OECD-Programm PISA weiter forciert worden.

In der internationalen Fachdiskussion werden mathematische und naturwissenschaftliche Kompetenzen – ebenso wie das Textverstehen – als basale Kulturwerkzeuge verstanden, die universelle Bedeutung haben für das Verständnis der natürlichen und sozialen Umwelt und für selbstständiges, verantwortliches Handeln in Beruf und Alltag. Allerdings erschöpft sich *Literacy* nicht in der Kenntnis und Nutzung einfacher Begriffe und Verfahren. Sie schließt vielmehr ein Verständnis zentraler theoretischer Konzepte und Arbeitsprinzipien sowie den kritischen Umgang mit mathematischen bzw. naturwissenschaftlichen Modellen ein. Dementsprechend werden mehrere Stufen von *Literacy* unterschieden, von einem anschaulichen Verständnis von Alltagsphänomenen über ein sinnvolles Anwenden elementarer fachlicher Modelle bis hin zur vollen Kommunikations- und Urteilsfähigkeit in diesen Bereichen.

Die Aufgaben der TIMSS-Grundbildungstests sollen nun – im Sinne dieses *Literacy*-Konzepts – prüfen, inwieweit zentrale Konzepte des mathematisch-naturwissenschaftlichen Unterrichts der Mittelstufe verstanden sind und in Alltagskontexten genutzt werden können. Als pragmatischer Ausgangspunkt der Testentwicklung diente die für TIMSS/II, die Mittelstufenstudie, entwickelte Ordnungsmatrix, die zum einen mehrere Sachgebiete unterscheidet und zum anderen beschreibt, welche Art

von Anforderungen beim Lösen der Aufgaben bewältigt werden müssen („Verhaltenserwartungen“). Fachwissenschaftler und Fachdidaktiker aus zehn Ländern entwickelten in mehreren Durchgängen gemeinsam mit Institutionen aus Kanada, Australien und den USA Sammlungen geeigneter Testaufgaben, die alle Bereiche der Matrix ausfüllen konnten und durchweg in Alltagskontexte eingebunden waren. Nach einer internationalen Felderprobung wurden Aufgaben mit schlechten Messeigenschaften, zum Beispiel mit einem *Bias* zu Gunsten einzelner Länder, ausgesondert. Die beiden nachfolgenden Tabellen geben an, wie sich die verbliebenen und in der Hauptuntersuchung genutzten Testaufgaben auf Sachgebiete und Verhaltenserwartungen verteilen.

Aufgaben, in denen lediglich Wissenselemente und Routinen abgeprüft werden, bilden in beiden Grundbildungstests die Minderheit. Überwiegend wird konzeptuelles Denken, Problemlösen und Anwenden gefordert.

Die empirische Prüfung der Testdaten ergab, dass es gerechtfertigt ist, von einer homogenen Leistungsdimension der mathematisch-naturwissenschaftlichen Grundbildung zu sprechen. Das Antwortformat (*Multiple Choice* oder offene Fragestellungen) kann dabei vernachlässigt werden. Die Anwendung von mathematischen und naturwissenschaftlichen Kenntnissen und Fertigkeiten im Alltag bildet – wie im *Literacy*-Konzept gefordert – einen integralen Bestandteil dieser Grundbildung. Neben der globalen Fähigkeitsdimension können – je nach Fragestellung – auch spezifische Skalen für die mathematische und die naturwissenschaftliche Grundbildung gebildet und unterschieden werden. Diese Unterscheidung zwischen den fachlichen

Tabelle 6: Mathematische Testaufgaben im Grundbildungstest nach Sachgebiet und Verhaltenserwartung

| Sachgebiet | Verhaltenserwartung | | | | Insgesamt |
|--|---------------------|-----------------------------------|--------------------------------------|---|-----------|
| | Wissen | Beherrschung von Routineverfahren | Beherrschung von komplexen Verfahren | Anwendungsbezogene Aufgaben und innermathematische Probleme | |
| Zahlen, Zahlverständnis | 1 | 10 | 2 | 7 | 20 |
| Algebraische Terme, Gleichungen, Graphen | 3 | 1 | – | 6 | 10 |
| Messen, Schätzen | 1 | 1 | 1 | 11 | 14 |
| Insgesamt | 5 | 12 | 3 | 24 | 44 |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Tabelle 7: Naturwissenschaftliche Testaufgaben im Grundbildungstest nach Sachgebiet und Verhaltenserwartung

| Sachgebiet | Verhaltenserwartung | | | | Insgesamt |
|-----------------------------|-----------------------------------|-----------------------------------|-------------------------------|---|-----------|
| | Verstehen von Einzelinformationen | Verstehen komplexer Informationen | Konzeptionalisieren, Anwenden | Experimentieren, Beherrschung von Verfahren | |
| Geowissenschaftliche Themen | – | 1 | 3 | – | 4 |
| Biologie | 5 | 3 | 1 | 2 | 11 |
| Physik/Chemie | 5 | 2 | 7 | 3 | 17 |
| Insgesamt | 10 | 6 | 11 | 5 | 32 |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Bereichen ist aus pädagogischer und schulpolitischer Sicht nützlich. Im Folgenden beschreiben wir daher Anforderungsmerkmale und Kompetenzstufen getrennt für die beiden Domänen.

Anforderungsmerkmale und Kompetenzstufen in der mathematischen Grundbildung

Die Kennwerte der individuellen Fähigkeit werden in TIMSS international auf einer Skala mit dem Mittelwert 500 und der Standardabweichung 100 angegeben. Auf dieser Skala sollen nun diskrete Niveaustufen unterschieden werden. Ziel ist es, den Mittelwert 500 sowie einige darüber oder darunter liegende Stufen (400, 600, 700) inhaltlich zu charakterisieren. Die Leitfrage hierfür lautet: Welche Aufgaben werden von Schülern, die beispielsweise den Fähigkeitskennwert 500 haben, mit ausreichender Sicherheit gelöst, nicht jedoch von Schülern auf der darunter liegenden Stufe (hier: 400)? („Ausreichende Sicherheit“ wird in TIMSS als Lösungschance von mindestens 2:1 bzw. Lösungswahrscheinlichkeit von mindestens 65 % definiert.)

Die nachfolgende Abbildung zeigt die TIMSS-Fähigkeitsskala und positioniert darauf vier Beispielaufgaben aus dem Test zur mathematischen Grundbildung. Diese Beispiele markieren die vier Kompetenzstufen. Gemeinsam mit anderen Markier-Items, die hier nicht im einzelnen wiedergegeben werden können, lassen sie folgende inhaltliche Interpretation der Kompetenzniveaus zu:

Stufe 1 (Score 400): Alltagsbezogene Schlussfolgerungen

Beispielaufgabe C6 ist die einzige Aufgabe, die bereits auf diesem untersten Kompetenzniveau mit ausreichender Sicherheit gelöst werden kann. Sie verlangt keinerlei explizit mathematisches Operieren, sondern nur die intuitive Überlegung: Je mehr Schritte jemand braucht, um eine bestimmte Entfernung zu überwinden, desto kleiner ist seine Schrittlänge.

Stufe 2 (Score 500): Anwendung von einfachen Routinen

Auf dieser Stufe können Aufgaben gelöst werden, die einfache Proportionalitätsüberlegungen, Prozentrechnungen (siehe Beispielaufgabe D13) oder Flächenberechnungen erfordern, die Standardstoff des Mathematikunterrichts der Mittelstufe sind.

Stufe 3 (Score 600): Bildung von Modellen und Verknüpfung von Operationen

Wer diese Fähigkeitsstufe erreicht, kann auch Aufgaben korrekt lösen, bei denen mehrere mathematische Operationen verknüpft werden müssen, im Beispiel B17 etwa Volumenberechnung und Verhältnisrechnung. Typisch ist auch, dass der mathematische Ansatz nicht im Aufgabentext selbst nahe gelegt wird, sondern erschlossen werden muss. Auf dieser Kompetenzstufe ist also mathematisches Modellieren möglich.

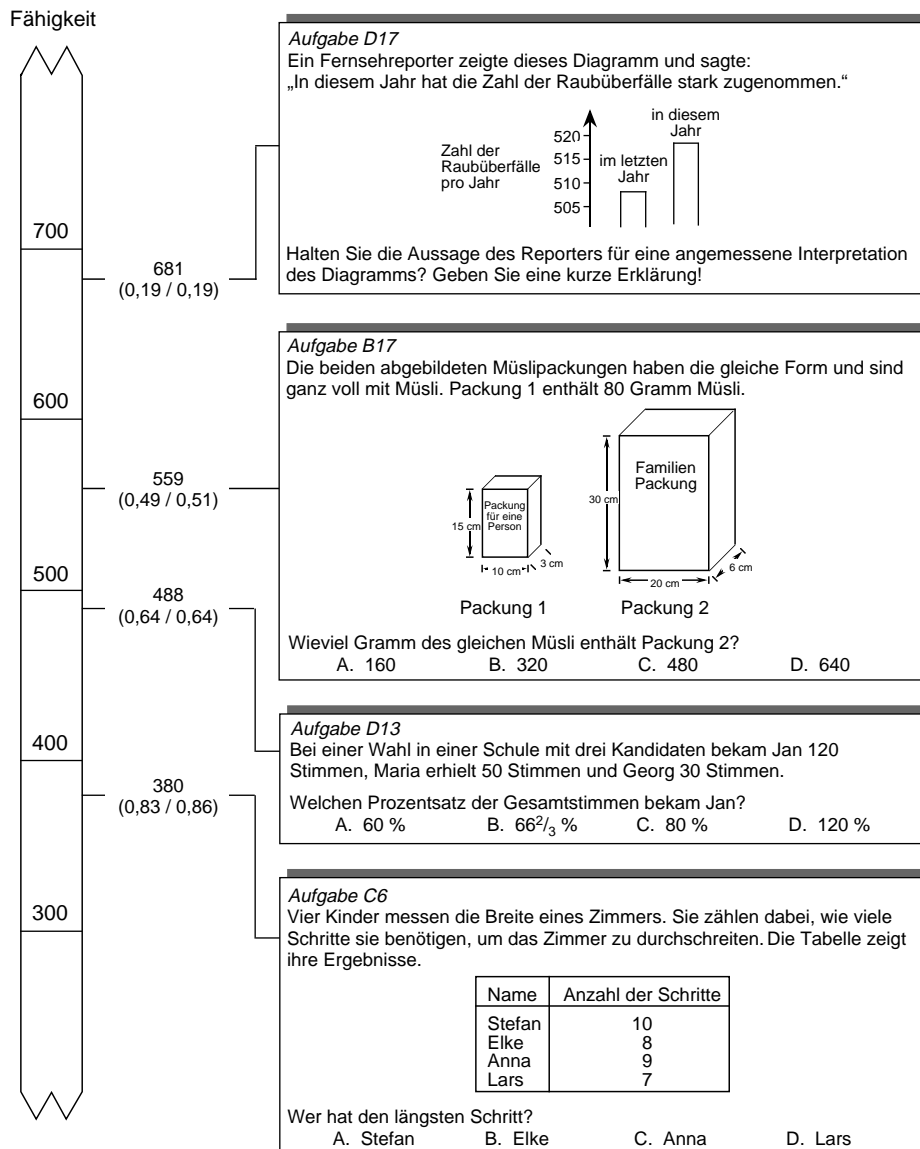
Stufe 4 (Score 700): Mathematisches Argumentieren

Wer diese höchste Stufe der mathematischen Grundbildung erreicht, vermag komplexe Diagramme zu erstellen und kritisch zu interpretieren. In der Beispielaufgabe D17 etwa muss der Bearbeiter erkennen, dass die dargestellte Zunahme der Zahl der Raubüberfälle nur etwa $1/50$ der Ausgangsbasis ausmacht und daher keineswegs als starke Zunahme interpretiert werden darf.

Diese empirisch gewonnene Abstufung von Kompetenzniveaus stimmt mit den Stufen von *Literacy* überein, die in der theoretischen Diskussion genannt wurden. Mathematische Grundbildung bewegt sich daher zwischen den Polen des erfahrungsgelunden Denkens einerseits, der kritischen Verwendung mathematischer Modelle andererseits.

Je weiter nach oben man auf der Kompetenzskala kommt, desto komplexer werden die mathematischen Ansätze (gemessen an der Zahl der Größen, die einbezogen werden müssen) und desto offener werden die Probleme, die noch bewältigt werden können. (Bei offenen Problemen muss man nicht nur eine Lösung ermitteln, sondern auch Lösungswege beschreiben oder vergleichen.) Diese beiden Aufgabenmerkmale konnten zusammen mit einer unabhängigen Experteneinschätzung des

Abbildung 2: Testaufgaben aus dem Bereich mathematischer Grundbildung – Beispiele nach Schwierigkeit



Die Werte an den Verbindungslinien zwischen den Beispielen und der Fähigkeitsssäule geben das für eine 65-prozentige Lösungswahrscheinlichkeit erforderliche Fähigkeitsniveau und die Werte in Klammern die relativen internationalen und deutschen Lösungshäufigkeiten an.

Literacy-Niveaus zwei Drittel der Schwierigkeitsvarianz erklären. Dieser Befund und die oben dargestellte Kompetenzstufung belegen, dass der TIMSS-Test zur mathematischen Grundbildung dem Anspruch der *Literacy*-Konzeption gerecht wird.

Anforderungsmerkmale und Kompetenzstufen in der naturwissenschaftlichen Grundbildung

Auch im naturwissenschaftlichen Bereich wurden vier Kompetenzstufen unterschieden. In der nachfolgenden Abbildung sind vier Beispielaufgaben aus der Biologie aufgeführt. Im Verbund mit weiteren Markieraufgaben, auch aus der Physik und der Biologie, die teilweise im Bericht abgedruckt sind, lassen diese Beispiele folgende Interpretation der Kompetenzstufen im Bereich der naturwissenschaftlichen Grundbildung zu:

Stufe 1 (Score 400): Naturwissenschaftliches Alltagswissen

Die Wahl der richtigen Antwortalternative in Beispielaufgabe B2 ist schon möglich, wenn man weiß, dass gesunde Ernährung etwas mit Vitaminen zu tun hat. Derartige Wissen setzt noch keinen systematischen Schulunterricht voraus.

Stufe 2 (Score 500): Erklärung einfacher alltagsnaher Phänomene

Diese zweite Kompetenzstufe markiert den internationalen Durchschnitt naturwissenschaftlicher Grundbildung. Hier muss man beispielsweise wissen, dass Infektionskrankheiten durch Übertragung von Krankheitserregern verursacht werden, und man muss dieses Wissen einsetzen, um alltagsnahe Phänomene wie eine Grippeerkrankung zu erklären (vgl. Beispielaufgabe D3).

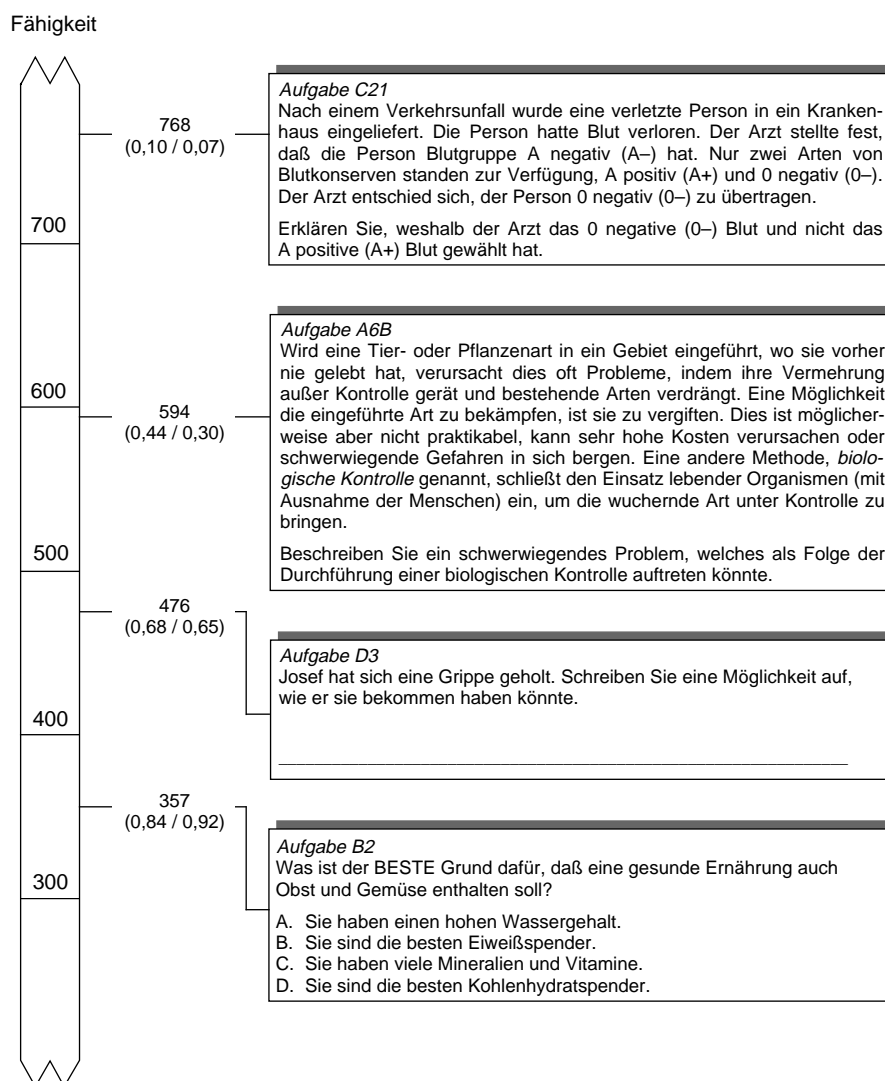
Stufe 3 (Score 600): Anwendung elementarer naturwissenschaftlicher Modellvorstellungen

Die Beispielaufgabe A6B erfordert ein Verständnis für Vorgänge in einem Ökosystem: Die Einführung eines neuen Organismus in ein solches Ökosystem kann unerwünschte Nebenfolgen haben, wie zum Beispiel die übermäßige Vermehrung des neuen Organismus, wenn dieser keine natürlichen Feinde besitzt. Auf dieser Kompetenzstufe werden demnach Modellvorstellungen angesprochen, die der naturwissenschaftliche Unterricht der Mittelstufe vermittelt.

Stufe 4 (Score 700): Verfügung über grundlegende naturwissenschaftliche Fachkenntnisse

Diese höchste Stufe der naturwissenschaftlichen Grundbildung setzt korrektes Fachwissen voraus, beispielsweise über Prozesse der Fotosynthese, über physika-

Abbildung 3: Testaufgaben aus dem Bereich naturwissenschaftlicher Grundbildung – Beispiele für Biologie nach Schwierigkeit



Die Werte an den Verbindungslinien zwischen den Beispielen und der Fähigkeitsäule geben das für eine 65-prozentige Lösungswahrscheinlichkeit erforderliche Fähigkeitsniveau und die Werte in Klammern die relativen internationalen und deutschen Lösungshäufigkeiten an.

liche Energiekonzepte oder (in Beispielaufgabe C21) über Blutgruppen im menschlichen Organismus.

Auch hier werden die Probleme, die noch bewältigt werden können, umso offener, je weiter nach oben man auf der Kompetenzskala kommt. Gemeinsam mit einer unabhängigen Einstufung des *Literacy*-Niveaus erklärt dieses Aufgabenmerkmal gut die Hälfte der Schwierigkeitsvarianz. Dieser Befund und die beschriebene Stufung der Kompetenzniveaus machen wiederum deutlich, dass der TIMSS-Test das theoretische Konzept naturwissenschaftlicher Literalität angemessen abbildet.

Weitere Belege für die Validität der Grundbildungstests liefert die Befragung von Experten der beruflichen Bildung, über die in Kapitel V des ersten dieser Zusammenfassungen zu Grunde liegenden Bandes berichtet wird. Berufsschullehrer sowie Vertreter von Industrie- und Handelskammern ordneten drei Viertel der mathematischen und etwa die Hälfte der naturwissenschaftlichen Aufgaben als „wichtig“ oder „sehr wichtig“ für die berufliche Ausbildung ein, wobei je nach Berufsfeld unterschiedliche Schwerpunkte gebildet werden.

4. Mathematisch-naturwissenschaftliche Grundbildung im internationalen Vergleich

Fragestellungen

Dieses Berichtskapitel verfolgt nicht die Absicht, die bekannten deskriptiven Befunde zum internationalen Vergleich noch einmal in extenso auszubreiten. Die Ergebnisse werden nur knapp als Hintergrundinformation zusammengefasst und durch einen Bericht zu spezifischen Anwendungsaufgaben ergänzt. Im Mittelpunkt stehen vielmehr drei analytische Fragestellungen:

- In einem ersten Schritt wird der Zusammenhang zwischen Haltekraft der Sekundarstufe II, Strukturmerkmalen von Bildungssystemen und erreichtem mathematisch-naturwissenschaftlichem Grundbildungsniveau untersucht. Diese Teilstudie geht von der Annahme aus, dass in der Beschulungsquote der Sekundarstufe II – unabhängig von der wirtschaftlichen Situation eines Landes und den organisatorischen Merkmalen des Bildungssystems – in gewisser Weise die gesellschaftlich und politisch geteilte Wertschätzung formalisierter Bildung auch unterhalb des akademischen Niveaus zum Ausdruck kommt.

- Nach einer Diskussion über die methodischen Voraussetzungen des internationalen Vergleichs im Rahmen von TIMSS/III und vor dem Hintergrund eines knappen Referats der wichtigsten Vergleichsergebnisse werden auf der Ebene ausgewählter mathematischer und naturwissenschaftlicher Aufgaben die spezifischen Stärken und Schwächen deutscher Schulabsolventen analysiert.
- In einem dritten Schritt schließlich werden die Befunde in Relation zu internationalen fachdidaktischen Gütemaßstäben und deutschen Lehrplanvorgaben interpretiert.

Haltekraft der Sekundarstufe II und Niveau der mathematisch-naturwissenschaftlichen Grundbildung

Die an TIMSS/III teilnehmenden Länder sind im Wesentlichen Industriestaaten. Umso überraschender ist die hohe Variabilität der Bildungsbeteiligung in der Sekundarstufe II. Auf der einen Seite stehen etwa Dänemark und Island mit einer Retentivitätsrate von unter 60 Prozent des durchschnittlichen Altersjahrgangs, auf der anderen Seite befinden sich die Niederlande oder Österreich mit einer Bildungsbeteiligung im Abschlussjahrgang von über 90 Prozent. In der Bundesrepublik beträgt – wie auch in der Schweiz – die Retentivitätsrate 84 Prozent eines durchschnittlichen Altersjahrgangs; 16 Prozent gehen in das Erwerbsleben, ohne das letzte Ausbildungsjahr erreicht zu haben.

Die Variabilität der Bildungsbeteiligung lässt sich nicht auf übliche strukturelle Systemmerkmale zurückführen. Die Retentivität variiert unabhängig von der Organisationsstruktur der Oberstufe, der Länge der gesetzlich vorgeschriebenen Schulpflicht, dem Ausbau der zur Hochschule führenden Bildungsgänge, aber auch unabhängig von Wohlstandsindikatoren wie dem Bruttonettoprodukt oder monetären Bildungsindikatoren wie dem relativen Anteil der Bildungsausgaben am Bruttonettoprodukt. In der Retentivitätsrate scheint unabhängig von wirtschaftlichen und organisatorischen Merkmalen so etwas wie die gesellschaftlich und politisch geteilte Wertschätzung formalisierter Bildung auch unterhalb des akademischen Niveaus und die damit korrespondierende Investitionsbereitschaft von Zeit und Anstrengung zum Ausdruck zu kommen.

In der unterschiedlichen Retentivität der Systeme liegt ein zentrales Untersuchungsproblem von TIMSS/III: Die Zielpopulationen können sich von Land zu Land unterscheiden. Denn das Ausschlusskriterium für Teile der Alterskohorte ist das Nichterreichen des Abschlussjahrgangs der Sekundarstufe II. Frühabgänger werden

also aus pragmatischen Gründen – sie sind in der Regel nicht mehr auffindbar – aus der Untersuchung ausgeschlossen. Dieses Ausschlusskriterium ist vom Leistungskriterium nicht unabhängig. Es sind im Allgemeinen die leistungsschwächsten Schüler, die vor Erreichen des Abschlussjahrs die Schule verlassen. Man darf davon ausgehen, dass mit abnehmender Retentivitätsrate die Selektivität der Untersuchungspopulation und damit auch der Stichprobe zunimmt.

Da mit sinkender Ausschöpfungsquote die Selektivität von Population und Stichprobe steigt, ist eine negative Korrelation zwischen Retentivitätsrate und Leistung im Grundbildungstest zu erwarten. Analoge Zusammenhänge sind auch in den früheren internationalen Vergleichsuntersuchungen gefunden worden, die sich auf den voruniversitären Mathematik- und Naturwissenschaftsunterricht beschränkten. Der bekannte negative Zusammenhang zwischen Retentivität und Leistung ist auch die Grundlage der Kritik an direkten Ländervergleichen ohne Berücksichtigung differentieller Ausschöpfungsquoten. Umso überraschender ist der empirische Befund der TIMSS/III-Grundbildungsuntersuchung, dass sich ein straffer positiver Zusammenhang zwischen Retentivitätsrate einerseits und Grundbildungsniveau andererseits nachweisen lässt. Die Korrelation zwischen Retentivität und mittlerem nationalen Grundbildungsniveau liegt bei $r = .56$. Betrachtet man nur die 25 Prozent der testleistungsstärksten Untersuchungsteilnehmer – eine Teilpopulation, für die das Argument differentieller Selektivität der Länderstichproben praktisch nicht mehr zutrifft –, so steigt die Korrelation auf $r = .72$.

Mit zunehmender Retentivität eines Bildungssystems steigt gleichzeitig die durchschnittliche Verweildauer der Schüler und Auszubildenden und der Anteil von Personen, die voruniversitäre Programme in den mathematisch-naturwissenschaftlichen Fächern besuchen. Beide Faktoren beeinflussen in spezifischer Weise das erreichbare Grundbildungsniveau positiv. Um die spezifische Bedeutung der Haltekraft eines Systems für das Grundbildungsniveau zu schätzen, müssen diese beiden Faktoren konstant gehalten werden. Dies ist durch die Kombination verschiedener regressionsanalytischer Modelle möglich.

Das Ergebnis ist eindeutig: Auch nach Kontrolle dieser Einflüsse zeigt sich ein substantieller spezifischer Effekt der Retentivität eines Bildungssystems. *Die Haltekraft in der Sekundarstufe II scheint ein Indikator für eine allgemeine positive Bildungseinstellung zu sein, die unterstützenden Einfluss auf das erreichbare Bildungsniveau ausübt. Dieser Befund unterstreicht die strategische Bedeutung, die das duale System im Rahmen der institutionalisierten Bildung in Deutschland besitzt.*

Stärken und Schwächen der deutschen Schulabsolventen

Die Befunde der multiplen Ländervergleiche auf der allgemeinen Ebene mathematisch-naturwissenschaftlicher Grundbildung sind konsistent. Das am Ende der schulischen und beruflichen Erstausbildung erreichte Niveau mathematisch-naturwissenschaftlicher Grundbildung liegt im internationalen Vergleich in einem mittleren Bereich. Schulabsolventen wichtiger europäischer Nachbarstaaten erreichen tendenziell oder deutlich bessere Leistungsergebnisse. Das Gesamtbild ist über alle verglichenen Teilpopulationen hinweg stabil. Auch bei der Betrachtung des oberen Leistungsviertels der Alterskohorte ergeben sich keine abweichenden Resultate. Der Leistungsrückstand der deutschen Schüler des Abschlussjahrgangs der Sekundarstufe II ist gegenüber der durch Schweden, die Niederlande, Norwegen und die Schweiz gebildeten Spitzengruppe substantiell. *Das Durchschnittsalter des Absolventenjahrgangs ist in Deutschland mit 19,5 Jahren relativ hoch. In den meisten TIMSS-Teilnehmerstaaten erreichen die Schul- und Ausbildungsabsolventen ein vergleichbares oder höheres mathematisch-naturwissenschaftliches Grundbildungsniveau in jüngerem Alter.*

Im differenzierten Vergleich der Besetzung von Fähigkeitsniveaus in Frankreich, den Niederlanden, Norwegen, der Schweiz und Deutschland zeigen sich die Besonderheiten der Leistungsverteilung der deutschen Schüler. Im Bereich der mathematischen und naturwissenschaftlichen Grundbildung sind das unterste Fähigkeitsniveau über- und die beiden obersten Fähigkeitsniveaus deutlich unterbesetzt.

Methodisch differenziertere Analysen auf der Ebene von Einzelaufgaben geben schließlich Auskunft über spezifische Stärken und Schwächen der deutschen Schulabsolventen. Die Ergebnisse der Analysen so genannter *differentieller Itemfunktionen* lassen sich knapp zusammenfassen: Deutsche Schulabsolventen sind – vermutlich unterrichts- und weniger lehrplanbedingt – tendenziell mit größeren Schwierigkeiten bei Aufgabenstellungen konfrontiert, die komplexe Operationen, die Anwendung mathematischer oder naturwissenschaftlicher Modellvorstellungen und selbstständiges fachliches Argumentieren verlangen. *Die relativen Stärken der deutschen Abschlussjahrgänge liegen eher bei der Lösung mathematischer Routineaufgaben und erfahrungsnaher naturwissenschaftlicher Aufgaben, die häufig im Rückgriff auf Alltagswissen und ohne entsprechenden Fachunterricht gelöst werden können.* Zwischen den Lehrplänen, die konzeptuelles Verständnis, die Fähigkeit, elementare Operationen zu verknüpfen, und den Transfer des Gelernten auf neue Zusammenhänge verlangen, und der Konkretisierung der Lehrplanvorschriften im Unterricht klafft eine Lücke. Im internationalen Vergleich scheint sie in den Naturwissenschaften größer als in der Mathematik zu sein.

Zwei Aufgabenbeispiele aus dem mathematischen Grundbildungstest verdeutlichen diesen Befund.

Abbildung 4: Mathematische Beispielaufgaben der Niveaustufe III und IV mit erhöhter relativer Lösungsschwierigkeit für deutsche Schülerinnen und Schüler im Vergleich zur Schweiz, zu Frankreich und Schweden

A12. Diese beiden Anzeigen sind in einer Zeitung in einem Land erschienen, in dem die Währungseinheit *zeds* ist.

| | |
|---|---|
| <p style="text-align: center;">GEBÄUDE A</p> <p style="text-align: center;">Büroräume zu vermieten</p> <p style="text-align: center;">85–95 Quadratmeter 475 <i>zeds</i> pro Monat</p> <p style="text-align: center;">100–120 Quadratmeter 800 <i>zeds</i> pro Monat</p> | <p style="text-align: center;">GEBÄUDE B</p> <p style="text-align: center;">Büroräume zu vermieten</p> <p style="text-align: center;">35–260 Quadratmeter 90 <i>zeds</i> pro Quadratmeter pro Jahr</p> |
|---|---|

Eine Firma ist daran interessiert, ein 110 Quadratmeter großes Büro in diesem Land für ein Jahr zu mieten. In welchem Bürogebäude, A oder B, sollte sie das Büro mieten, um den niedrigeren Preis zu bekommen? Wie rechnen Sie?

Relative Lösungshäufigkeit
International: .49 Deutschland: .45 Frankreich: .66 Schweden: .70 Schweiz: .66

B25. Wie aus der Skizze ersichtlich ist, verläuft ein Pfad diagonal durch das rechteckige Feld. Berechnen Sie die Fläche des Feldes OHNE den Pfad. Notieren Sie Ihren Lösungsweg!

Relative Lösungshäufigkeit
International: .38 Deutschland: .29 Frankreich: .44 Schweden: .51 Schweiz: .51

Mathematische und naturwissenschaftliche Literalität

In der internationalen mathematik- und naturwissenschaftsdidaktischen Literatur zeichnet sich eine gewisse Verständigung über Grundzüge einer wünschenswerten mathematisch-naturwissenschaftlichen Allgemein- und Grundbildung der nachwachsenden Generation ab. Die Hauptaspekte mathematischer und naturwissenschaftlicher Literalität lassen sich mit einer begrenzten Zahl von Kompetenzen beschreiben. Zur mathematischen Grundbildung gehören:

- die Fähigkeit, die Anwendbarkeit mathematischer Konzepte und Modelle auf alltägliche und komplexe Problemstellungen zu erkennen,
- die Fähigkeit, die einem Problem zu Grunde liegende mathematische Struktur zu sehen,
- die Fähigkeit, Aufgabenstellungen in geeignete Operationen zu übersetzen, und
- ausreichende Kenntnis und Beherrschung von Lösungsroutinen.

Die Grunddimensionen von *Scientific Literacy* sind danach:

- Vertrautheit mit der natürlichen Welt und Kenntnis ihrer Verschiedenheit und Einheit,
- Verständnis zentraler naturwissenschaftlicher Konzepte und Prinzipien,
- Kenntnis der Interdependenz von Naturwissenschaften und Technik,
- epistemologische Vorstellungen von der konstruktiven Natur der Naturwissenschaften sowie Kenntnis ihrer Stärken und Grenzen,
- Verständnis der Grundzüge naturwissenschaftlichen Denkens,
- Anwendung von naturwissenschaftlichem Wissen auf Sachverhalte des persönlichen und sozialen Lebens.

Wählt man die internationalen fachdidaktischen *Benchmarks* und die deutschen Lehrplanvorgaben bzw. die einheitlichen Standards des mittleren Abschlusses als Bezugspunkte für die Interpretation der Befunde des internationalen Leistungsvergleichs, zeigt sich, dass TIMSS/III wichtige und differenzierende Antworten auf offene Fragen der Grundbildungsdiskussion gibt. Zunächst wird deutlich, dass sowohl die internationalen als auch die deutschen Standards die Messlatte für eine wünschenswerte mathematisch-naturwissenschaftliche Grundbildung sehr hoch legen. In vielen Industriestaaten erreicht die Mehrzahl der jungen Erwachsenen im Abschlussjahrgang der Sekundarstufe II die vorgegebene Zielmarke nicht. *Auch in Deutschland gibt es eine erhebliche Diskrepanz zwischen den an den Fachunterricht gerichteten normativen Erwartungen und den tatsächlich erreichten Resultaten.* Die TIMSS-Befunde bestätigen in gewisser Weise jene Skeptiker, die generelle Zweifel an der Universalisierbarkeit einer mathematisch-naturwissenschaftlichen Grundbildung auf dem hohen angestrebten Niveau äußern. Gleichzeitig wird diese Skepsis

aber auch relativiert. *Denn die Resultate zeigen ebenfalls, dass es in einer ganzen Reihe europäischer Industriestaaten gelingt, für nahezu drei Viertel der nachwachsenden Generation mindestens das Niveau funktionaler Literalität und für etwa ein Drittel das Niveau konzeptuellen und prozeduralen Verständnisses zu erreichen und gleichzeitig den Anteil wirklich schwacher Lerner auf ein Minimum zu beschränken.* Diese beiden oberen Fähigkeitsstufen der mathematisch-naturwissenschaftlichen Grundbildung sind nicht durch Alltagserfahrungen, sondern ausschließlich durch systematischen Unterricht erreichbar.

Die Ergebnisse des internationalen Leistungsvergleichs haben selbstverständlich keine inhärente normative Funktion, auch wenn dies die Verfasser von Ranglisten immer wieder unterstellen möchten. Was in einem Bildungssystem mathematisch-naturwissenschaftliche Grundbildung ausmachen soll, ist eine bildungstheoretische und letztlich politische Frage, die durch empirische Forschung nicht beantwortet werden kann. Empirische Ergebnisse können aber sehr wohl ein Element der Realitätskontrolle im normativen Diskurs sein, indem sie auf Diskrepanzen zwischen Erwartungen und Ergebnissen aufmerksam machen und auf prinzipiell Erreichbares hinweisen.

5. Mathematische und naturwissenschaftliche Grundbildung beim Übergang von der Schule in den Beruf

Das Kapitel zur mathematisch-naturwissenschaftlichen Grundbildung beim Übergang von der Schule in den Beruf beschäftigt sich im Wesentlichen mit zwei Fragestellungen. Die beiden ersten Abschnitte liefern einen vorwiegend analytischen Beitrag zur mathematisch-naturwissenschaftlichen Grundbildung am Ende der Sekundarstufe II mit dem Schwerpunkt auf berufliche Bildungsgänge. Der dritte Abschnitt fokussiert motivationale Orientierungen gegenüber dem Ausbildungsberuf und fragt danach, welche Faktoren der betrieblichen Ausbildung die Entwicklung des Interesses und der Bindung an den Ausbildungsberuf begünstigen. Wir wollen die Zusammenfassung der Ergebnisse deshalb in zwei Schritten vornehmen, indem wir zunächst auf den Leistungs- und anschließend auf den motivationalen Aspekt eingehen.

Die Analysen zur mathematisch-naturwissenschaftlichen Grundbildung beim Übergang von der Schule in den Beruf orientieren sich an den folgenden Punkten:

- In einem ersten Schritt fassen wir die Befunde zu institutionellen und regionalen Leistungsunterschieden zusammen.

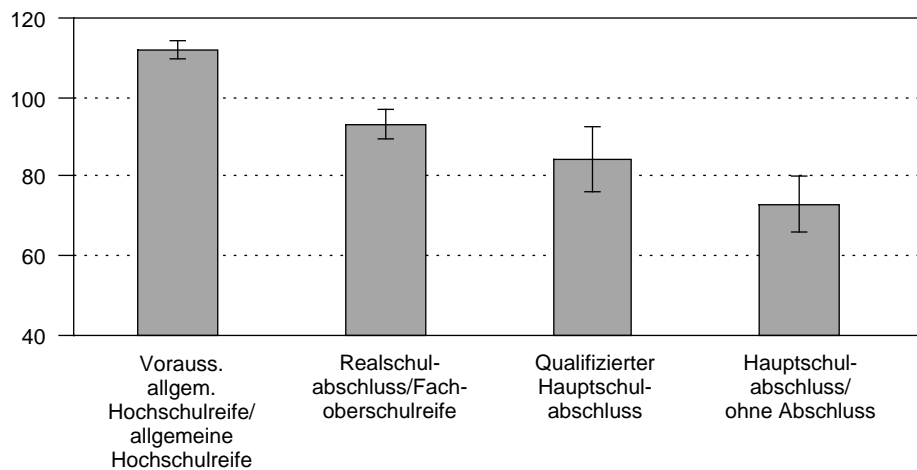
- In einem zweiten Schritt untersuchen wir die mathematisch-naturwissenschaftliche Grundbildung von Schülern nach der Sachgebietsnähe ihrer Berufsausbildung. Wir gehen davon aus, dass die Selbstselektion in den Ausbildungsberuf aufgrund fachspezifischer Interessen und selbstwahrgenommener Fähigkeiten sowie fachbezogenes berufliches Lernen systematische Unterschiede in den Leistungsniveaus erklären.
- Drittens gehen wir der Frage nach, welche Fähigkeiten auf den Gebieten der mathematischen und naturwissenschaftlichen Grundbildung von Experten und Abnehmern beruflicher Bildung beim Übergang in den Beruf erwartet werden und inwieweit die Leistungen von Teilnehmern an kaufmännischen, handwerklich-technischen und sozialpflegerischen Bildungsgängen diesen Erwartungen entsprechen.
- Viertens wird geprüft, ob Schülerinnen und Schüler in kaufmännischen, handwerklich-technischen und sozialpflegerischen Bildungsgängen relative Stärken und Schwächen auf den Gebieten der mathematischen und naturwissenschaftlichen Grundbildung besitzen, die mit differentiellen Lerngelegenheiten in der Berufsausbildung in Zusammenhang stehen.

Institutionelle und regionale Leistungsunterschiede

Personen mit unterschiedlichem allgemeinbildenden Schulabschluss – bzw. proaktivem Schulabschluss im Falle der Besucher einer gymnasialen Oberstufe – unterscheiden sich erwartungsgemäß im Mittel in ihrer mathematisch-naturwissenschaftlichen Grundbildung statistisch signifikant und praktisch bedeutsam. Am größten ist die Differenz zwischen Personen, die eine allgemeine Hochschulreife erlangt haben oder voraussichtlich erwerben werden, und den Schulabgängern mit einem Realschulabschluss. Die Differenz beträgt rund zwei Drittel einer Standardabweichung. Die Unterschiede zwischen den mittleren Grundbildungsleistungen von Personen mit Realschulabschluss, qualifiziertem Hauptschulabschluss und einfachem Hauptschulabschluss bzw. Abgangszeugnis betragen jeweils rund eine Drittel Standardabweichung.

Vergleicht man die Leistungsabstände zwischen den Schulformen mit den entsprechenden Befunden für die 8. Jahrgangsstufe, so zeigt sich die bereits in TIMSS/II beschriebene Leistungsschere zwischen den Schulformen tendenziell erweitert. Insbesondere scheinen die Leistungsunterschiede zwischen Haupt- und Realschule zum Abschluss der Sekundarstufe zu wachsen. In einem regionalen Leistungsver-

Abbildung 5: Mathematisch-naturwissenschaftliche Grundbildung nach Schulabschluss



IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

gleich wird deutlich, dass sich alte und neue Bundesländer zwar nicht signifikant unterscheiden. Eine genauere Inspektion der Daten zeigt allerdings, dass die Realschulabsolventen in den neuen Ländern ein vergleichsweise niedriges Grundbildungsniveau erreichen. Die stark expandierte Realschule der neuen Länder scheint noch nicht die angemessenen Leistungsstandards gefunden zu haben. Damit werden Befunde aus TIMSS/II bestätigt.

Betrachtet man schließlich die drei Bildungsgruppen der Sekundarstufe II – Allgemeinbildung, theoretisch-orientierte Berufsbildung (Fachgymnasien und Fachoberschulen) und praktisch-orientierte Berufsbildung (Berufs- und Berufsfachschulen) – gleichzeitig unter dem Gesichtspunkt des allgemeinen Schulabschlusses und der erreichten Grundbildung, so wird deutlich, dass beim Übergang in die Sekundarstufe II institutionelle und individuelle Selektionsprozesse ineinander greifen. Die Sekundarstufe II ist ein leistungsmäßig stratifiziertes System, das sowohl durch an Abschlüsse gebundene Berechtigungssysteme als auch durch Selbstselektion stabilisiert wird.

Leistungsvergleiche von beruflichen Bildungsgängen

Für einen Vergleich von beruflichen Bildungsgängen wurden die in der TIMS-Studie vertretenen betrieblichen und überbetrieblichen Ausbildungsgänge von Experten der Berufsbildungsforschung drei großen Gruppen zugeordnet, zu denen jeweils mathematiknahe, techniknahe und mathematik- und technikferne Bildungsgänge gehören. Die mathematiknahen Bildungsgänge vereinigen kaufmännische Berufe und Ausbildungsgänge im Bereich Wirtschaft und Verwaltung. Den techniknahen Bildungsgängen wurden Metall- und Elektroberufe sowie Bau- und Bau-nebenberufe zugeordnet. Ausbildungsgänge in den Bereichen Landwirtschaft, Hauswirtschaft, Pflege, Körperpflege usw. wurden zu mathematik- und technikfernen Bildungsgängen zusammengefasst. Die Leistungsvergleiche in den Bereichen mathematischer und naturwissenschaftlicher Grundbildung weisen substantielle Niveaувorsprünge von mathematik- und techniknahen gegenüber mathematik- und technikfernen Bildungsgängen aus. Die Niveauunterschiede betragen zwischen zwei Drittel und einer Standardabweichung. Mathematik- und techniknahe Bildungsgänge unterscheiden sich nicht signifikant voneinander.

Kontrolliert man nach dem allgemeinen Schulabschluss und nach dem Geschlecht der Personen, zeigt sich lediglich im Bereich der mathematischen Grundbildung noch ein praktisch bedeutsamer Niveaувorsprung mathematik- und techniknaher Bildungsgänge. Bei Betrachtung eigens skaliertes Untertests wie der biologischen und physikalischen Grundbildung stellen wir leichte Differenzverschiebungen fest, die als Ausdruck relativer Stärken und Schwächen von Auszubildenden gedeutet werden können. So ist der Niveaувorsprung von Schülern und Auszubildenden in mathematik- und technikfernen Bildungsgängen im Bereich der biologischen Grundbildung geringer. Die gleiche Tendenz zeigt sich bei Schülern und Auszubildenden in techniknahen Bildungsgängen im Bereich der physikalischen Grundbildung. Diese differentiellen Befunde lassen zwei Interpretationen zu: Es kann sich hierbei sowohl um Effekte der fachspezifischen Selbstselektion in den Ausbildungsberuf aufgrund von Interessen und selbst wahrgenommener Fähigkeiten als auch um Effekte fachbezogenen beruflichen Lernens handeln. Mithilfe der TIMS-Studie ist es nicht möglich, diese beiden Einflussgrößen empirisch voneinander zu trennen. Dies würde ein längsschnittliches Untersuchungsdesign voraussetzen. Unter Vorbehalt war es uns möglich, Testleistungen von Schülern und Auszubildenden unterschiedlicher Ausbildungsjahre miteinander zu vergleichen. Hier zeigt sich, dass das Ausbildungsjahr keinen Unterschied macht. Für einen förderlichen Einfluss der Berufsausbildung ließe sich der mit aller Vorsicht zu behandelnde Befund werten, dass Auszubildende in fachrichtungsnahen Sachgebieten zu besseren Leistungen fähig sind, wenn sie den jeweiligen Fachunterricht positiv erleben und beurteilen. Hier

deutet sich eine standardsichernde Funktion des Berufsschulunterrichts an, die wir jedoch nicht überbetonen wollen. Es bleibt weiteren Untersuchungen vorbehalten, die hier angedeuteten Befunde mithilfe eines längsschnittlichen Untersuchungsdesigns zu prüfen. *Die Analysen sprechen cum grano salis dafür, dass im Wesentlichen der in der Sekundarstufe I besuchte Bildungsgang über das mathematisch-naturwissenschaftliche Grundbildungsniveau von Schülerinnen und Schülern am Ende der Berufsausbildung entscheidet.*

Gegenüberstellung von Anforderungen und Leistungen

Leistungsvergleiche zwischen Berufsgruppen geben deskriptive Auskünfte über Kompetenzniveaus, liefern jedoch kein Tableau normativer Zielvorgaben. Die Leistungen am Übergang von der Ausbildung in den Beruf bedürfen der Betrachtung im Kontext von beruflichen Bildungsvorstellungen. Deshalb haben wir die mathematisch-naturwissenschaftliche Grundbildung vor dem Hintergrund von Erwartungen der Abnehmer untersucht. In einer Zusatzuntersuchung haben wir je 19 Aufgaben aus den Bereichen der mathematischen und naturwissenschaftlichen Grundbildung durch Experten und Abnehmer beruflicher Bildung im Hinblick auf ihre berufliche Relevanz einstufen lassen. Es handelte sich hierbei um 35 aus den TIMSS-Testschulen stammende Berufsschullehrer und um 50 Vertreter von Industrie- und Handels- bzw. Handwerkskammern. Die Einstufungen wurden für kaufmännische, handwerklich-technische und sozialpflegerische Bildungsgänge vorgenommen.

In der mathematischen Grundbildung sollten Teilnehmer an beruflichen Bildungsgängen gut drei Viertel aller Aufgaben mit hinreichender Sicherheit lösen. Dass die Messlatte der Experten hier allerdings hoch liegt, mag darin zum Ausdruck kommen, dass derartige Leistungen einigermaßen sicher nur von denen erzielt werden, die zur Hochschule führende Bildungsgänge durchlaufen haben. Dies trifft beispielsweise für die angehenden Bankkaufleute in der TIMS-Studie zu, von denen die meisten die allgemeine Hochschulreife besitzen. Kaufmännische und handwerklich-technische Berufe lösen die von den Experten eingestufteten Aufgaben im Mittel zu 50 Prozent, sozialpflegerische nur zu 30 Prozent. Im naturwissenschaftlichen Bereich sieht die Situation etwas anders aus: Experten und Abnehmer stufen nur etwa die Hälfte der Aufgaben als beruflich relevant ein. Dabei liegen die Anforderungen in sozialpflegerischen Berufen mit knapp 60 Prozent und im handwerklich-technischen Bereich mit ca. 50 Prozent relevant eingestufte Aufgaben erwartungsgemäß höher als in kaufmännischen Bildungsgängen, für die nur ein Drittel der Aufgaben relevant sind. Hier liegen die Erfolgsquoten im Mittel bei 53 Prozent für kaufmännische

nische und handwerklich-technische Berufe und bei 39 Prozent für sozialpflegerische Berufe. In der naturwissenschaftlichen Grundbildung ist demnach für sozialpflegerische Bildungsgänge eine Diskrepanz zwischen Anforderungen und Leistungen festzustellen. *Zusammenfassend konstatieren wir insbesondere im Bereich der mathematischen Grundbildung für die Mehrheit der Auszubildenden eine Diskrepanz zwischen den Kompetenzen, die von Abnehmern erwartet werden, und den am Ende der Sekundarstufe II erzielten Ergebnissen.*

Relative Stärken und Schwächen und differentielle Lerngelegenheiten

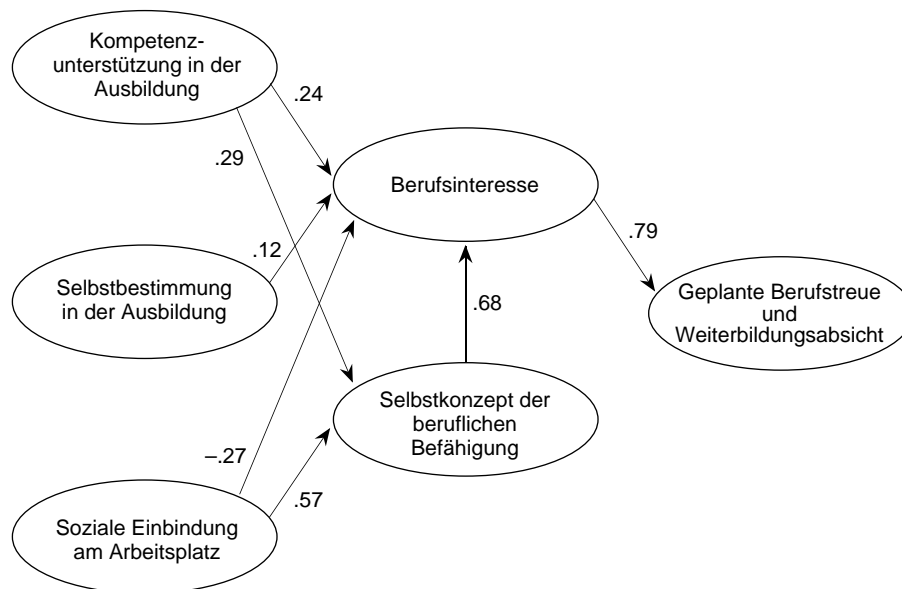
Methodisch differenziertere Analysen auf der Ebene von Einzelaufgaben geben Auskunft über spezifische Stärken und Schwächen der Schülerinnen und Schüler beruflicher Bildungsgänge, die über einen normativen Vergleich hinausgehen. Die Ergebnisse der Analysen *differentieller Itemfunktionen* lassen sich auf folgende Weise zusammenfassen: Die Lösung von Aufgaben der mathematischen und naturwissenschaftlichen Grundbildung fällt Schülern und Auszubildenden dann leichter, wenn die Problemstellung in einen berufsfeldspezifischen Kontext eingebunden ist. Bei diesen Aufgabentypen gelingt ihnen vermutlich die entsprechende Modellbildung bzw. die Übertragung in ein Situationsmodell besser. Wird beispielsweise danach gefragt, welche Funktion die weißen Blutkörperchen haben, erzielen Schülerinnen und Schüler sozialpflegerischer Berufe bei einem insgesamt geringeren Leistungsniveau sogar etwas bessere Leistungen (44 %) als Teilnehmer an kaufmännischen und handwerklich-technischen Bildungsgängen (42 % bzw. 39 %). Weitere Analysen zeigen, dass derartige relative Stärken und Schwächen vor allem in der naturwissenschaftlichen Grundbildung in hohem Maße mit den Lerngelegenheiten im Fachunterricht korrespondieren.

Neben dem Erwerb von berufsspezifischen Kompetenzen in der Ausbildung gilt die Entwicklung einer intrinsisch geprägten motivationalen Orientierung gegenüber dem Ausbildungsberuf als ein Kriterium für eine erfolgreiche Berufsausbildung. Empirische Untersuchungen weisen darauf hin, dass eine auf Selbstbestimmung und Kompetenzerleben beruhende intrinsische Berufsmotivation auch die Motivation zur individuellen Weiterbildung und Qualifikation steigert. Gerade vor dem Hintergrund einer ständigen Veränderungen unterliegenden Arbeitswelt wird es für Arbeitnehmer immer wichtiger, den wandelnden Anforderungsstrukturen im Berufsleben mit der Einstellung zu begegnen, dass Weiterbildungs- und Qualifizierungsprozesse für langfristigen beruflichen Erfolg notwendig sind. In der nationalen Erweiterung von TIMSS/III haben wir diesen Ausbildungskontexten und motiva-

tionalen Orientierungen nachgespürt. Die auf die Auszubildenden im dualen System beschränkten Analysen verfolgen die Beantwortung der Frage: Welche Bedeutung haben berufliche Sozialisationserfahrungen für die Entwicklung einer intrinsisch motivierten Berufsorientierung und einer Bindung an den Ausbildungsberuf?

In Anlehnung an die Selbstbestimmungstheorie von Deci und Ryan sind wir davon ausgegangen, dass Erlebnisqualitäten in der betrieblichen Ausbildung, die dem Bedürfnis nach Kompetenzunterstützung, Selbstbestimmung und sozialer Eingebundenheit entgegenkommen, sowohl für die Entwicklung des Selbstkonzepts der beruflichen Befähigung als auch für eine intrinsisch geprägte Berufsorientierung eine wichtige Rolle spielen. Ferner sollte sich zeigen, dass eine intrinsisch geprägte Berufsmotivation auch die Bereitschaft zur Weiterentwicklung der beruflichen Qualifikation erhöht. Das Beziehungsgefüge in Abbildung 6 stellt das Ergebnis der Analysen in etwas vereinfachter Form dar. Es bestätigt im Wesentlichen die getroffenen Annahmen. Alle dort aufgeführten Einflussgrößen sind statistisch signifikant; die

Abbildung 6: Beziehungsgefüge zwischen subjektiven Ausbildungsumwelten, Kompetenzerleben, Berufsinteresse und beruflichen Plänen



Werte können jeweils zwischen -1 (= stark negativer Einfluss) und $+1$ (= stark positiver Einfluss) variieren. Werte nahe Null würden besagen, dass kein Einfluss besteht.

Die soziale Eingebundenheit stellt sich dabei, sofern sie sich auf Berufserfahrung und Kompetenzerleben bezieht, als ein wichtiger Entwicklungskontext hierfür heraus. Gleichzeitig kann die soziale Einbindung am Arbeitsplatz jedoch zu einer Minderung der intrinsischen Berufsmotivation führen, wenn die sozialen Beziehungen nicht primär mit der Entwicklung von Berufskompetenz zu tun haben. In den Ausbildungsberufen, in denen mathematische und technische Kompetenzen besonders gefragt sind, ließ sich ebenfalls ein positiver Einfluss der mathematisch-naturwissenschaftlichen Grundbildung auf das Selbstkonzept feststellen (ohne Abb.). *Aus pädagogisch-psychologischer Perspektive weisen die Ergebnisse auf die Notwendigkeit von Bedingungen in der Ausbildung hin, die möglichst allen Auszubildenden eine optimale Befriedigung ihrer Bedürfnisse nach Kompetenzerfahrung, Selbstbestimmung und sozialer Einbindung im Kontext berufsbezogener Tätigkeiten ermöglichen.*

6. Mathematisch-naturwissenschaftliche Grundbildung ausländischer Schulausbildungsabsolventen und schichtenspezifische Einflüsse beim Übergang auf die Sekundarstufe II

Grundbildung ausländischer Schulausbildungsabsolventen

Trotz vielfacher Probleme, die ausländische Mitbürger bzw. Aussiedler im deutschen Bildungssystem haben, hat in den letzten Jahren die Bildungsbeteiligung von ausländischen Kindern und Jugendlichen kontinuierlich zugenommen. Dennoch unterscheiden sich ihre Bildungskarrieren nach wie vor von diesen, die Kinder und Jugendliche durchlaufen, die in der Bundesrepublik geboren bzw. ausschließlich mit deutscher Muttersprache aufgewachsen sind. Schwierigkeiten mit der deutschen Sprache werden primär für diese Unterschiede verantwortlich gemacht.

Da die Schule unter anderem auf die Anforderungen künftiger Berufstätigkeit vorbereitet ist davon auszugehen, dass die jugendlichen ausländischen Schüler auch andere Berufsperspektiven entwickeln als Schüler mit ausschließlich deutschem kulturellen Hintergrund. Es zeigt sich, dass sich die unterdurchschnittliche Beteiligung ausländischer Jugendlicher an den höheren Schulformen bei der weiteren Gestal-

tung des Bildungsgangs in der Tat deutlich bemerkbar macht. Dies auch dann, wenn man den sozioökonomischen Hintergrund der Jugendlichen mitberücksichtigt.

Darüber hinaus wird deutlich, dass die Bildungsbeteiligung der Jugendlichen umso ungünstiger ausfällt, je enger die Bindung der Familie an eine nicht deutsche Heimatkultur ist. Die Analysen zeigen ferner, dass keine substantielle Benachteiligung von Jugendlichen nicht deutschsprachiger Familien innerhalb der verschiedenen Bildungsgänge besteht, das heißt, innerhalb eines Bildungszweigs müssen Jugendliche nicht mit schlechteren Schulleistungen rechnen, nur weil sie einen „ausländischen“ Hintergrund haben. Zwischen den Bildungsgängen, insbesondere beim Übergang auf die Schulformen der gegliederten Sekundarstufe, lassen sich Disparitäten in der Verteilung der Jugendlichen partiell bzw. nicht deutscher Muttersprache nachweisen. Jugendliche ausländischer Herkunft finden sich deutlich seltener in den gymnasialen Oberstufen bzw. der theoretisch orientierten beruflichen Bildung (z.B. Fachoberschule).

Das insgesamt schlechtere Abschneiden der ausländischen Schüler in mathematisch-naturwissenschaftlichen Fächern, so die Schlussfolgerung, geht nur zum Teil auf ihre sprachlichen Schwierigkeiten zurück. Kulturell gebundene Unterschiede in den elterlichen Bildungsvorstellungen – zumal für die weiblichen Jugendlichen – sowie schichtenspezifische Effekte erweisen sich als besonders starke Determinanten für die gewählten Ausbildungsgänge.

Übergang auf die Sekundarstufe II

An allen wichtigen Gelenkstellen unseres Schulwesens sind Wahlentscheidungen über den weiteren Bildungsverlauf zu treffen, auf die die Eltern erheblichen Einfluss haben und an denen folglich schichtenspezifische Einflüsse wirksam werden. Wichtige Entscheidungen sind zum einen bei dem Übergang von der Grund- auf die weiterführende Schule zu treffen als auch am Ende der Sekundarstufe I, wenn eine praktisch orientierte Ausbildung begonnen oder ein höherer Bildungsabschluss angestrebt werden kann.

Die Bildungswünsche, die Eltern für ihre Kinder haben, stehen bekanntlich in deutlichem Zusammenhang ihres eigenen Berufs- und Bildungshintergrunds. Wie die Analysen anhand der TIMSS/III-Daten zeigen, lassen sich elterlicher Steuerungseffekte auf Bildungsentscheidungen an der Übergangsschwelle zum Ende der Sekundarstufe I, aber auch jenseits schichtenspezifischer Effekte erkennen. Jugendliche zum Beispiel, die insgesamt gute Schulleistungen und eine vergleichsweise

frühe Technikorientierung zeigen, werden von allen Eltern (d.h. ungeachtet der sozialen Herkunft) in gleichem Ausmaß dazu ermuntert, sich auf ein Fachhochschulstudium hin zu orientieren.

Für Absolventen mit mittlerem Abschluss hingegen, die ein weniger klares Neigungsprofil zeigen, werden Effekte der sozialen Herkunft deutlich sichtbar: Eltern mit höherer Schulbildung ermutigen ihre Kinder in dieser Situation systematisch häufiger zum Besuch der gymnasialen Oberstufe, während Eltern unterer Sozialschichten die Berufs- und Bildungsunsicherheiten stärker problematisieren und eher auf eine frühzeitige Berufsentscheidung und somit eher auf eine traditionelle Berufsausbildung hinwirken.

7. Fachleistungen im voruniversitären Mathematik- und Physikunterricht: Theoretische Grundlagen, Kompetenzstufen und Unterrichtsschwerpunkte

Zielpopulation für die TIMSS-Tests im voruniversitären Bereich sind ausschließlich jene Schülerinnen und Schüler, die bis zum Ende ihrer Sekundarschulausbildung als Teil ihrer Qualifizierung für einen späteren Hochschulbesuch in Mathematik bzw. Physik unterrichtet wurden. In Deutschland sind dies Schüler der gymnasialen Oberstufen, die entsprechende Grund- oder Leistungskurse besucht haben. Das Testkonzept für die voruniversitäre Mathematik und Physik zielt daher auf fachliche, „akademische“ Kompetenzen und nicht – wie die Tests zur Grundbildung – auf alltagsbezogene Anwendung mathematisch-naturwissenschaftlicher Konzepte.

Maßstab für die Aussagekraft (Validität) der voruniversitären Tests sind die fachliche Breite und Tiefe sowie die Passung zu Lehrplänen und Unterrichtspraxis in der Sekundarstufe II, zu den Prüfungsanforderungen des Abiturs und zu den Anforderungen im Grundstudium. Durch Befragung von Fachdidaktikern, Lehrplanexperten, Lehrern der beteiligten Schulen und Hochschullehrern haben wir diese Aspekte der Testvalidität geprüft. Hier gehen wir im Einzelnen folgenden Fragen nach:

- Zunächst werden Konzeption und Aufbau der beiden Tests beschrieben. In diesem Zusammenhang wird auch geprüft, ob über verschiedene Sachgebiete und Verhaltenserwartungen hinweg jeweils eine homogene Fähigkeitsskala für Mathematik und Physik gebildet werden kann.

- Die Anforderungsmerkmale der mathematischen Aufgaben werden detailliert untersucht. Auf der Basis von Expertenurteilen können wir ermitteln, von welchen Anforderungen die Schwierigkeit der Testitems abhängt. Ferner werden qualitative Stufen mathematischer Kompetenz (*Proficiency Levels*) unterschieden.
- Analoge Untersuchungen gelten schwierigkeitsbestimmenden Anforderungen und Kompetenzstufen in der Physik.
- Zusätzlich werden TIMSS-Aufgaben aus der Perspektive von Lehrplanexperten und Fachleitern an den Testschulen untersucht. Zum einen geht es um die Frage, ob die TIMSS-Aufgaben Inhalte der Lehrpläne und der Unterrichtspraxis angemessen repräsentieren, das heißt um die curriculare und die unterrichtliche Validität der Tests. Zum anderen lassen sich, basierend auf konkreten Aufgaben und Anforderungsmerkmalen, Diskrepanzen zwischen Curriculum und schulischer Praxis aufzeigen.
- Abschließend untersuchen wir die Tests aus der Perspektive von Hochschullehrern, um ihre kriteriale Validität zu bestimmen.

Konzeption und Aufbau der Tests für die gymnasiale Oberstufe

Beide Testteile stützen sich auf umfangreiche internationale Curriculumanalysen, die für die Oberstufe ebenso wie für die Mittelstufe zu einem Rahmenkonzept führten, das nach Sachgebieten und Verhaltenserwartungen unterscheidet. Nach umfangreichen Entwicklungsarbeiten und Vorerprobungen verblieben für Mathematik und Physik jeweils 65 Aufgaben. Etwa die Hälfte der Punkte waren in offenen Fragen zu erreichen, die andere Hälfte in *Multiple Choice*-Items. Passend zu dem Ziel, sich eng an die Curricula der Schulfächer anzulehnen, nahmen in der Mathematik Aufgaben, die Wissen und die Beherrschung von Standardroutinen beinhalten, einen breiten Raum ein; gleichgewichtig wurden aber komplexe Verfahren und Problemlösungen gefordert. In der Physik wurde eine Balance erreicht zwischen Quantifizierung und mathematischer Modellierung einerseits, qualitativen und auf Experimente bezogenen Aufgaben andererseits.

Strukturanalysen wiesen nach, dass der Mathematiktest über die verschiedenen Sachgebiete und Verhaltenserwartungen hinweg eine homogene Leistungsskala erzeugt. Die Physikleistungen deutscher Oberstufenschüler sind etwas weniger konsistent. Vor allem in der Wärmelehre besteht eine deutliche Diskrepanz zwischen Anforderungen der Lehrpläne und tatsächlichem Unterricht. Die internationale

Tabelle 8: Testaufgaben für den voruniversitären Mathematikunterricht nach Sachgebiet und Verhaltenserwartung

| Sachgebiet | Verhaltenserwartung | | | | Insgesamt |
|------------------------------------|---------------------|-----------------------------------|--------------------------------------|---|-----------|
| | Wissen | Beherrschung von Routineverfahren | Beherrschung von komplexen Verfahren | Anwendungsbezogene Aufgaben und innermathematische Probleme | |
| Zahlen, Gleichungen und Funktionen | 1 | 7 | 2 | 7 | 17 |
| Analysis | 2 | 8 | – | 5 | 15 |
| Geometrie, analytische Geometrie | 5 | 6 | 3 | 9 | 23 |
| Wahrscheinlichkeit, Statistik | 1 | 2 | 1 | 3 | 7 |
| Aussagenlogik, Beweise | – | 1 | – | 2 | 3 |
| Insgesamt | 9 | 24 | 6 | 26 | 65 |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Tabelle 9: Testaufgaben aus dem Bereich voruniversitärer Physik nach Sachgebiet und Verhaltenserwartung

| Sachgebiet | Verhaltenserwartung | | | Insgesamt | |
|---|---|---|-------------|-----------|---|
| | Verstehen von Einzelinformationen und komplexen Informationen | Konzeptualisieren, Anwenden Qualitativ | Quantitativ | | Experimentieren, Beherrschung von Verfahren |
| Mechanik | 1 | 2 | 8 | 5 | 16 |
| Elektrizität und Magnetismus | 4 | 1 | 10 | 1 | 16 |
| Wärmelehre | 4 | 2 | 2 | 1 | 9 |
| Wellen und Schwingungen | 4 | 2 | 3 | 1 | 10 |
| Teilchen-, Quanten-, Astrophysik, Relativitätstheorie | 4 | 3 | 4 | 3 | 14 |
| Insgesamt | 17 | 10 | 27 | 11 | 65 |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Physik-Kompetenzskala kann aber mit einem eindimensionalen Modell hinreichend gut beschrieben werden.

Anforderungsmerkmale und Kompetenzstufen in der voruniversitären Mathematik

An vier Punkten der mathematischen Fähigkeitsskala haben wir geprüft, welche spezifischen Leistungen jeweils erbracht werden müssen, um das entsprechende Niveau zu erreichen. Die beiden folgenden Abbildungen verdeutlichen diese Kompetenzstufen mit entsprechenden Aufgaben aus der elementaren Statistik und der Analysis; in Band 2 unseres ausführlichen Abschlussberichts finden sich zahlreiche weitere Beispiele.

Stufe 1 (Score 400): Elementares Schlussfolgern

Zur Interpretation des Diagramms in Beispielaufgabe J12 ist noch kein exaktes mathematisches Wissen erforderlich; es genügt die intuitive Abschätzung eines Durchschnittswertes. Wer nur die unterste Kompetenzstufe erreicht, kommt über derartiges elementares Schlussfolgern nicht hinaus.

Stufe 2 (Score 500): Anwendung einfacher mathematischer Begriffe und Regeln

Auch unsere zweite Beispielaufgabe aus der Statistik (L15), ist ohne Konzepte der Oberstufenmathematik lösbar. Sie setzt allerdings das Konzept der linearen Funktion voraus, das in der Mittelstufe eingeführt wird. Beispielaufgabe K3, die im weiteren Sinne dem Sachgebiet Analysis zugehört, setzt lediglich voraus, dass man Geschwindigkeits-Zeit-Graphen kennt und den Begriff „Beschleunigung“ mit der Steigung dieses Graphen verbindet – auch dies Konzepte, die noch in der Mittelstufe oder zu Beginn der Oberstufe vermittelt werden.

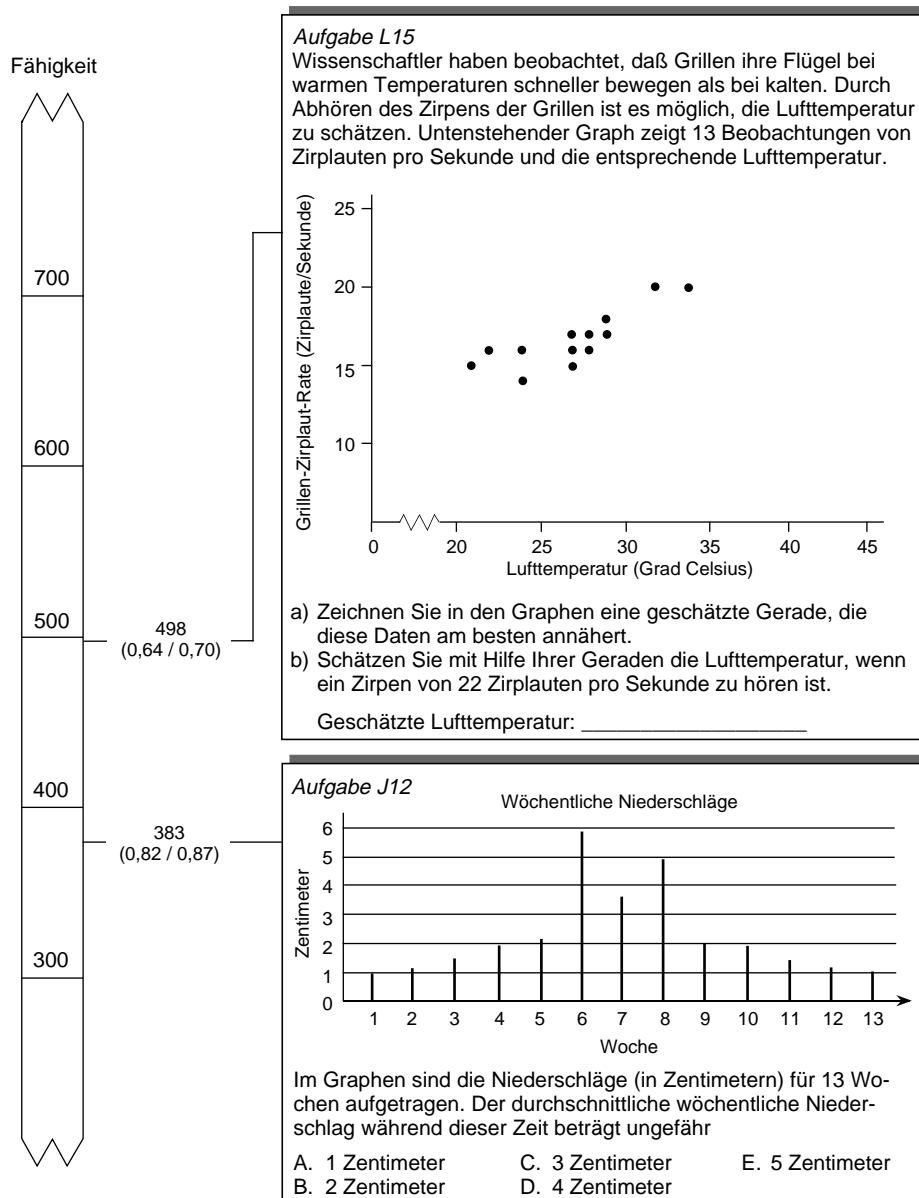
Stufe 3 (Score 600): Anwendung von Lerninhalten der Oberstufe

Erst auf dieser Kompetenzstufe werden typische Standardaufgaben der Oberstufenmathematik mit hinreichender Sicherheit gelöst. Hierzu gehören Kurvendiskussionen (siehe Beispielaufgabe K5) und die Bestimmung der Summe einer geometrischen Reihe, die häufig im Unterricht behandelt wird (L5).

Stufe 4 (Score 700): Selbstständiges Problemlösen auf Oberstufenniveau

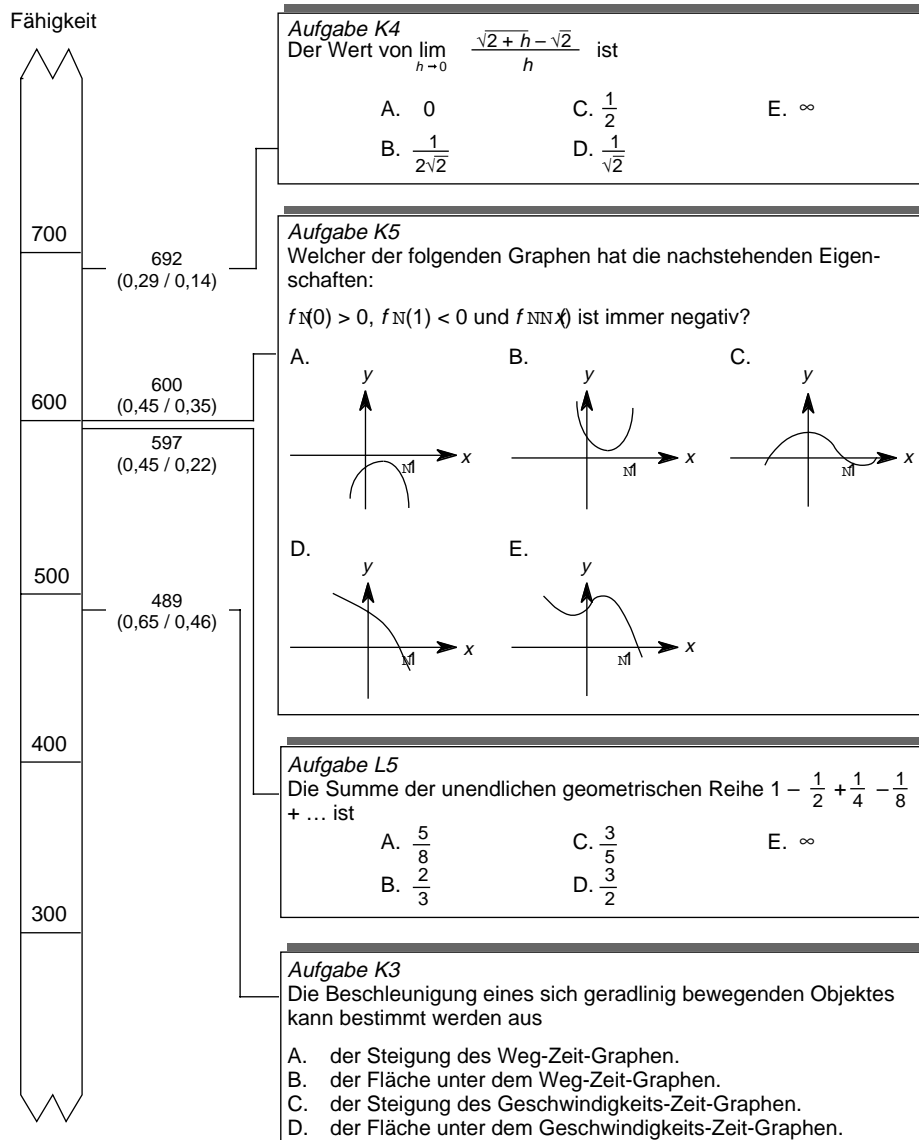
Um in Beispielaufgabe K4 den Grenzwert zu bestimmen, kann man unterschiedliche Wege beschreiten, zum Beispiel die binomische Formel anwenden oder die Ableitung der Wurzel aus x verwenden. Für den Bearbeiter bedeutet dies, dass er einen

Abbildung 7: Beispielaufgaben zur voruniversitären Mathematik, Teilgebiet Statistik



Die Werte an den Verbindungslinien zwischen den Beispielen und der Fähigkeitssäule geben das für eine 65-prozentige Lösungswahrscheinlichkeit erforderliche Fähigkeitsniveau und die Werte in Klammern die relativen internationalen und deutschen Lösungshäufigkeiten an.

Abbildung 8: Beispielaufgaben zur voruniversitären Mathematik, Teilgebiet Analysis



Die Werte an den Verbindungslinien zwischen den Beispielen und der Fähigkeitsäule geben das für eine 65-prozentige Lösungswahrscheinlichkeit erforderliche Fähigkeitsniveau und die Werte in Klammern die relativen internationalen und deutschen Lösungshäufigkeiten an.

eigenen Lösungsweg finden und dabei kreativ sein Wissen aus unterschiedlichen Bereichen der Mathematik einsetzen muss.

Nach einem eigens entwickelten Kategoriensystem zur Beschreibung von Anforderungsmerkmalen haben zehn Fachdidaktiker die mathematischen TIMSS-Aufgaben bewertet. Ihrem Urteil zufolge erfordern die Aufgaben vor allem die Kenntnis von mathematischen Begriffen und Verfahren. In den oberen Stufen der Skala steigen die Anforderungen an das Fachwissen, aber auch an das Problemlösen, die Beherrschung von Prinzipien mathematischen Denkens und den Umgang mit formalen Modellen. Der Test zur voruniversitären Mathematik passt somit zu den fachlichen und didaktischen Ansprüchen an den gymnasialen Mathematikunterricht.

Anforderungsmerkmale und Kompetenzstufen in der voruniversitären Physik

Im Bereich der Physik lässt sich noch eine fünfte, fachdidaktisch sehr bedeutsame Kompetenzstufe identifizieren. Aus dem umfangreichen Beispielmateriale unseres Berichts haben wir hier die Aufgaben zur Mechanik ausgewählt.

Stufe 1 (Score 450): Lösen von Routineaufgaben mit Mittelstufenwissen

Das unterste Kompetenzniveau indiziert einen erfolgreichen Umgang mit Standardstoffen der Mittelstufe. Aufgabe E5 beispielsweise ist eine typische Aufgabe zum „freien Fall“.

Stufe 2 (Score 550): Anwendung von Faktenwissen zur Erklärung einfacher Phänomene der Oberstufenphysik

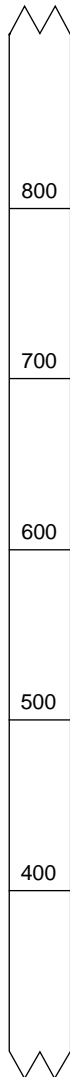
Die beiden Aufgaben, die für diese zweite Kompetenzstufe charakteristisch sind, beinhalten Themen der Teilchenphysik und damit des Oberstufenunterrichts. Zur Lösung sind allerdings nur relativ einfache Faktenkenntnisse erforderlich, beispielsweise das Wissen, dass im Inneren von Sternen Kernfusionen stattfinden.

Stufe 3 (Score 650): Anwendung physikalischer Gesetze zur Erklärung experimenteller Effekte auf Oberstufenniveau

Mit dieser Kompetenzstufe wird ein Kern physikalischer Kompetenz erreicht, nämlich die Kenntnis physikalischer Gesetze und klassischer Experimente, auf die jene Gesetze zutreffen. Hierzu gehören der Impulserhaltungssatz (Beispielaufgabe G12), die Bewegungsgesetze der schiefen Ebene (H1), das Gay-Lussacsche Gesetz, die Grundgleichungen des Elektromagnetismus und das Rutherford'sche Experiment aus der Teilchenphysik.

Abbildung 9: Beispielaufgaben zur voruniversitären Physik, Teilgebiet Mechanik

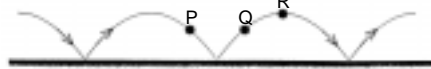
Fähigkeit



840
(0,16 / 0,07)

Aufgabe G15

Die Abbildung zeigt die Bewegung eines Balls, der bei vernachlässigtem Luftwiderstand auf dem Boden springt.

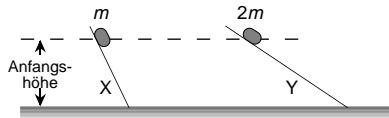


Zeichnen Sie Pfeile in die Abbildung ein, die die Richtung der Beschleunigung des Balls in den Punkten P, Q und R angeben.

650
(0,39 / 0,24)

Aufgabe H1

Zwei Kästen der Massen m und $2m$ gleiten die schiefen Ebenen X bzw. Y hinab. Sie starten aus dem Ruhezustand in derselben Höhe. Die beiden Ebenen sind unterschiedlich stark geneigt und weisen eine zu vernachlässigende Reibung auf.

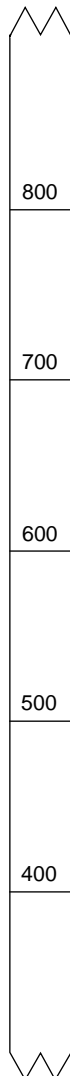


Welche der folgenden Aussagen ist NICHT richtig?

- A. Am oberen Ende der Ebenen hat einer der Kästen eine halb so große potentielle Energie wie der andere.
- B. Die Kästen haben am unteren Ende der schiefen Ebenen dieselben Geschwindigkeiten.
- C. Die Kästen benötigen dieselbe Zeit, um das untere Ende der schiefen Ebenen zu erreichen.
- D. Der Kasten auf der Ebene X erfährt eine höhere Beschleunigung als der Kasten auf der Ebene Y.

noch Abbildung 9: Beispielaufgaben zur voruniversitären Physik, Teilgebiet Mechanik

Fähigkeit

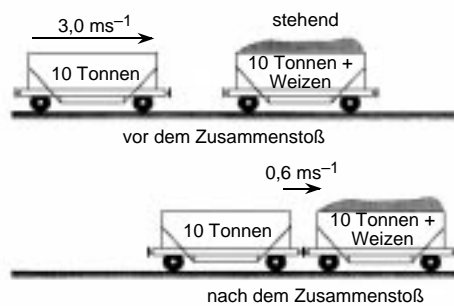


647
(0,36 / 0,24)

380
(0,79 / 0,75)

Aufgabe G12

Ein leerer Eisenbahnwagen mit einer Masse von 10 Tonnen ($1,0 \times 10^4$ kg) fährt mit der Geschwindigkeit $3,0 \text{ ms}^{-1}$. Er prallt auf einen identischen stehenden Wagen, der mit Weizen beladen ist. Während des Zusammenstoßes koppeln die beiden Wagen an und bewegen sich gemeinsam mit der Geschwindigkeit $0,6 \text{ ms}^{-1}$. Die Situationen vor und nach dem Zusammenstoß sind in den Abbildungen unten dargestellt.



Benutzen Sie diese Informationen, um die Masse des Weizens im beladenen Wagen zu berechnen. Schreiben Sie alle Ihre Arbeitsschritte auf.

Aufgabe E5

Ein Stein wird aus dem Ruhezustand in einen tiefen Schacht fallen gelassen. Nach 2 s schlägt er auf dem Boden auf.

Wie tief ist der Schacht? Der Einfluß des Luftwiderstands auf den fallenden Stein kann vernachlässigt werden. Rechnen Sie mit der Fallbeschleunigung $g = 9,8 \text{ ms}^{-2}$.

- A. 4,9 m
- B. 9,8 m
- C. 19,6 m
- D. 39,2 m
- E. 78,4 m

Die Werte an den Verbindungslinien zwischen den Beispielen und der Fähigkeitssäule geben das für eine 65-prozentige Lösungswahrscheinlichkeit erforderliche Fähigkeitsniveau und die Werte in Klammern die relativen internationalen und deutschen Lösungshäufigkeiten an.

Stufe 4 (Score 750): Selbstständiges fachliches Argumentieren und Problemlösen

Auf dieser Kompetenzstufe werden auch offene Fragestellungen zur Oberstufenphysik, die eigenständige Lösungsansätze und zum Teil divergente Denkprozesse erfordern, mit hinreichender Sicherheit gelöst. Beispielsweise muss ein Experiment konzipiert oder die Bahn von Teilchen in einem elektrischen Feld skizziert werden.

Stufe 5 (Score 850): Überwinden von Fehlvorstellungen

Die fachdidaktische Forschung hat nachgewiesen, dass alltagsgebundene Fehlvorstellungen auch bei fortgeschrittenen Lernern oft anzutreffen sind. So wird häufig die Richtung der Beschleunigung eines Objekts mit dessen Bewegungsrichtung verwechselt (vgl. Beispielaufgabe G15, in der Pfeile nach *unten* einzuzeichnen sind). Interessanterweise kennzeichnet die Überwindung solcher Fehlvorstellungen das höchste Kompetenzniveau der voruniversitären Physik.

Experten aus der Physikdidaktik kamen zu der Einschätzung, dass die TIMSS-Aufgaben vor allem fachliches Wissen erfordern, etwa ein qualitatives Verständnis physikalischer Begriffe und die Kenntnis typischer Experimente. Schwierige Aufgaben erfordern insbesondere ein Verständnis für funktionale Zusammenhänge zwischen physikalischen Größen – wobei das qualitative Verständnis wichtiger ist als die mathematische Darstellung – und das Überwinden von Fehlvorstellungen. Der TIMSS-Test zur voruniversitären Physik entspricht demnach dem Diskussionsstand der Fachdidaktik.

Lehrpläne, Unterrichtsinhalte und Studienanforderungen aus der Perspektive von TIMSS

Nach den Einschätzungen von Lehrplanexperten aller Bundesländer entspricht die Verteilung der voruniversitären Mathematik- und Physikaufgaben auf unterschiedliche Anforderungsstufen in etwa derjenigen im Abitur. Wichtiger noch ist, dass die Experten die curriculare Validität der TIMSS-Aufgaben bestätigten. Für jede Testaufgabe und jedes Bundesland wurde – jeweils für Grund- und Leistungskurse getrennt – gefragt, ob der Lehrstoff, der hier geprüft wird, Gegenstand des Curriculums sei. Die Antwort fiel sowohl in der Mathematik als auch in der Physik in über 90 Prozent der Fälle positiv aus. Zusätzlich wurde gefragt, wann das entsprechende Thema erstmals eingeführt wird, und in der Mathematik haben wir eigens die Lehrpläne der Jahrgangsstufe 13 (aus allen Ländern, die eine 13. Jahrgangsstufe vorsehen) mit den Anforderungen des TIMSS-Tests verglichen. Es zeigte sich, dass TIMSS auch den Lehrstoff einbezogen hat, der in Deutschland im gymnasialen Abschlussjahrgang behandelt wird.

Deutliche Schwankungen bestehen allerdings in Abhängigkeit vom Kursniveau sowie von Sachgebiet und Anforderungsmerkmalen der Aufgaben. Aufgaben zur Wahrscheinlichkeitsrechnung und Statistik sind eher in Mathematikgrundkursen abgedeckt als in Leistungskursen, bei der Analysis ist es umgekehrt. Offenbar sehen die Lehrpläne spezifische Profile für Grund- und Leistungskurse vor. Ähnlich in der Physik: Hier sind Aufgaben, die alltagsbezogene Anwendungen ansprechen, eher im Grundkurs abgedeckt, gehobene fachliche Wissensinhalte und Mathematisierungen hingegen in den Leistungskursen.

Die Lehrer der Testschulen sagen in etwa 80 Prozent der Fälle, dass der Lehrstoff von TIMSS-Aufgaben im Unterricht behandelt wird. Dieser Wert dokumentiert eine erfreulich hohe Unterrichtsvalidität der Tests. Die Diskrepanz zu den etwa 10 Prozentpunkte höheren Angaben der Lehrplanexperten zeigt jedoch, dass curriculare Richtlinien an den Schulen nicht voll umgesetzt werden. Dies betrifft insbesondere die Grundkurse: In der Mathematik überwiegt hier der Stoff aus der Analysis, obwohl laut Lehrplan Wahrscheinlichkeitsrechnung und Statistik höhere Bedeutung besitzen. In der Physik behandeln Grundkurse – entgegen dem Lehrplan – eher *selten* alltagsbezogene Fragestellungen.

Je zehn Kernfragen der beiden voruniversitären Tests wurden von 250 Hochschullehrern unterschiedlicher Disziplinen im Hinblick auf Anforderungen des Grundstudiums bewertet. Im Durchschnitt wurden 70 Prozent der Mathematikaufgaben sowie 56 Prozent der Physikaufgaben als studienrelevant eingestuft, wobei jede Disziplin charakteristische Schwerpunkte setzte.

Als Fazit der Validitätseinschätzungen verschiedener Expertengruppen lässt sich festhalten, dass die Aufgaben der TIMSS-Tests zur voruniversitären Mathematik und Physik sowohl die Lehrpläne der Grund- und Leistungskurse als auch die unterrichtliche Realität und die Anforderungen verschiedener Studienfächer in hohem Ausmaß widerspiegeln. Aussagen über den Leistungsstand deutscher Abiturienten auf der Basis dieser Tests besitzen daher hohe Gültigkeit.

8. Fachleistungen im voruniversitären Mathematik- und Physikunterricht im internationalen Vergleich

Fragestellungen des internationalen Vergleichs

Wie in den übrigen Kapiteln unseres Abschlussberichts zu TIMSS/III steht auch beim internationalen Vergleich der Fachleistungen im voruniversitären Mathematik- und Physikunterricht nicht die Deskription, sondern die Analyse im Mittelpunkt des Interesses. Wir wollen deshalb den deskriptiven internationalen Vergleich auch nur als knappe Hintergrundinformation wieder aufnehmen. Drei analytische Fragestellungen bestimmen den Vergleich:

- Im ersten Schritt wurde – ähnlich wie bei der Analyse des mathematisch-naturwissenschaftlichen Grundbildungsniveaus – der Zusammenhang zwischen der Retentivitätsquote in der Sekundarstufe II eines Bildungssystems, der Expansion voruniversitärer „akademischer“ Bildungsgänge und den mathematischen und physikalischen Fachleistungen im voruniversitären Mathematik- und Physikunterricht untersucht. Im Hintergrund dieser Frage steht die verbreitete Sorge, dass mit einer Öffnung voruniversitärer Bildungsgänge – in Deutschland also vor allem der Oberstufe des Gymnasiums – das Anspruchsniveau dieser Bildungsgänge gesenkt und damit auch Spitzenleistungen beeinträchtigt werden könnten.
- Ein zweiter analytischer Schwerpunkt beschäftigt sich mit dem qualitativen Verständnis von Leistungsunterschieden und ihrer praktischen Bedeutung. Daher werden verschiedene methodische Wege eingeschlagen.
- Schließlich werden auf der Grundlage methodisch komplexer Analysen von Einzelaufgaben die Stärken und Schwächen des mathematischen und physikalischen Verständnisses von Oberstufenschülern untersucht.

Expansion voruniversitärer Bildungsgänge und Fachleistungen im Mathematik- und Physikunterricht

Obwohl die an TIMSS/III teilnehmenden Länder im Wesentlichen Industriestaaten sind, variiert die Bildungsbeteiligung an der Sekundarstufe II erheblich. Aufgrund der straffen Korrelation zwischen Bildungsbeteiligung und mittlerem nationalen Grundbildungsniveau haben wir die Haltekraft eines Schulsystems in der Sekundarstufe II als Indikator für die gesellschaftliche und politische Wertschätzung schu-

lischer Bildung interpretiert. Der relative Besuch voruniversitärer Bildungswege in den an TIMSS/III teilnehmenden Ländern variiert ebenfalls beträchtlich. Gleichzeitig ist die Expansion voruniversitärer „akademischer“ Bildungsgänge weitgehend unabhängig von der Haltekraft eines Bildungssystems. Die internationalen Beteiligungsquoten an voruniversitären Bildungsprogrammen liegen zwischen 18 und 80 Prozent der einschlägigen Alterskohorte bei einem Mittelwert von 44 Prozent und einer Standardabweichung von 17 Prozent. *Deutschland befindet sich mit einem relativen Besuch der gymnasialen Oberstufe von 25 Prozent der Alterskohorte ähnlich wie die anderen deutschsprachigen Länder am unteren Rand der Verteilung.*

Infolge der internen Differenzierung und Wahlmöglichkeiten auf der voruniversitären Ebene gibt der relative Besuch dieser Bildungsgänge nur eingeschränkt Auskunft über die Teilnahmequoten am voruniversitären Mathematik- oder Physikunterricht. Die Teilnahmequoten am Mathematikunterricht schwanken zwischen 2 und 75 Prozent der Alterskohorte bei einem Mittelwert von 10 Prozent und einer Standardabweichung von 16 Prozent. Die Variabilität der Teilnahme am Physikunterricht ist deutlich geringer. Die Extremwerte liegen bei 2 und 39 Prozent der Altersgruppe bei einem Mittelwert von 14 Prozent und einer Standardabweichung von 10 Prozent.

Deutschland hat für den Mathematikunterricht mit 25 Prozent eine hohe Beteiligungsrate, die nur in Österreich mit 33 Prozent und Slowenien mit 75 Prozent übertroffen wird. Mit einer Teilnahmequote am Physikunterricht von 9 Prozent der Altersgruppe liegt Deutschland dagegen unter dem internationalen Durchschnitt.

Die Öffnung voruniversitärer Bildungsgänge ist nicht nur in Deutschland, sondern auch in anderen Ländern immer wieder strittig. Skeptiker befürchten ein generelles Sinken des Anspruchsniveaus und insbesondere eine Beeinträchtigung möglicher Spitzenleistungen; Befürworter sehen in der hohen Beteiligung an diesen Bildungsprogrammen einen Indikator für die kognitive Mobilisierung der nachwachsenden Generation und die Anpassungsfähigkeit eines modernen Schulsystems. In den früheren internationalen Leistungsvergleichen, die den akademischen Bereich der Sekundarstufe II einbezogen haben, wurde deshalb regelmäßig auch nach dem Zusammenhang zwischen Selektivität der akademischen Programme und dem erreichten Leistungsniveau gefragt. Erwartungsgemäß fand man immer negative Korrelationen zwischen Teilnahmequote und mittlerem Leistungsniveau. Betrachtete man jedoch allein die Leistungsspitze eines Jahrgangs – etwa die 1 oder 5 Prozent testleistungsstärksten Untersuchungsteilnehmer –, war ein Zusammenhang nicht mehr nachweisbar. Dieser Befund wurde als Hinweis darauf interpretiert, dass ein Ausbau der zur Hochschule führenden Bildungswege Spitzenleistungen nicht beeinträchtigt.

Ein damit konsistentes Ergebnis haben wir im Rahmen von TIMSS/III für die mathematisch-naturwissenschaftliche Grundbildung berichtet.

Hinsichtlich des Zusammenhangs zwischen relativem Besuch der Mathematik- und Physikkurse und den erreichten Fachleistungen in TIMSS/III haben wir ähnliche Ergebnisse erwartet. Bezüglich der generellen Auswirkungen eines Ausbaus des akademischen Bereichs ließ sich jedoch keine gerichtete Hypothese begründen. Rechnet man bei einer Expansion des voruniversitären Bereichs mit einer generellen Senkung der Leistungsstandards, wird man auch nach Kontrolle der Beteiligungsdaten am voruniversitären Mathematik- und Naturwissenschaftsunterricht eine negative Korrelation zwischen der Breite des voruniversitären Bereichs und den Fachleistungen erwarten. Wenn man allerdings davon ausgeht, dass bei einer Öffnung der zur Hochschule führenden Bildungsgänge die interne Selektion in die Fachkurse effektiver verlaufen kann, da ein größerer Personenkreis für die Auswahl zur Verfügung steht, wird man eher mit einem positiven Zusammenhang rechnen.

Die regressionsanalytischen Prüfungen ergeben ein klares Bild: Mit abnehmender Selektivität der voruniversitären Mathematik- und Physikkurse sinken die mittleren Fachleistungen, *ohne* dass die Ergebnisse der Leistungsspitze beeinträchtigt würden. *Im Gegenteil:* Es deuten sich positive Auswirkungen einer verbreiterten Basis für die Erzielung von Spitzenleistungen an. Dieser positive Effekt verbesserter interner Selektion wird ebenfalls für den relativen Besuch der voruniversitären Bildungsgänge insgesamt sichtbar. Mit der Öffnung vorakademischer Bildungsgänge werden offensichtlich neue Ressourcen erschlossen, die eine verbesserte interne Lenkung und Auswahl mit positiven Auswirkungen selbst für Spitzenleistungen ermöglichen.

Befunde des internationalen Vergleichs zum Mathematikunterricht und spezifische Stärken und Schwächen von Schülerinnen und Schülern der gymnasialen Oberstufe

Aus methodischen Gründen und Gründen der Fairness wurden multiple Vergleiche unter Berücksichtigung unterschiedlicher Jahrgangsanteile durchgeführt. Auch Variationen der Unterrichtszeit, die für den Mathematikunterricht in der Oberstufe zur Verfügung steht, wurden in Rechnung gestellt. Die Befunde sind konsistent: Die Leistungen deutscher Schüler liegen im unteren Mittelbereich. Dies gilt auch für den Vergleich der Gruppe der leistungsstärksten Schüler. Bei den Unterschieden zur Schweiz, zu Frankreich, Schweden oder Slowenien handelt es sich um qualitative Niveausprünge mit praktischer Bedeutung. In paarweisen Vergleichen ausgewählter Länder wurden diese Befunde vertieft. Aus den Einzelvergleichen geht hervor, dass

Exzellenz in ganz unterschiedlicher organisatorischer Form erreicht werden kann. Erfolgreiche Programme teilen jedoch eine Gemeinsamkeit: Sie sind klar strukturiert und stellen die Kontinuität des Lernens sicher.

Neben der allgemeinen Information über die Positionierung deutscher Schüler im internationalen Vergleich ist es wichtig, zu wissen, in welchen Bereichen des mathematischen und naturwissenschaftlichen Denkens die Leistungen unserer Schüler besonders stark vom internationalen Mittel oder den Werten anderer ausgewählter Länder abweichen – wo also ihre spezifischen Stärken und Schwächen liegen. Deshalb wurden mit methodischen aufwendigeren Verfahren die Lösungen ausgewählter Testaufgaben, deren Anforderungsmerkmale bekannt sind, mit Ergebnissen aus den Ländern Österreich, Schweiz, Frankreich und Schweden verglichen. Die so genannten differentiellen Itemanalysen zeigen, dass für Abiturienten in Deutschland die relative Aufgabenschwierigkeit zunimmt, wenn diese höheren Kompetenzstufen zuzuordnen sind. *Der bedenkliche Befund lautet also: Je anspruchsvoller eine Aufgabe, umso mehr fallen die deutschen Abiturienten hinter Schülern anderer europäischer Länder zurück.*

Die von Experten eingeschätzten Anforderungsmerkmale erlauben detaillierte Angaben zu den Hintergründen: Die Ergebnisse sind umso ungünstiger, je mehr konzeptuelle Kenntnisse und qualitatives Verständnis gefordert werden, je höher in den Jahrgangsstufen die Wissensanforderungen sind, je stärker arithmetische und algebraische Fertigkeiten verlangt werden. Die relativen Schwächen der deutschen Oberstufenschüler liegen demnach sowohl im Bereich des konzeptuellen als auch des prozeduralen mathematischen Wissens. Aber auch bei Aufgaben, die Anforderungen an mathematische Modellierung, Übersetzungsleistungen und Problemlösen betreffen, zeigen sich tendenziell Schwächen deutscher Schüler. Relative Stärken sind lediglich bei Aufgaben zu verzeichnen, in denen graphische Darstellungen mit Koordinatensystemen interpretiert werden müssen bzw. bei denen bildliches Denken bedeutsam ist. Im Umgang mit visuellen Repräsentationen – und nur hier – scheint eine Stärke des deutschen Mathematikunterrichts zu liegen.

Fachleistungen im voruniversitären Physikunterricht im internationalen Vergleich

Der internationale Vergleich der Physikleistungen ergibt für Deutschland ähnliche Ergebnisse wie sie für Mathematik erzielt worden sind. Ein erster Schritt eines fairen Vergleichs ist die Gegenüberstellung der testleistungstärksten 10 Prozent der Alterskohorte. Bei diesem Vergleich liegen die Physikleistungen deutscher Schüler mit de-

nen aus Österreich, der Schweiz, Kanada und Frankreich in einem breiten mittleren Bereich. Der Abstand zu den beiden Spitzenreitern Slowenien und Schweden ist mit über einer Standardabweichung sehr groß. Eine solche Differenz signalisiert einen qualitativen Sprung im Leistungsniveau. Vergleicht man in einem zweiten Schritt die 5 Prozent testleistungsbesten Schüler der Alterskohorte – in Deutschland sind dies überwiegend Schüler des Physikleistungskurses –, ändert sich am Gesamtbild wenig, allerdings liegen die Testleistungen der deutschen Schüler jetzt im oberen Bereich des Mittelfelds.

Unter den vertiefenden Einzelvergleichen ist insbesondere der Blick auf Norwegen und Schweden lehrreich. Norwegen und Schweden haben gleichermaßen ein Gesamtschulsystem mit differenzierter Oberstufe eingerichtet. In der Sekundarstufe II werden unterschiedliche Programme, die zur Hochschulreife und/oder zum Übergang in den Beruf führen, als Wahlmöglichkeiten angeboten. In die TIMSS-Untersuchung zum voruniversitären Physikunterricht hat Schweden die Teilnehmer am naturwissenschaftlichen und technologischen Programm einbezogen, das einen deutlichen Akzent auf die naturwissenschaftliche Ausbildung legt. Norwegen rechnet zur Zielpopulation jene Schülerinnen und Schüler des allgemeinen akademischen Programms, die innerhalb dieses Bildungsgangs einen dreijährigen Physikursus belegt haben. In Schweden haben sich 16 Prozent der Alterskohorte in die beiden naturwissenschaftlich ausgerichteten Programme eingeschrieben; in Deutschland besuchen rund 9 Prozent einen Physikgrund- oder -leistungskurs. Dennoch erreichen die schwedischen Abiturienten bei nur geringfügig höherer Unterrichtszeit bereits in der Gesamtstichprobe ein Leistungsergebnis im Physiktest, das mehr als eine halbe Standardabweichung über den deutschen Befunden liegt. Bei einem Vergleich äquivalenter Jahrgangsteile erhöht sich die Differenz auf mehr als eine Standardabweichung. Dies ist ein sehr großer Unterschied mit hoher praktischer Bedeutung.

Im Vergleich zu Norwegen, das einen ähnlichen Jahrgangsteil wie das deutsche Gymnasium in den Physikursen unterrichtet, liegen die deutschen Abiturienten zwei Drittel Standardabweichungen zurück. Unterschiede dieser Größenordnung stehen für einen qualitativen Sprung im physikalischen Verständnis. Schweden und Norwegen sind gute Beispiele für die Tatsache, dass man in einem Gesamtschulsystem mit differenzierter Oberstufe, in dem für inhaltliche Konsistenz von Programmen und Kontinuität des Lernens gesorgt wird, Spitzenleistungen erreichen kann, die weit über dem Niveau gymnasialer Leistungskurse liegen. *Das für solche Leistungsergebnisse die Spezifität der curricularen Programme von ganz erheblicher Bedeutung sein dürfte, zeigt der intraschwedische Vergleich zwischen den Mathematik- und Physikleistungen von Besuchern desselben Oberstufenprogramms.* Die beiden natur-

wissenschaftlichen und technologischen Programme der schwedischen Sekundarstufe II sind – im Unterschied etwa zum mathematisch-naturwissenschaftlichen Zug des Lycée d'Enseignement Général in Frankreich – klar naturwissenschaftlich und nicht primär mathematisch orientierte Bildungsgänge. Dies bildet sich erwartungsgemäß in den internationalen Leistungsbefunden ab. Schweden erreicht überragende Ergebnisse in Physik und leicht überdurchschnittliche Resultate in Mathematik.

Ähnlich wie im Bereich der voruniversitären Mathematik wurde auch für die Physik nach Stärken und Schwächen des von deutschen Schülerinnen und Schülern erreichten Fachverständnisses gefragt. Als Vergleichsländer wurden wiederum Österreich, die Schweiz, Frankreich und Schweden herangezogen. Die Aufgabenanalysen zeigen über alle Vergleichsländer hinweg vor allem zwei Besonderheiten des Leistungsprofils der deutschen Schüler: Die deutschen Abiturienten tun sich besonders schwer mit Aufgaben, die die Überwindung typischer Fehlvorstellungen verlangen oder besondere konzeptuelle Kenntnisse voraussetzen. Vergleichsweise erfolgreich sind sie hingegen bei offenen Aufgabenformaten.

Der erste Befund lässt unmittelbar vermuten, dass es der Fachdidaktik in Deutschland nicht gelungen ist, ihre intensiven Forschungen über Schülervorstellungen in die Unterrichtspraxis hinein zu vermitteln. *In keinem der vier Vergleichsländer wirken sich Alltagsvorstellungen so stark negativ aus wie in Deutschland.* Im Vergleich zu einzelnen Referenzländern lässt sich auch bei weiteren Anforderungsmerkmalen, die mit einem qualitativen Verständnis physikalischer Konzepte und Symbole verknüpft sind, ein Leistungstief der deutschen Schüler feststellen. *Die (relative) Stärke der deutschen Oberstufenschüler liegt dagegen im formal-quantitativen Umgang mit Physik.*

9. Kurswahlen, Motivation und selbstreguliertes Lernen im Mathematik- und Physikunterricht der gymnasialen Oberstufe

Neben den Analysen der mathematischen und physikalischen Kompetenzen im engeren Sinne bildet die Untersuchung der institutionellen Einbettung des Motivations- und Lerngeschehens im Mathematik- und Physikunterricht einen weiteren Schwerpunkt der Untersuchung. Im Mittelpunkt der Analysen steht das Kurssystem der gymnasialen Oberstufe in seiner Bedeutung für motiviertes und selbstreguliertes Lernen. Folgende Fragestellungen wurden bearbeitet:

- Am Anfang stehen die Beschreibung und Analyse des Kurs- und Fachwahlverhaltens in der gymnasialen Oberstufe. Die Ergebnisse machen den breiten Kern der obligatorischen Allgemeinbildung sichtbar, der trotz aller Wahlmöglichkeiten gesichert bleibt.
- Daran schließt sich die Analyse der Kurswahlmotive und deren Bedeutung für verständnisvolles Lernen an. Aufgabe dieses Untersuchungsteils ist es unter anderem, festzustellen, ob die öffentlich immer wieder vorgetragene Vermutung, Kurswahlen in der reformierten gymnasialen Oberstufe seien vor allem durch Strategien der Anstrengungsminimierung und Punkteoptimierung beeinflusst, empirische Substanz besitzt.
- Zentral für die Bewertung einer intentionsgemäßen Umsetzung der Oberstufenkonzeption ist die Frage nach den Unterschieden zwischen Grund- und Leistungskursen. Sind Grundkurse ausgedünnte Leistungskurse oder haben sie konzeptionsgemäß jeweils eigene didaktische Gestalt gefunden, die ihren besonderen Zielsetzungen entspricht?
- Unter den Gesichtspunkten von Studierfähigkeit und Wissenschaftspropädeutik steht die Förderung der Selbstregulation des Lernens, der Bereitschaft und Fähigkeit, Verantwortung für das eigene Lernen zu übernehmen, im Mittelpunkt des Auftrags der gymnasialen Oberstufe. Die Verfügbarkeit und der Einsatz von Lernstrategien im Mathematik- und Physikunterricht sind Untersuchungsgegenstand von TIMSS/III.

Kurswahlen in der gymnasialen Oberstufe

Die gymnasiale Oberstufe ist ein komplexes System, das versucht, Verpflichtung und Wahlfreiheit in ein ausbalanciertes Verhältnis zu bringen, sodass vertiefte Allgemeinbildung und die individuelle Verantwortung von jungen Erwachsenen für ihre Lernbiographie jeweils zu ihrem Recht kommen. Bei aller Komplexität sind die strukturbildenden Elemente der gymnasialen Oberstufe klar und gering an der Zahl. Dies sind:

- die Ordnung von Grund- und Leistungskursen,
- die Erwartung der Lernangebote in Aufgabenfeldern und die Festlegung von Belegverpflichtungen sowie
- die Abiturprüfung und die Festlegung der Gesamtqualifikation.

Die Belegverpflichtungen definieren Deutsch, eine fortgeführte Fremdsprache, Geschichte oder ein gesellschaftswissenschaftliches Fach mit festem geschichtlichen Schwerpunkt, Mathematik, eine Naturwissenschaft, ein musisch-künstlerisches Fach, Sport und – je nach Ländervorschrift – Religion als verbindliche Fächer. Mit diesen sieben bzw. acht Fächern ist der Bereich der obligatorischen Bildung im internationalen Vergleich sehr breit angelegt. Darüber hinaus gibt es einschränkende Regeln für die Wahl der Leistungskurse. Ähnlich wie die curricularen Festlegungen wirkt auch die Regelung der Abiturprüfung standardisierend auf die gesamte Arbeit der Oberstufe.

Die Grundprinzipien der gymnasialen Oberstufe stützen sich wechselseitig und sorgen insgesamt für eine relativ große – von Reformern als zu weitgehend beklagte – Standardisierung des Bildungsgangs trotz der Öffnung von Wahlmöglichkeiten und Stärkung der persönlichen Verantwortung. Dafür sind aber nicht nur die Beleg- und Einbringvorschriften als solche verantwortlich. Vielmehr besitzt das System eine interne, oft übersehene Logik, die zu einer Privilegierung von Englisch, Mathematik, den Naturwissenschaften – in der Regel Biologie – und Deutsch als Kern- und Leistungsfächer führt. Fällt die Wahl der Leistungskurse auf diese Fächer, werden gleichzeitig die individuellen Freiheitsgrade bei der Wahl aller anderen Fächer, insbesondere der Grundkurse maximiert.

Tabelle 10: Leistungskurswahlen nach Fächern, Ländergruppen und Geschlecht (in % des letzten Oberstufenjahrgangs)

| Fächer | Länder | | Geschlecht | | Insgesamt |
|-------------|-------------|-------------|------------|--------|-----------|
| | Alte Länder | Neue Länder | Männer | Frauen | |
| Englisch | 36,4 | 46,1 | 32,7 | 42,6 | 38,9 |
| Mathematik | 32,4 | 40,6 | 47,1 | 26,2 | 34,4 |
| Biologie | 26,4 | 37,2 | 21,3 | 44,6 | 29,1 |
| Deutsch | 24,6 | 31,2 | 17,1 | 32,4 | 28,3 |
| Erdkunde | 12,7 | 9,7 | 14,3 | 10,4 | 11,8 |
| Geschichte | 11,0 | 5,6 | 14,4 | 7,8 | 11,2 |
| Physik | 6,2 | 12,6 | 15,0 | 3,0 | 8,2 |
| Französisch | 8,3 | 3,0 | 2,8 | 9,5 | 6,9 |
| Kunst | 6,4 | – | 3,8 | 6,1 | 5,5 |
| Chemie | 5,4 | 5,1 | 7,2 | 4,2 | 5,4 |
| Sozialkunde | 6,6 | – | 4,8 | 3,0 | 4,9 |
| Russisch | 0,6 | 4,0 | 0,8 | 2,0 | 1,4 |

IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

Tabelle 10 zeigt, dass Englisch, Mathematik, Biologie und Deutsch tatsächlich zahlenmäßig die Gewinner der Oberstufe sind. Auf diese vier Fächer konzentrieren sich die Wahlen der Leistungskurse, wobei es charakteristische Unterschiede zwischen alten und neuen Ländern sowie zwischen Männern und Frauen gibt. Unter den Naturwissenschaften hat sich Biologie zur schulischen Leitdisziplin entwickelt – auf Kosten von Physik und insbesondere Chemie (über mögliche Ursachen wird weiter unten berichtet). 65 Prozent aller Leistungskurskombinationen entfallen auf Verbindungen der ersten sieben in Tabelle 10 aufgeführten Fächer. Kaum ein europäisches Land besitzt in den voruniversitären Bildungsgängen eine vergleichbar hohe Standardisierung des Programms. Von Beliebigkeit der Wahlen oder einer Konzentration auf so genannte „leichte“ Fächer kann keine Rede sein.

Kurswahlmotive

In TIMSS/III wurden Kurswahlmotive retrospektiv erfasst. Unterschieden wurden folgende Wahlmotive: Entfaltung der eigenen Kompetenz, Interesse, Beruf, Kontakt zu Mitschülern, Punkteoptimierung, Vermeidung von Lehrern und Wahlbeschränkung an der Schule. Die Analyse der Wahlmotive ergibt eine klare und über die Fächer hinweg stabile zweifaktorielle Struktur.

Markier-Items der ersten Hauptkomponente sind Kompetenzerfaltung und Interesse, die durch die Berufsperspektive ergänzt werden. Bemerkenswerterweise ist diese Motivlage auch mit dem Ziel, ein möglichst gutes Punktergebnis im Abitur zu erreichen, zu vereinbaren. Dies ist durchaus im Sinne der Idee der gymnasialen Oberstufe, nach der dem jungen Erwachsenen verstärkte Verantwortung für den eigenen Bildungsgang in die Hände gegeben werden soll. Dieses Motiv ist jedoch nicht dominant. Die zweite Hauptkomponente ist durch soziale Wahlmotive bestimmt – sei es, dass man bestimmte Lehrer vermeiden oder den engeren Kontakt zu Mitschülern erhalten will. Diese sozialen Motive werden durch die strategische Überlegung ergänzt, Punkte zu optimieren; sie können aber auch mit Wahlbeschränkungen an der jeweiligen Schule einhergehen. Der erste Faktor ist für die Kurs- und Fachwahl deutlich am wichtigsten.

In allen Fächern sind die Motive „Kompetenz“ und „Interesse“ gefolgt von „Berufsperspektiven“ prävalent. Die Kompetenz- und Interessenwahlen sind am ausgeprägtesten in den ästhetisch-expressiven Fächern Kunst, Musik und Sport. Die Unterschiede zwischen den Fächern sind gering. Auch in den Fächern Mathematik und Physik sind Leistungskurswahlen primär Kompetenz- und Interessenwahlen. Im Fach Mathematik berichten zwischen 70 und 80 Prozent der Befragten, dass für die

Wahl ihres Leistungskurses die Gesichtspunkte, persönliche Leistungsstärken entfalten oder persönlichen Interessen nachgehen zu können, wichtig oder sehr wichtig sind. Für die Wahl des Leistungskurses Physik ist dieses Bild noch ausgeprägter.

Motivations- und Kompetenzprofile in Grund- und Leistungskursen

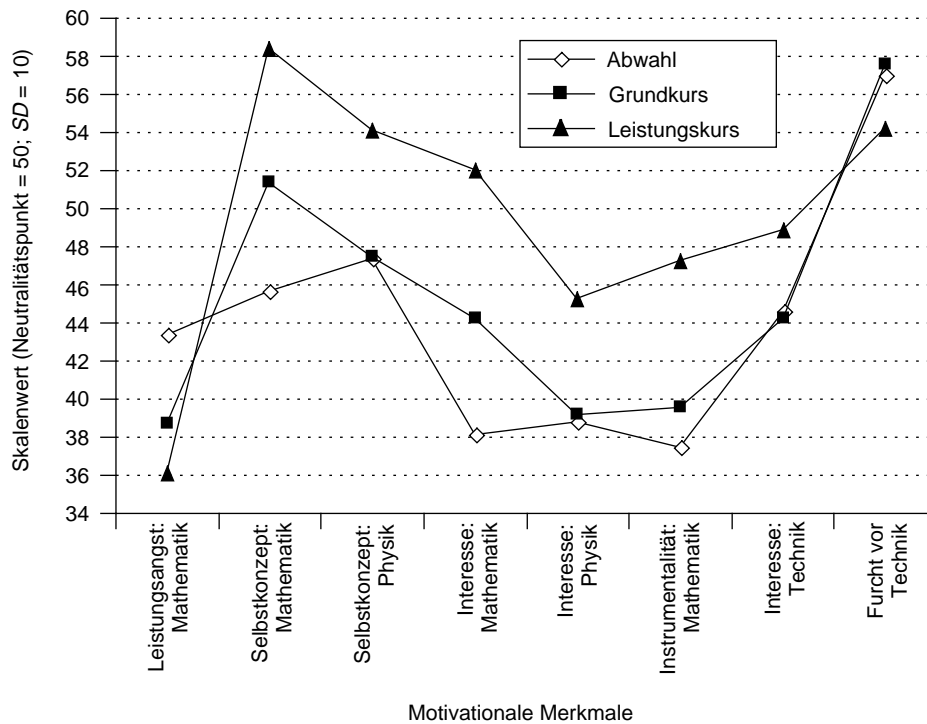
Die Prävalenz der sachbezogenen Motive bei den Kurswahlen führt hinsichtlich Vorwissen und motivationalen Orientierungen der Kursteilnehmer zu sehr unterschiedlichen Arbeitsbedingungen in den einzelnen Veranstaltungen.

Hinsichtlich der motivationalen Orientierungen unterscheidet sich die Situation im obligatorischen Fach Mathematik deutlich von der im hoch selektiven Fach Physik. Die Teilnehmer am mathematischen Leistungskurs wissen, was sie können: Sie schätzen ihre fachliche Befähigung sehr hoch ein und haben gleichzeitig die geringste Leistungsangst. Das Interesse an Mathematik liegt deutlich im positiven Bereich. Hinsichtlich der praktischen Nützlichkeit von Mathematik herrscht Unentschiedenheit. Gegenüber technischen Entwicklungen ist Skepsis erkennbar. Das Motivationsprofil des Grundkurses weist einen ähnlichen Verlauf aus; es ist gegenüber dem Leistungskursprofil im Wesentlichen nur im Niveau versetzt. Im mathematischen Grundkurs ist die Distanz zur Technik besonders ausgeprägt. Differenziertere Analysen zeigen, dass dennoch im Leistungskurs Mathematik eine nicht unerhebliche Anzahl von Schülern zu finden ist, die das Motivationsprofil von Grundkursteilnehmern aufweisen. Dies sind die Kosten der sich auf das Fach Mathematik konzentrierenden Leistungskurswahlen.

Im Fach Physik setzen sich die Oberstufenschülerinnen und -schüler, die sich gegen das Fach entschieden haben, in ihren Einstellungen deutlich von denen ab, die das Fach beibehalten haben. Die Gruppe der Nichtwähler ist durch niedrige Kompetenzüberzeugung, Interesselosigkeit und Technikdistanz gekennzeichnet. Grund- und Leistungskursteilnehmer unterscheiden sich nur geringfügig.

In einem weiteren Schritt wurde analysiert, ob sich kursspezifische Wissensprofile identifizieren lassen, die als Ergebnis der curricularen und didaktischen Konzeption von Grund- und Leistungskursen verstanden werden können. Selbstverständlich unterscheiden sich Grund- und Leistungskurse im mittleren Kompetenz*niveau* bei allerdings breiten Überlappungen der Leistungsverteilungen. Für diese Niveauunterschiede sind sowohl die Eingangsselektion als auch der erweiterte Unterricht verantwortlich, ohne dass man die Effekte im Rahmen der Querschnittsunter-

Abbildung 10: Motivationsprofile nach Kurswahl im Fach Mathematik



IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

suchung trennen könnte. *Inhaltliche* Wissensprofile ließen sich jedoch auch auf der Basis der Analyse ausgewählter Aufgaben nicht ausmachen.

Motivation, Kurswahlen und Fachleistungen: Ein multivariates Erklärungsmodell

Es wurde ein theoretisches Modell entwickelt, das die simultane Analyse der Zusammenhänge von Motivation und Kurswahl einerseits und Motivation und Fachleistung andererseits erlaubt. Ausgangspunkt waren die Annahmen, dass das Selbstkonzept der fachspezifischen Befähigung, das Sachinteresse, die Nützlichkeits einschätzung des Sachgebiets und die Bedeutung, die subjektiv wichtige Personen dem Lernen von Mathematik und Physik beimessen, sowohl die Kurswahl als auch ver-

mittelt über die Kurswahl, die Leistungsergebnisse beeinflussen. Für Mathematik zeigen die Ergebnisse, dass die beiden wichtigsten Einflussgrößen das Selbstkonzept der Fachkompetenz und das Sachinteresse darstellen. Sie beeinflussen die Kurswahl und die Fachleistung direkt und indirekt. Nützlichkeitsabwägungen und das Ansehen der Mathematik in der Familie beeinflussen die Kurswahl. Im Fach Physik reduzieren sich die signifikanten Einflussgrößen auf die Selbsteinschätzung der Fachkompetenz und das Sachinteresse.

Diese Befunde haben große praktische Bedeutung. Sie belegen, dass die Entscheidungen über Kurswahlen und letztlich auch über die erreichbaren Fachleistungen bereits im Unterricht der Mittelstufe fallen. Über das Weiterlernen in der gymnasialen Oberstufe entscheidet ein verständnisorientierter Unterricht in der Mittelstufe, der subjektives Kompetenzerleben durch Lernfortschritte vermittelt und individuell sinnstiftend erlebt wird, sodass sich Interessen entfalten können. In den interessen- und kompetenzgesteuerten Kurswahlen wird nicht zuletzt das Ergebnis des Mittelstufenunterrichts sichtbar.

Selbstreguliertes Lernen

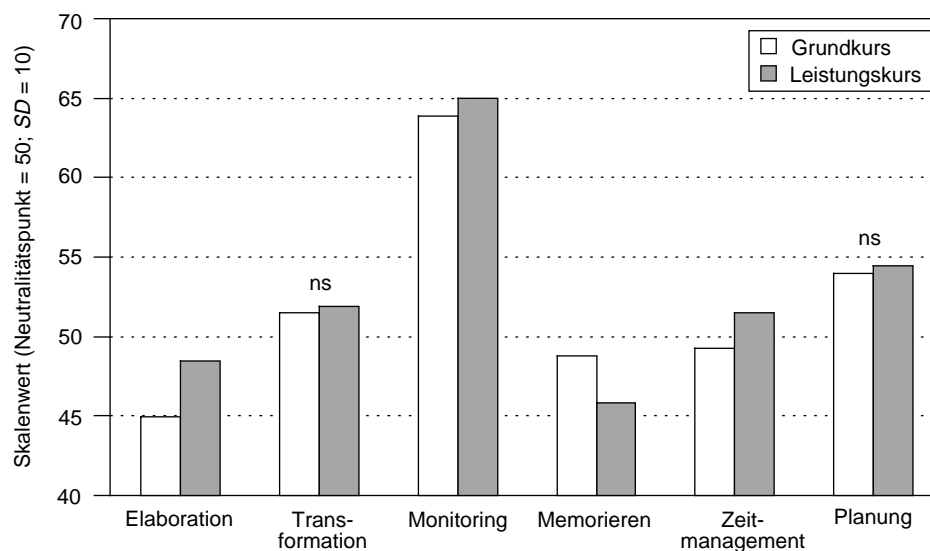
Unter den Gesichtspunkten von Studierfähigkeit und Wissenschaftspropädeutik steht die Förderung der Selbstregulation des Lernens, das heißt der Bereitschaft und Fähigkeit, Verantwortung für das eigene Lernen zu übernehmen, dieses ökonomisch zu planen, selbstständig zu steuern und zu überwachen und im praktischen Vollzug gegen konkurrierende Intentionen abzuschirmen, im Mittelpunkt des Auftrags der gymnasialen Oberstufe, auch wenn die Zugangswege zu diesen Arbeitsformen bereits in der Mittelstufe angebahnt werden müssen. Selbstreguliertes Lernen lässt sich als zielorientierter Prozess des aktiven und konstruktiven Wissenserwerbs beschreiben, der auf dem reflektierten Zusammenspiel von kognitiven und motivational-emotionalen Ressourcen einer Person beruht.

Im Rahmen von TIMSS/III wurde versucht, zentrale Komponenten dieses Modells in der nationalen Zusatzstudie zu erfassen. Es wurden fachspezifische Lernstrategien und mit dem Sachinteresse die motivationale Orientierung in bereichsspezifischer Ausprägung erhoben. Man unterscheidet drei große Gruppen von Lernstrategien: kognitive und metakognitive Strategien sowie Strategien des Ressourcenmanagements. Zu den kognitiven Strategien zählen Memorier-, Elaborations- und Transformationsstrategien. Memorierstrategien sind vor allen Dingen Wiederholungsstrategien, die dazu dienen, Neugelertes im Arbeitsspeicher zu halten, aber auch die Übernahme der Informationen in das Langzeitgedächtnis zu unterstützen. Eine

zweite Untergruppe der kognitiven Strategien bilden so genannte Elaborationsstrategien, die durch sinnstiftendes Vorgehen ausgezeichnet sind. Elaborationsstrategien dienen dazu, innerhalb neu zu lernender Stoffe Sinnstrukturen herauszuarbeiten, Lerninhalte mit bereits gespeichertem Wissen möglichst sinnvoll und dicht zu vernetzen und die Übertragbarkeit des Neugelerten auf andere Kontexte zu erproben. Diese Strategien unterstützen insbesondere Encodierungs-, Erwerbs- und Transferprozesse. Die zweite große Gruppe der Lernstrategien bilden die so genannten metakognitiven Strategien, zu denen Strategien der Planung, Überwachung und Regulation des eigenen Lernens gehören. Eine flexible Verfügung über metakognitive Strategien gilt als entscheidende Voraussetzung selbstgesteuerten Lernens. Strategien des Ressourcenmanagements schließlich werden als Stützstrategien bezeichnet. Zum internen Ressourcenmanagement werden die Überwachung von Anstrengung und Aufmerksamkeit sowie die planvolle Nutzung der Lernzeit gerechnet.

Die Nutzung effizienter und verständnisfördernder Lernstrategien hat einmal für das Erreichen guter Leistungsresultate im Unterricht der gymnasialen Oberstufe instrumentelle Bedeutung, darüber hinaus sind die Verfügbarkeit von Lernstrategien und ihr situationsadäquater Einsatz als Elemente selbstregulierten Lernens ein

Abbildung 11: Lernstrategien in Mathematik nach Kursniveau



IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

eigenständiges Ziel des Unterrichts, das über diesen hinaus auf selbstständiges Weiterlernen verweist. Geht man von den in sich konsistenten Beschreibungen des Lernverhaltens aus, das uns Schülerinnen und Schüler der gymnasialen Oberstufe für die Vorbereitung auf Mathematik und Physik berichten, so wird dieses Ziel je nach Strategietyp unterschiedlich gut erreicht. Nach den Selbstberichten sind Oberstufenschüler im Allgemeinen gut in der Lage, ihren Arbeitsprozess während der Vorbereitung auf den Unterricht systematisch zu planen und zu überwachen. Sie regulieren ihre Aufmerksamkeit und prüfen, ob sie das Gelernte auch tatsächlich behalten haben. Dagegen sind verständnisorientierte Wissenserwerbsstrategien, mit denen man versucht, sinnstiftende Bezüge innerhalb des neuen Stoffs herzustellen, diesen systematisch mit dem Vorwissen zu verbinden und das neu Gelernte in variierenden Kontexten zu erproben, selten anzutreffen. Besonders auffällig ist das niedrige Niveau des berichteten Einsatzes von Elaborationsstrategien bei der Vorbereitung auf den mathematischen Grundkurs, in dem gleichzeitig das Auswendiglernen stärker zum Zuge kommt.

10. Soziale Vergleichsprozesse im Kurssystem der gymnasialen Oberstufe

Leistungsgruppierungen stellen ein zentrales Merkmal des deutschen Schulsystems dar. Auch innerhalb des Gymnasiums findet nach Eintritt in die Oberstufe eine Differenzierung statt, nämlich durch die Wahl unterschiedlicher Kursniveaus (Grund- vs. Leistungskurs). Die Bildung leistungshomogener Gruppen, ob in Kurssystemen oder durch unterschiedliche Schulformen im Sekundarbereich I, kann insbesondere im oberen Leistungsbereich negative Konsequenzen für den weiteren Wissenserwerb verhindern. Den besonderen Fördereffekt des Gymnasiums in der Sekundarstufe I belegen Analysen der Daten der Kohorten-Längsschnittstudie „Bildungsprozesse und psychosoziale Entwicklung im Jugendalter (BIJU)“.

Schüler im unteren Leistungsbereich am Ende der Primarstufe profitieren vom Übergang in die differenzierte Sekundarstufe offenbar weniger im Wissenszuwachs. Im Bereich der psychosozialen Entwicklung sind aber durchaus differentielle Effekte zu erwarten, da der ungünstige Leistungsvergleich mit sehr viel besseren Schülern entfällt, was positive Auswirkungen auf das Selbstwertgefühl bzw. die Selbstwahrnehmungen eigener Begabung hat. Für leistungsstarke Schüler hat der Übergang auf das Gymnasium hinsichtlich ihrer selbst wahrgenommenen Fähigkeiten den entgegengesetzten Effekt. Gehörten diese Schüler in der Grundschule noch zu den Besten, so erleben sie nach dem Übergang auf das Gymnasium, dass viele andere in der

Leistung ebenbürtig oder besser sind. Soziale Vergleiche führen hier eher zu einem Absinken fähigkeitsbezogener Selbstkonzepte und des Selbstwertgefühls. Dieser Prozess mündet oft darin, dass das mittlere Fähigkeitsselftkonzept oder das Selbstwertgefühl auf den verschiedenen Schulformen im Laufe der Sekundarstufe I konvergiert. Das Bewusstsein von Jugendlichen, mit dem Gymnasium den leistungsstärksten Zweig zu besuchen, kann diese Kosten der Leistungsgruppierung im oberen Bereich offenbar nicht kompensieren.

Das Phänomen, wonach systematische Leistungsgruppierungen Effekte auf selbstbezogene Kognitionen haben können, wird als *big-fish-little-pond-effect* (BFLPE) beschrieben. Zwei Schüler mit gleicher Leistungsfähigkeit, die aber auseinander liegende Schulen mit unterschiedlichen Leistungsniveaus besuchen, sollten demzufolge ein unterschiedliches Selbstkonzept eigener Begabung (= fachbezogenes Selbstvertrauen) entwickeln, das heißt, der Schüler (*big fish*) in der schwächeren Schule (*in a little pond*) sollte ein höheres Begabungselftkonzept als der entsprechende Schüler in der leistungsstärkeren Schule (*in a big pond*) haben.

Dieser BFLPE muss sich im deutschen Schulsystem natürlich nicht nur zwischen Schulformen abspielen, vielmehr ist er auch innerhalb von Schulformen zu erwarten, sofern die Leistungen zwischen den Schulen variieren, was in Deutschland typischerweise der Fall ist. Zwei gleich starke Realschüler, einer auf einer leistungsstärkeren, einer auf einer leistungsschwächeren Schule, sollen sich ebenfalls bezüglich ihrer Fähigkeitselftbilder im Sinne der obigen Ausführungen unterscheiden.

Ein offenes Problem, speziell bezogen auf die gymnasiale Oberstufe, ist, welche Effekte die erneute Fähigkeitsgruppierung nach Leistungs- und Grundkursen im Sekundarbereich II auf psychosoziale Variablen hat. Ziel dieses Kapitels war es daher, anhand der TIMSS-Oberstufenkohorte den akademischen Kontext von jungen Erwachsenen in seinen Auswirkungen auf das fachbezogene Selbstvertrauen (Selbstkonzept der Begabung) und die Lernmotivation in Form des Interesses zu untersuchen. Berücksichtigt wurde das Fach Mathematik.

Die in TIMSS durchgeführten Analysen ergeben in der Tat einen negativen BFLPE der mittleren Leistungsstärke von Schulen auf das Interesse und Selbstkonzept der Begabung, das heißt für das fachbezogene Selbstvertrauen und Interesse, dass es bei Schülern identischer individueller Leistungsstärke auf besonders leistungsstarken Gymnasien geringer ausfällt als an schwächeren Gymnasien.

Neben der mittleren Schulleistung hat das Kursniveau aber keinen zusätzlich negativen Effekt, vielmehr zeigt sich ein positiver Einfluss auf das Selbstvertrauen und

Interesse. Offenbar sind sich Leistungskurschüler der besonderen Leistungsstärke des Leistungskurses bewusst, und diese positive Bewertung der Gruppe schlägt sich auch auf die eigene Fähigkeitseinschätzung durch.

Bezüglich des Interesses ist der positive Einfluss des Kursniveaus noch stärker. Dieser Befund macht deutlich, dass ein auf tieferes Verständnis abzielender Unterricht im Leistungskurs vermutlich interessefördernd ist. Die intensivere Auseinandersetzung mit mathematischen Inhalten öffnet bei Schülern den Blick für die besondere Relevanz der Mathematik bei der Lösung vieler angewandter Probleme. Dieser Bezug zum Alltag ist in der Regel auch interessefördernd.

Praktische Implikationen

Im Hinblick auf den BFLPE ist man gut beraten, mit voreiligen Schlüssen aus unseren Befunden für die Schulpraxis vorsichtig zu sein. Die generelle Abschaffung von Leistungsdifferenzierung lässt sich jedenfalls mit den vorgestellten Analysen aus TIMSS *nicht* rechtfertigen. Vermutlich müsste auch die Art und Weise Berücksichtigung finden, wie die sozialen Vergleiche, die den BFLPE ja mittelbar erzeugen, in den Interaktionsprozess der Klasse eingebunden sind. Vor dem Hintergrund des jeweiligen Klassenklimas können kompetitive Elemente entweder spielerisch-sportlichen Charakter annehmen oder aber aggressiv-abgrenzend eingefärbt sein. So wurde argumentiert, dass ein sehr aggressiv-kompetitives Schul- bzw. Klassenklima langfristig negative Konsequenzen nicht nur im psychosozialen, sondern auch im kognitiven Bereich hat. Unterricht, ob in der Mittel- oder Oberstufe kann natürlich Bezugsgruppeneffekten entgegensteuern, indem zum Beispiel Rückmeldungen stärker mit individueller Bezugsnormorientierung (BNO) gegeben werden. Hier bieten sich zukünftige Studien an, die untersuchen, ob der BFLPE in Klassen bzw. Kursen, deren Lehrkraft eine individuelle BNO verfolgt, verschwindet, wohingegen er in Klassen mit sozialer BNO umso stärker auftritt.

11. Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht

Die Vermittlung adäquater epistemologischer Überzeugungen über die Grundlagen und Leistungsfähigkeit spezifischer Fächer und deren Grenzen gehört zum zentralen Anliegen des Oberstufenunterrichts. Unter den Begriffen „epistemologische Überzeugungen“ (*Epistemological Beliefs*) bzw. „Weltbilder“ (*World Views*) werden jene

Vorstellungen und subjektiven Theorien subsumiert, die Personen über das Wissen und den Wissenserwerb generell oder in spezifischen Domänen entwickeln. Sie thematisieren, wie in beiden Disziplinen Erkenntnisse gewonnen werden, welche praktische Relevanz die Erkenntnisse besitzen und vieles mehr.

Ausgangspunkt der rezenten Arbeiten zu epistemologischen Überzeugungen ist die Annahme, dass diese intuitiven Theorien die Art der Begegnung mit der erkennbaren Welt vorstrukturieren. Sie beeinflussen Denken und Schlussfolgern, Informationsverarbeitung, Lernen, Motivation und schließlich auch die akademische Leistung. Verschiedene Autoren sprechen von epistemologischen Orientierungen als mentalen Prozessen höherer Ordnung, die kognitive Vorgänge steuern. Bei der Mehrzahl der Autoren verbindet sich mit der Analyse epistemologischer Überzeugungen eine entwicklungspsychologische Perspektive, nach der es einen altersgradierten Komplexitätsgewinn gibt. Dieser Prozess kann in spezifischer Weise durch Schulunterricht gestützt oder behindert werden.

Wenn epistemologische Überzeugungen in den Fachdidaktiken behandelt werden – wie es in den Mathematik- und Naturwissenschaftsdidaktiken der Fall ist –, so geschieht dies nicht nur unter einem instrumentellen Blickwinkel, um Lernprozesse zu befördern, sondern immer auch mit einem bildungstheoretischen Anspruch. Denn die fachbezogenen intuitiven Theorien über Wissen und Wissensgenese geben Antworten darauf, welche Fragen in einem Fachgebiet überhaupt legitimerweise gestellt und mit welchen Verfahren und mit welcher Sicherheit beantwortet werden können. Die Klärung dessen, was Fächer oder Disziplinen in ihrer spezifischen Fokussierung leisten und wo ihre Grenzen liegen, gehört unserer Überzeugung nach zum Bildungsauftrag eines jeden Schulfachs. Im pädagogischen Bereich haben epistemologische Überzeugungen also immer gleichzeitig instrumentelle und substantielle Bedeutung. Deshalb werden auch die *Epistemological Beliefs* zum Kern der mathematischen und naturwissenschaftlichen Grundbildung (*Literacy*) gezählt.

Für beide in Deutschland untersuchten Altersgruppen von TIMSS (Population II und III) wurde der Versuch unternommen, mithilfe kurzer Fragebogen Aspekte des mathematischen und naturwissenschaftlichen Weltbilds von Schülern zu erfassen. Die detaillierte Beschreibung der berücksichtigten Konstrukte und ihrer Operationalisierungen befindet sich im Kapitel VI des zweiten Bandes. Die wesentlichen Befunde in der Oberstufenkohorte sollen im Folgenden geschildert werden.

Zentrales Moment des mathematischen Weltbilds von Schülerinnen und Schülern der gymnasialen Oberstufe ist die schematisch-algorithmische Ausrichtung von Mathematik und des Mathematikunterrichts. Die große Mehrheit der Befragten

stimmt Aussagen zu, wie: „Mathematik ist Behalten und Anwenden von Definitionen, Formeln, mathematischen Fakten und Verfahren“ oder „Mathematik betreiben heißt: allgemeine Gesetze und Verfahren auf spezielle Aufgaben anwenden“. Damit korrespondiert, dass Oberstufenschülern eine relativistische wissenschaftstheoretische Position, die den konstruktiven und prozessualen Charakter von Mathematik betont, deren Weiterentwicklung von Kreativität und Imagination abhängig ist, nicht vertraut ist. Allerdings ist die instrumentelle Bedeutung von Mathematik, Wirtschaft oder Technik fast jedem Gymnasiasten bewusst. Das Grundmuster der epistemologischen Überzeugungen ist sowohl im Grund- als auch im Leistungskurs anzutreffen, auch wenn das mathematische Weltbild von Leistungskurschülern etwas differenzierter ist.

Auch im physikalischen Weltbild von Gymnasiasten der Oberstufe lässt sich eine weitgehend geteilte Grundvorstellung identifizieren, in der sich die ontologische Überzeugung einer allmählichen Entdeckung des Bauplans des Universums mit der Vorstellung vom Systemcharakter physikalischen Wissens verbindet. Danach existieren in der Natur physikalische Gesetze, die von den Physikern Schritt für Schritt entdeckt werden. Physikalische Theorien systematisieren menschliche Erfahrungen, die vor allem im Experiment gemacht werden. Physik ist danach eine Leistung des Entdeckens. Diese Basisvorstellung wird in der internationalen Forschungsliteratur als traditionell-empiristisches Wissenschaftsbild bezeichnet. Die Vorstellung von Wissenschaft als einer Konstruktionsleistung ist in diesem Weltbild ein Fremdkörper.

Ähnlich wie im Fach Mathematik ist die gesellschaftliche Relevanz von Physik unbestritten. Nach der Überzeugung der großen Mehrheit der Oberstufenschülerinnen und -schüler ist Physik der Motor der technologischen Entwicklung und zielt darauf, die praktischen Probleme der Menschheit zu lösen. Grund- und Leistungskurschüler unterscheiden sich in ihrer physikalischen Weltsicht systematisch. Die Richtung der Unterschiede ist jedoch bemerkenswert. Je länger und intensiver sich Schüler in der Schule mit Physik beschäftigen, desto ausgeprägter scheinen empiristische Wissenschaftsvorstellungen zu sein und desto weiter reichen die Erkenntnis- und Wahrheitsansprüche, die mit Physik verbunden werden. Dieses Ergebnis widerspricht offensichtlich fachdidaktischen Zielvorstellungen.

Unbestritten ist in der rezenten Forschungsliteratur die Bedeutung epistemologischer Überzeugungen für leistungsthematisches Handeln. Sie beeinflussen Denken und Schlussfolgern, Informationsverarbeitung, Lernen, Motivation und schließlich auch die akademische Leistung. Die anhand von Strukturgleichungsmodellen vorgenommene Überprüfung des Zusammenhangs von epistemologischen

Überzeugungen einerseits und motivationalen Orientierungen, Nutzung von Lernstrategien und Fachleistungen andererseits zeigt eine hohe Vergleichbarkeit der Befunde über beide Fächer hinweg. Die Skalen zu den mathematischen Weltbildern weisen direkte oder indirekte, das heißt über Mediatoren vermittelte Effekte auf die Mathematikleistung auf. Erwartungsgemäß sind die Einflüsse einer realistischen und schematischen Mathematikkonzeption negativ. Der Einfluss der schematischen Konzeption ist über Fachinteresse und Memorierstrategien vermittelt. Personen mit der Überzeugung, Mathematik sei das bloße Anwenden von Lösungsalgorithmen auf vorgegebene Aufgaben, sind weniger interessiert, verwenden vermehrt Oberflächenstrategien beim Lernen und erreichen niedrigere Leistungen. Ebenfalls erwartungsgemäß sind die positiven Einflüsse der beiden anderen Dimensionen des mathematischen Weltbilds. Personen mit relativistischer Konzeption äußern größeres Interesse an Mathematik, und als mögliche Folge ausgeprägteren Interesses liegt auch ihre Leistung höher.

Analoge Zusammenhänge zwischen epistemologischen Überzeugungen, Mediatoren und Fachleistungen lassen sich für das Fach Physik nachweisen. Die Einflüsse einer realistischen und schematischen Physikkonzeption auf das Fachverständnis sind negativ und dabei teils über das Interesse, teils über Lernstrategien vermittelt. Die beiden Dimensionen, in denen die Überzeugung von der instrumentellen Bedeutung der Physik thematisiert wird, haben sowohl einen über das Interesse vermittelten indirekten als auch einen direkten Einfluss auf die Fachleistung. Insgesamt zeigen die Analysen, dass epistemologische Überzeugungen ein wichtiges, bislang nicht ausreichend gewürdigtes Element motivierten und verständnisvollen Lernens in der Schule darstellen.

12. Unterrichtsqualität und Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe

Ein wichtiger Bereich der TIMSS/III-Analysen widmet sich der didaktischen Gestaltung des Mathematik- und Physikunterrichts in der gymnasialen Oberstufe und der Frage, ob unterschiedliche Zielsetzungen der Schule etwa kognitiver oder motivationaler Art miteinander verträglich sind und zur gleichen Zeit gleich gut erreicht werden können. Drei Fragestellungen strukturieren diesen Untersuchungsteil:

- Lassen sich für Grund- und Leistungskurse charakteristische Unterrichtsmuster identifizieren und wie gut lassen sich aufgrund von didaktischen Merkmalen des

Mathematik- und Physikunterrichts Fachkompetenz und Fachverständnis vorhersagen?

- Nicht selten wird die Ansicht vertreten, dass ein lernintensiver Unterricht möglicherweise Kosten im motivationalen oder sozialen Bereich habe. Kompetenzgewinne könnten dann etwa mit Interessenverlust oder erhöhtem Wettbewerb in der Klassengemeinschaft bezahlt werden. Im Rahmen der nationalen Zusatzuntersuchung wurden systematisch mögliche Verträglichkeitsprobleme unterschiedlicher Zielsetzungen untersucht.
- Im Anschluss an diesen Untersuchungsteil wurde ferner die allgemeine psychosoziale Situation von Oberstufenschülerinnen und -schülern im Kurssystem analysiert.

Verständnisvolles Lernen im Fachunterricht

Im Rahmen von TIMSS/III wurde eine Schülerbefragung zum Mathematik- und Physikunterricht durchgeführt, die theoretisch an Konzeptionen verständnisvollen Lernens anschließt. Die Schülerinnen und Schüler wurden gebeten anzugeben, wie häufig bestimmte Tätigkeiten in ihrem Mathematik- bzw. Physikunterricht auftreten, die Hinweise auf didaktische Akzentsetzungen geben. Die gewählten Indikatoren lassen vorsichtige Rückschlüsse auf Verständnisorientierung des Unterrichts durch anspruchsvolle kognitive Eigentätigkeit von Schülern, die Auftretenshäufigkeit von Modellbildung und Anwenden des Gelernten in neuen Kontexten, den Stellenwert des naturwissenschaftlichen Experiments, die Bedeutung der Schulung von Fertigkeiten und schließlich auch auf die Rezeptivität des Unterrichts zu.

Die Analysen ergeben für den Mathematikunterricht folgendes Bild: Aus Schülersicht ist der Mathematikunterricht der gymnasialen Oberstufe bemerkenswert variationsarm. Vorherrschend sind zwei miteinander korrespondierende Schritte: Sobald die Lehrkraft einen mathematischen Gedankengang entwickelt und vorgestellt hat, folgen in der Schülerarbeitsphase das Lösen von Gleichungen und die Übung von Rechenfertigkeiten. Inwieweit die Entwicklung des mathematischen Themas allein in der Hand der Lehrkraft liegt oder primär im lehrergeleiteten Unterrichtsgespräch erfolgt, kann aufgrund des TIMSS/III-Fragebogens nicht entschieden werden. Insgesamt nehmen Schüler den Mathematikunterricht jedoch als rezeptive und fertigkeitenorientierte Veranstaltung wahr. Variabilität lässt sich am ehesten in der Dimension der Verständnisorientierung von Aufgabenstellungen erkennen.

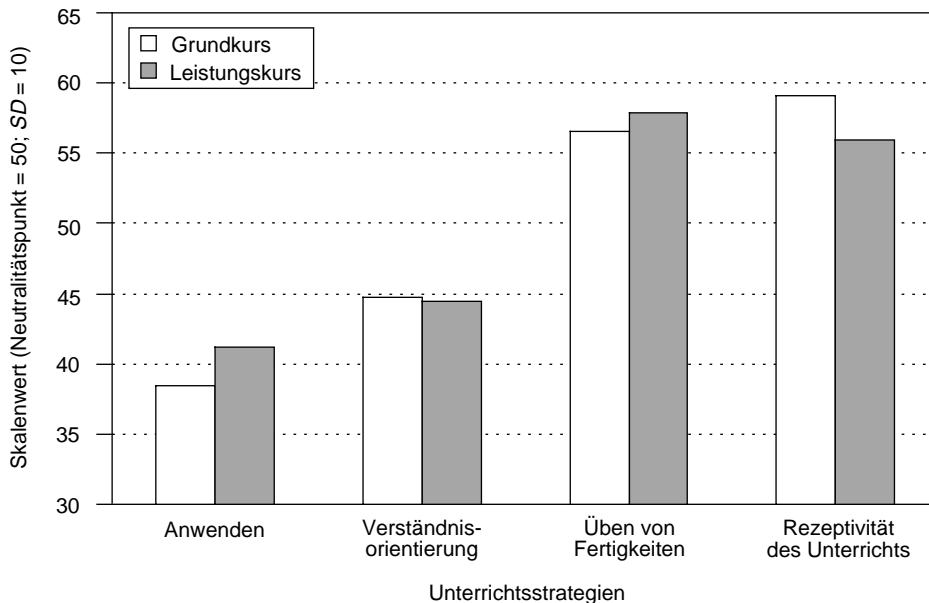
Die didaktischen Grundmuster werden über Kursniveaus hinweg durchgehalten. Es ist nicht zu erkennen, dass Grund- und Leistungskurse differentiellen didaktischen Konzeptionen folgten. Dies gilt sowohl in curricularer als auch methodischer Hinsicht. Grundkurse scheinen, wenn man es salopp ausdrückt, ausgedünnte Leistungskurse zu sein und überdies in methodischer Hinsicht schlechter bedacht zu werden. Sie sind noch etwas stärker rezeptiv und fertigungsorientiert angelegt als Leistungskurse.

Infolge der geringen didaktischen Variabilität des Unterrichts erklären die erfassten Merkmale der Unterrichtsgestaltung auch nur einen relativ geringen Anteil der Leistungsvarianz zwischen Kursen. Als wichtigste Unterrichtsdimension erwies sich eine verständnisorientierte Unterrichtsführung, die besonderen Wert auf kognitiv anspruchsvolle Tätigkeiten von Schülerinnen und Schülern legt. In einem Unterricht, in dem die Schüler diese Dimension besonders ausgeprägt wahrnehmen, kommen häufiger Nicht-Routine-Aufgabenstellungen vor, die Problemlöseleistungen erfordern, legt der Lehrer Wert auf die Erklärung und Analyse von mathematischen Zusammenhängen und spielt schließlich mathematisches Modellieren eine größere Rolle. Damit wird eine Dimension im Kernbereich des Mathematikunterrichts erfasst, die man als durch anspruchsvolle Aufgaben initiierte kognitive Eigentätigkeit von Schülern bezeichnen könnte. Das Ausmaß der Rezeptivität des Unterrichts steht dagegen in einem negativen Zusammenhang mit den erzielten Leistungsergebnissen. In einem primär rezeptiv angelegten Unterricht verfolgen Schülerinnen und Schüler, wie der Lehrer oder die Lehrerin einen mathematischen Gedankengang entwickelt, übertragen den Tafelanschrieb in ihr Schulheft und memorieren Regeln und Verfahren. Nach den Schülerangaben bilden diese Tätigkeiten ein Grundgerüst des Mathematikunterrichts. Abbildung 12 gibt einen Überblick über die Verteilung der unterschiedlichen Unterrichtsstrategien.

Um Anhaltspunkte für die didaktische Gestaltung des Physikunterrichts zu gewinnen, sind den Teilnehmern an Physikkursen ähnliche Fragen wie für Mathematik gestellt worden. Die zentralen didaktischen Dimensionen, für die Indikatoren gewonnen werden sollten, sind wiederum Verständnisorientierung, Rezeptivität und Anwendungsorientierung des Unterrichts. Ergänzend kamen für das Fach Physik eine Reihe von nationalen Fragen nach der experimentellen und erfahrungsorientierten Ausrichtung des Unterrichts hinzu.

Die Analysen zeigen folgendes Bild: Im Physikunterricht ist die didaktische und methodische Variabilität aus Schülersicht größer als im Mathematikunterricht. Dennoch lässt sich auch im Physikunterricht ein Muster identifizieren, das den Unterricht über Kursniveaus hinweg prägt. Der Physikunterricht der gymnasialen Ober-

Abbildung 12: Unterrichtsstrategien in Mathematik nach Kursniveau



IEA. Third International Mathematics and Science Study.

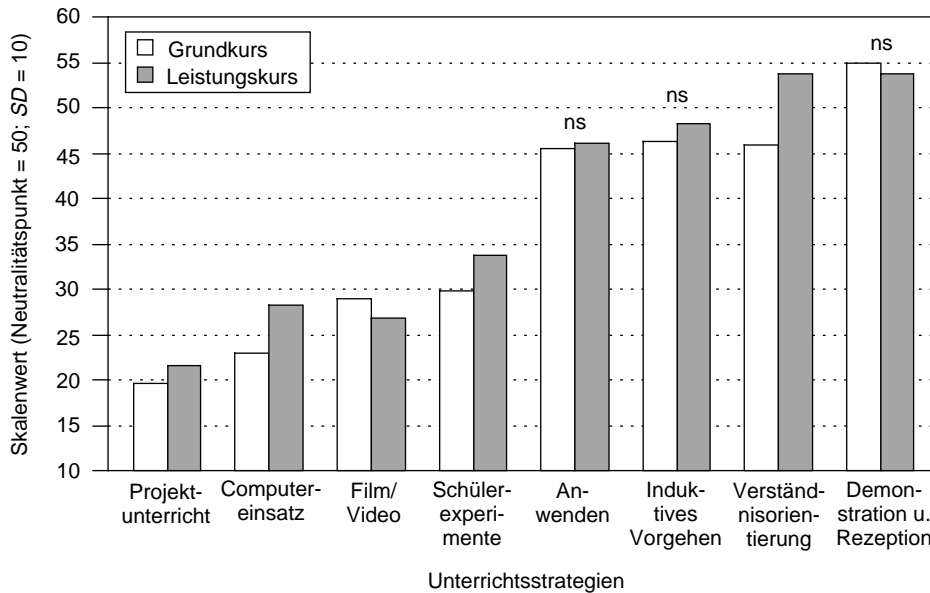
© TIMSS/III-Germany

stufe scheint vornehmlich Demonstrationsunterricht zu sein, in dem Lehrkräfte mithilfe des Vorführexperiments einen physikalischen Gedankengang entwickeln. Das gelenkte Schülerexperiment ist selten, und die Entwicklung von Experimenten durch Schüler kommt praktisch nicht vor (vgl. Abb. 13).

Im Physikunterricht lassen sich mit neun didaktischen und methodischen Merkmalen als Prädiktoren 40 Prozent der Leistungsvariation zwischen Kursen erklären. Der Zusammenhang zwischen Unterrichtsmerkmalen und Fachleistungen ist in Physik höher als in Mathematik. Interpretiert man das optimale Erklärungsmodell extensiv, erzielt folgender Physikunterricht günstigere Leistungsergebnisse:

- Die Lehrkraft legt Wert auf kognitiv-anspruchsvolle Aufgaben und das theoretische Verständnis von physikalischen Modellen.
- Sie unterstützt den Aneignungsprozess durch theoretisch gut vorbereitete Experimente, die auch unter Nutzung des Rechners durchgeführt oder ausgewertet werden.
- Dabei sind Schüler- und Lehrereperiment jedoch nicht Bestandteil eines induktiven Vorgehens, bei dem der handelnde Umgang mit Versuchsanordnungen

Abbildung 13: Unterrichtsstrategien in Physik nach Kursniveau



IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

- bzw. die Beobachtung von Phänomenen der theoretischen Fragestellung vorgelagert sind.
- Schließlich wird die verfügbare Unterrichtszeit zur Erarbeitung und Konsolidierung von Sachverhalten optimal ausgenutzt.

Multiple Zielerreichung und die psychosoziale Situation von Schülerinnen und Schülern in der gymnasialen Oberstufe

Schule ist ein Unternehmen, das unterschiedliche und teilweise konkurrierende Zielsetzungen gleichzeitig verfolgt. In gelingenden Bildungsprozessen vollzieht sich nach allgemein geteilter Vorstellung die gleichmäßige Entwicklung kognitiver, sozialer und emotionaler Fähigkeiten. Alle Schulzielbestimmungen der Länder nehmen diese Gedanken in der einen oder anderen Formulierung auf. Darüber hinaus soll die Schule auch Ansprüchen distributiver Gerechtigkeit genügen, nach denen sich die möglichst optimale Förderung eines jeden Einzelnen mit der Sicherung der

notwendigen Qualifikationsvoraussetzungen für eine verantwortliche Teilhabe an Beruf und Gesellschaft für alle verbindet.

Es ist keineswegs selbstverständlich, dass es gelingt, die unterschiedlichen Ziele gleichzeitig und gleich gut zu erreichen. Vielmehr ist in der praktischen Pädagogik die Ansicht verbreitet, dass im Alltag von Schule und Unterricht erhebliche Verträglichkeitsprobleme auftreten, die Optimierungsentscheidungen verlangen. Dieses Argument wird immer dann virulent, wenn in Teilbereichen Probleme bei der Zielerreichung diagnostiziert werden. Fallen die Fachleistungen weniger gut als erwartet aus, könnte diese Schwäche zum Beispiel durch Stärken im Erwerb fachübergreifender Kompetenzen kompensiert werden.

Wie sieht nun die Befundlage der Forschung aus? Für die Sekundarstufe I gibt es mehrfach replizierte Ergebnisse, die besagen, dass eine optimale individuelle Förderung und der gleichzeitige Ausgleich von Leistungsunterschieden in einem Spannungsverhältnis stehen, das mehr oder weniger optimal balanciert, aber im Rahmen des üblichen Klassenunterrichts kaum aufgelöst werden kann. Für die Verträglichkeit kognitiver und affektiver Zielkriterien von Unterricht gibt es ebenfalls einige Arbeiten, deren Ergebnisse weitgehend konsistent sind. Die Befunde belegen keine Verträglichkeitsprobleme, sondern eher kumulativ wirkende Effekte.

Die Ergebnisse unserer Analysen bestätigen in der Grundstruktur einschlägige Befunde aus der Mittelstufe und differenzieren diese in einigen wichtigen Punkten. Von einem grundlegenden strukturellen Kompatibilitätsproblem bei mehrdimensionaler Zielerreichung kann offensichtlich nicht die Rede sein. Innerhalb von institutionell definierten Domänen lässt sich eine Kopplung von kognitiven und motivationalen Zielkriterien nachweisen, die in einem stabilen Verständnis- und Motivationssyndrom auf institutioneller Ebene zum Ausdruck kommt. Ein steigendes durchschnittliches Leistungsniveau innerhalb eines Kurses wird nicht mit Interessenverlusten erkauft; ganz im Gegenteil: Kompetenz und Sachinteresse stützen sich auch auf Lerngruppenebene wechselseitig. Die Zusammenhänge zwischen kognitiven und motivationalen Kriterien sind in der gymnasialen Oberstufe erwartungsgemäß straffer als in der noch wenig ausdifferenzierten Mittelstufe oder gar der Grundschule. Gemeinsame Basis einer mehrdimensionalen Zielerreichung innerhalb eines Fachgebiets scheinen verständnisorientierte Unterrichtsstrategien zu sein, die vermutlich für das Verständnis- und Motivationssyndrom verantwortlich sind. Repetitive und rezeptive Unterrichtsführung, aber auch strukturarme Unterrichtsformen, wie sie das induktive Vorgehen im Physikunterricht offenbar darstellt, stehen in negativem Zusammenhang sowohl mit kognitiven als auch mit motivationalen Kriterien.

Generalisierte Einstellungen oder selbstbezogene Kognitionen – wie zum Beispiel das Selbstwertgefühl oder die Schulfreude – variieren dagegen weitgehend unabhängig von bereichsspezifischen Verständnis- und Motivationszusammenhängen. Die Qualität von Einzelkursen, sei sie herausragend oder weniger gut, schlägt also in der Regel – man könnte auch sagen, glücklicherweise – nicht unmittelbar auf generalisierte Merkmale der Person durch.

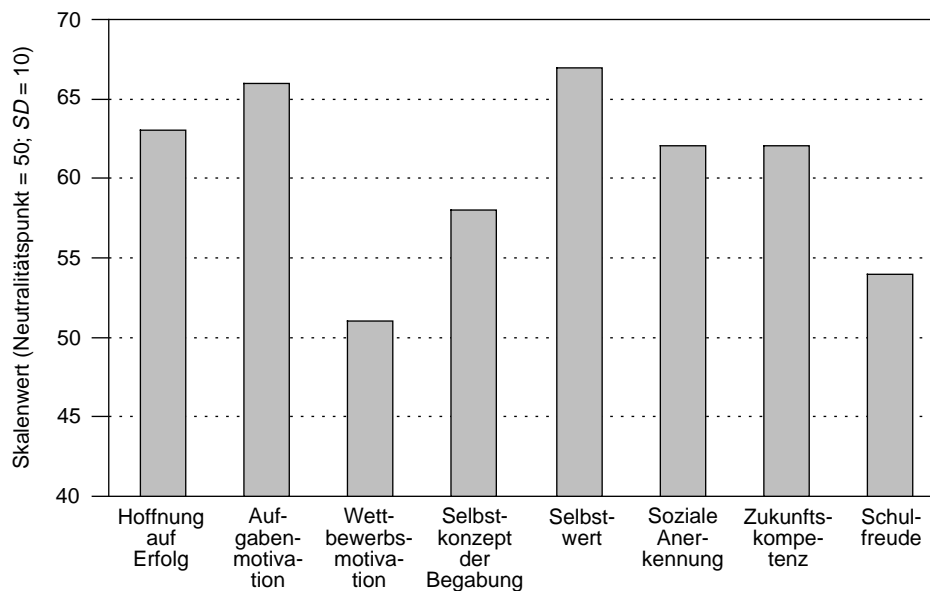
Wirft man schließlich einen Blick auf die Selbstangaben, die Auskünfte über die allgemeine psychosoziale Situation von Schülerinnen und Schülern an der gymnasialen Oberstufe geben, so widersprechen die Befunde jenem düsteren Bild, das von Schulkritikern gelegentlich entworfen wird (Entfremdung von der Schule, hemmungsloser Wettbewerb, instrumentelle Orientierung an selbstwertschädigendem Leistungsdruck, soziale Bindungslosigkeit und existentielle Sorge um die persönliche Entwicklung nach der Schule).

Die Analyseergebnisse weisen durchgehend auf eine positive Selbstwahrnehmung des größten Teils der Oberstufenschülerinnen und -schüler hin. Die tendenziell positive Selbstwahrnehmung gilt in allen untersuchten Dimensionen für mindestens 85 Prozent der Befragten. Ein Überblick über das vorherrschende Wahrnehmungsmuster gibt Abbildung 14. Besonders auffällig ist das hohe Selbstwertgefühl, das mit der Vorstellung selbstwertschädigenden Leistungsdrucks überhaupt nicht zusammengeht. Ebenso widerspricht die gerade im Vergleich zur Wettbewerbsmotivation hohe Ausprägung der intrinsischen Aufgabenorientierung Behauptungen vom schrankenlosen Leistungswettbewerb. Oberstufenschüler nehmen sich in der überwiegenden Mehrzahl aufgabenorientiert, erfolgsmotiviert und gleichzeitig sozial anerkannt wahr. Auch die Schulunlust scheint sich in Grenzen zu halten und oftmals übertrieben zu werden. Immerhin geben 75 Prozent der Befragten an, dass die Aussage: „Ich bin froh, dass ich noch zur Schule gehe“, für sie eher oder völlig zutrefte. Auch die persönliche Zukunft sehen Gymnasiasten in der großen Mehrzahl nicht als verbaut, sondern als kompetent gestaltbar an.

13. Institutionelle und regionale Variabilität und die Sicherung gemeinsamer Standards in der gymnasialen Oberstufe

Einen weiteren Untersuchungsschwerpunkt bilden Analysen mit eindeutig institutionellem Fokus. Der gemeinsame Fluchtpunkt dieser Analysen ist die Frage nach der institutionellen Sicherung verständnisorientierter Lernprozesse und vergleich-

Abbildung 14: Allgemeine psychosoziale Situation von Schülerinnen und Schülern der Oberstufe (Mittelwerte)



IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

barer Leistungsstandards. Uns liegt sehr daran, die Integration der Gesichtspunkte des verständnisvollen Lernens und der Standardsicherung nachdrücklich zu betonen. Wir beschreiben Leistungsstandards deshalb auch immer mit den qualitativen Begriffen inhaltlich definierter Fähigkeitsniveaus. Auf dieser theoretischen Grundlage werden Fragestellungen zum Zusammenhang von erreichten Leistungsniveaus und institutionellen Merkmalen, wie sie etwa die Bildungsbeteiligung an voruniversitären Programmen, die Dauer von Bildungsgängen, die Prüfungsstruktur oder die Vorgabe von Benotungsstandards darstellen, entwickelt. Diese Analysen sollen den Blick für die Komplexität institutioneller Bedingungen im Schulwesen schärfen. Über die Bearbeitung folgender Fragestellungen soll an dieser Stelle zusammenfassend berichtet werden:

- Über die Notwendigkeit einer Balance von kursübergreifenden und kursinternen Maßstäben bei der Notenvergabe in der Oberstufe
- Standardsicherung durch die Abiturprüfung: Zentralabitur oder dezentrale Prüfungsorganisation?
- In 12 oder 13 Schuljahren zum Abitur?

Die notwendige Balance zwischen kursübergreifenden und kursinternen Bewertungsgrundsätzen

Das Abitur und die Erteilung der Hochschulreife haben eine Schlüsselstellung im deutschen Berechtigungssystem, mit dem die Ausübung gesellschaftlicher Funktionen an Ausbildungsabschlüsse gebunden wird. Diese Abschlüsse dokumentieren das Durchlaufen eines förmlichen Bildungsgangs und das Erreichen eines bestimmten Qualifikationsniveaus. Formale Abschlüsse können ihre Zertifizierungsfunktion allerdings nur erfüllen, wenn sie von einem generalisierten Vertrauen getragen werden, dass die an einen Abschluss gebundenen Qualifikationserwartungen trotz Enttäuschungen in Einzelfällen eingelöst werden. Die Expertenkommission der Kultusministerkonferenz zur „Weiterentwicklung der Prinzipien der gymnasialen Oberstufe und des Abiturs“ formulierte deshalb vier Erwartungen an die allgemeine Hochschulreife. Danach dürfen erwartet werden:

- die Einhaltung von ausreichenden Mindeststandards,
- Auskunft über individuelle Leistungsprofile,
- überschulische Vergleichbarkeit der Standards und
- interindividuelle Differenzierung der Prädikate.

Die Frage der schulübergreifenden Vergleichbarkeit von Beurteilungsmaßstäben ist angesichts der Bedeutung der Hochschulreife im Berechtigungssystem gar nicht abweisbar. Sie kann jedoch nicht sinnvoll innerhalb der Dichotomie von Vergleichbarkeit versus Unvergleichbarkeit beantwortet werden, da diese Alternative die spezifische Logik der pädagogischen Arbeit im Unterricht außer Acht lässt. Ein erfolgreicher Unterricht verlangt neben einem professionellen Bewusstsein für lerngruppenübergreifende Leistungsstandards auch immer die Berücksichtigung lerngruppeninterner Referenzmaßstäbe. Unter einer analytisch fruchtbaren Perspektive wird man das Problem der Vergleichbarkeit von Beurteilungsmaßstäben nur als (lokal) auszutarierende Balance zwischen unterschiedlichen, aber funktional gleichermaßen legitimen Bezugsnormen formulieren können. Man wird also zu fragen haben, welche Belastungen das Vertrauen in die Gerechtigkeit der Leistungsbewertung und Abschlussvergabe verträgt und wie weit sich lerngruppeninterne Referenznormen unter pädagogischen Gesichtspunkten von institutionsübergreifenden Standards unterscheiden dürfen.

TIMSS/III kann nicht auf alle implizierten Fragen antworten, wohl aber einige kritische Punkte beleuchten. Ein Blick auf die Notenverteilung im Mathematik- und Physikunterricht gibt bereits eine erste Antwort auf die Frage nach der interindividuellen Differenzierungsfähigkeit schulischer Beurteilungen. Lehrkräfte an gymnasialen Oberstufen nutzen die Notenskala in der Breite von 1 bis 5 vollständig aus.

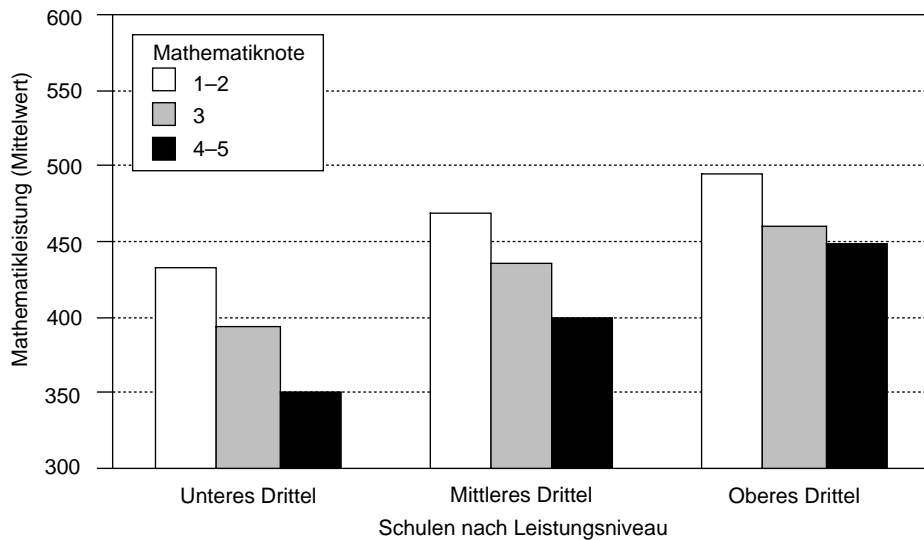
Dass Personen, die Physik und insbesondere Mathematik abgewählt haben, häufiger schlechtere Noten erhalten, ist ein erster Hinweis auf die Validität der Beurteilung. Erwartungsgemäß ist auch der Anteil schlechter Zensuren im praktisch obligatorischen Mathematikgrundkurs höher als im Grundkurs des selektiven Fachs Physik. Gerade diese Differenz der Notenverteilung weist auf ein Standardbewusstsein von Fachlehrern an gymnasialen Oberstufen hin, das sich nicht nur an kursinternen Normen orientiert, sondern auch kurs- und schulübergreifende Kriterien in Anschlag bringt. Diese Befunde werden durch eine Analyse des Zusammenhangs von Notenstufen und dem Niveau der erreichten Fachleistungen bestätigt. Die Fachnote erklärt *über alle Schulen hinweg* einen nennenswerten Anteil der Leistungsvariation.

Andererseits ist aber auch die Überlappung der Leistungsverteilungen unterschiedlicher Notenstufen, die auf eine erhebliche Variabilität der Bewertungsgrundsätze hinweist, nicht zu übersehen. Um das Ausmaß der systematischen Schwankungen anschaulich zu machen, haben wir die Schulen unserer Stichprobe nach Fachleistungen – getrennt nach Fächern und Kursniveaus – gruppiert. Die Schulen wurden jeweils in ein unteres, mittleres und oberes Leistungsdrittel eingeteilt, um anschließend die mit den Notenstufen korrespondierenden mittleren Fachleistungen vergleichen zu können (vgl. Abb. 15). Unter Konstanzhaltung der Notenstufe unterscheiden sich die Schulen in der erreichten Fachleistung. In Schulen mit höherem Leistungsniveau wird über alle Notenstufen hinweg strenger beurteilt. Auf Grundkursniveau sind in Schulen des oberen Leistungsdrittels die Mindeststandards besonders hoch. Die mittleren Differenzen zwischen den leistungsstärksten und leistungsschwächsten Schulen können zwei oder sogar drei Notenstufen betragen.

Unter den Gesichtspunkten der Verteilungsgerechtigkeit mögen diese Schwankungen in der Strenge der Leistungsbewertung groß erscheinen, vielleicht sogar Anlass zur Sorge geben. Umso wichtiger ist es, nachdrücklich auf die schulübergreifend geteilten Leistungsvorstellungen der Lehrkräfte an gymnasialen Oberstufen hinzuweisen. Im impliziten professionellen Gerechtigkeitsbild von Oberstufenlehrkräften scheinen sich schulübergreifende Gütekriterien und kursinterne Referenznormen, die für ein motiviertes Arbeiten gar nicht außer Acht gelassen werden können, in einer Balance zu befinden. Eine Regressionsanalyse von Fachnoten auf Fachleistungen in der Gesamtstichprobe und relativen Fachleistungen im Kurs zeigt, dass beide Beurteilungsaspekte in das in der Notenvergabe zum Ausdruck kommende Urteil von Fachlehrkräften eingehen, wobei der spezifische Anteil der schulübergreifenden Standards in der Regel bedeutsamer ist.

Ferner darf man nicht übersehen, dass es sich bei der Variabilität von Bewertungsmaßstäben zum überwiegenden Teil um *kurs-* und *fachspezifische*, nicht aber um

Abbildung 15: Mathematikleistung in Grundkursen nach Schulniveau und Note (Mittelwerte)



IEA. Third International Mathematics and Science Study.

© TIMSS/III-Germany

schulspezifische Phänomene handelt, sodass im Gesamtergebnis eines Abiturzeugnisses Gerechtigkeitsdefizite der einzelnen Noten zu einem vermutlich nicht unerheblichen Teil austariert werden. Die Tatsache, dass der beste bekannte allgemeine Prädiktor für den Studienerfolg die Durchschnittszensur des Abiturzeugnisses ist – und nicht die Durchschnittszensur der Fächer Deutsch, Englisch und Mathematik oder der Abiturprüfungsfächer, aber auch nicht die Summe gewichteter Einzelnoten –, findet wahrscheinlich in diesem Umstand eine Erklärung.

Man muss sich diese komplexe Befundlage vergegenwärtigen, wenn man die Variabilität von Bewertungsmaßstäben beurteilen möchte. Hält man die Schwankungsbreite für zu groß, weil sie das Vertrauen in die mit der Hochschulreife erteilten Berechtigungen schwächen könnte, so ist Abhilfe wohl kaum in einer Verschärfung der Beurteilung in Schulen mit weniger günstigen Leistungsergebnissen zu suchen. Angesichts des Gleichgewichts von kursübergreifenden und kursinternen Maßstäben und der schon jetzt – vor allen Dingen im mathematischen Grundkurs – strengen Benotung wird man durch Maßnahmen dieser Art nur die motivationale Grundlage der pädagogischen Arbeit gefährden können. Eine Homogenisierung der Standards

dürfte letztlich nur durch eine Verbesserung des Unterrichts gerade in den Klassen und Kursen, deren Leistungen nicht den berechtigten Erwartungen entsprechen, zu erreichen sein. Hier liegt die eigentliche Herausforderung einer professionellen Schulentwicklung.

Standardsicherung durch die Abiturprüfung

In der Diskussion über Qualitätsentwicklung und Qualitätssicherung spielt die Gestaltung der Abiturprüfung immer wieder eine wichtige Rolle. Befürworter eines Zentralabiturs bezweifeln, ob eine dezentrale Abiturprüfung äquivalente Funktionen bei der Qualitätssicherung erfüllen könne. Umgekehrt gehen Anhänger eine dezentralen Abiturprüfung von eine Äquivalenz beider Organisationsformen hinsichtlich der Qualitätssicherung aus. Sie befürchten allerdings negative Rückwirkungen einer zentralen Prüfung auf Unterrichts- und Lernprozesse.

Es lässt sich zeigen, dass beide Argumentationen theoretisch nicht stichhaltig sind. Unter dem Gesichtspunkt der Qualitätssicherung ist es plausibler, von differentiellen Wirkungen der Prüfungsform auszugehen, die von der jeweiligen Selektivität eines Unterrichtsfachs abhängen. Hinsichtlich möglicher Rückwirkungen auf individuelles Lernverhalten oder motivationale Orientierungen lässt sich keine gerichtete Hypothese theoretisch begründen.

Die Analysen erbrachten folgende Befunde:

- Im wenig selektiven Fach Mathematik unterscheiden sich die Leistungsmittelwerte in Ländern mit Zentralabitur und dezentraler Prüfungsordnung zu Gunsten des Zentralabiturs in nicht trivialer Weise. Erwartungsgemäß ist auch die Leistungsstreuung in Ländern mit Zentralabitur reduziert und zwar insbesondere auf Grundkursniveau.
- Wiederholt man die Analysen für das Fach Physik, lassen sich keine Effekte der Organisationsform der Abiturprüfung nachweisen. In obligatorischen oder häufig gewählten Kursen der gymnasialen Oberstufe scheint das Zentralabitur Leistungsniveaus zu stabilisieren und für eine größere Leistungshomogenität zu sorgen. Mit zunehmender Selektivität des Fachs oder des Kurses verschwindet der Effekt der Organisationsform. Das zentrale Korrektiv hat über die fachintern wirksamen professionellen Standards hinaus keine zusätzliche normierende Wirkung.

- Bei der Prüfung, ob die Organisationsvarianten unterschiedlich normierend wirken, ließ sich feststellen, dass im Grundkurs Mathematik unter den Bedingungen des Zentralabiturs Mindeststandards strenger definiert werden. Für den mathematischen Leistungskurs und für die beiden Kursniveaus im Fach Physik sind unterschiedliche Bewertungsmaßstäbe nicht nachweisbar.
- Um Argumente für die Beantwortung der Frage zu gewinnen, ob bei einem Zentralabitur möglicherweise mit unerwünschten Nebenwirkungen auf Lernprozesse zu rechnen sei, wurden die Verteilungen der Schülerinnen und Schüler auf die von uns definierten mathematischen Fähigkeitsniveaus verglichen. In Ländern mit Zentralabitur ist der Anteil von Schülerinnen und Schülern, die sich auf dem untersten Fähigkeitsniveau befinden, deutlich geringer als in Ländern mit dezentralem Abitur. Der Vergleich zeigt aber auch, dass von einer Beeinträchtigung verständnisvollen Lernens durch das Zentralabitur keine Rede sein kann.
- In einem letzten Schritt wurden mögliche negative motivationale Auswirkungen des Zentralabiturs geprüft. Die Ergebnisse sprechen gegen einschlägige Befürchtungen.

Fasst man die Ergebnisse unserer Analysen zusammen, so ist zunächst zu betonen, dass die Befunde keinen Schluss auf die Überlegenheit der einen oder anderen Organisationsform der Abiturprüfung erlauben. Die Ergebnisse sind differentiell und überdies auf die Fächer Mathematik und Physik beschränkt. Im Rahmen dieser Begrenzung lässt sich zeigen, dass ein Zentralabitur bei obligatorischen Grundkursen vor allem im unteren Leistungsbereich standardsichernd wirken kann. Werden Fächer oder Kurse selektiv angewählt, verliert sich dieser Effekt gegenüber der Selbstregulation der professionellen Beurteilungspraxis der Oberstufenlehrkräfte. Von negativen Rückwirkungen des Zentralabiturs auf das individuelle Lern- und Motivationsgeschehen kann keine Rede sein. Die ermittelten Befunde sprechen im Fach Mathematik eher für eine ordnende und entlastende Wirkung externer Standardsetzungen. Die Effektstärken sind aber gering.

In 12 oder 13 Schuljahren zum Abitur?

Die Dauer der Schulzeit bis zur Erlangung der allgemeinen Hochschulreife beträgt in Deutschland nach dem Hamburger Abkommen 13 Jahre. An dieser Festlegung wurde bis zur deutschen Vereinigung nicht ernsthaft gerüttelt. Diese Sachlage änderte sich schlagartig mit der deutschen Vereinigung und dem Wunsch der Mehrzahl der neuen Länder, trotz Anpassung des Schulsystems an das westdeutsche Mo-

dell die Vergabe der allgemeinen Hochschulreife nach 12 Schuljahren beizubehalten. Begründet wurde dieser Wunsch weniger mit dem Hinweis, dass sich diese Variante in der DDR bewährt habe, sondern vielmehr mit dem Argument der Homogenisierung europäischer Bildungsgänge und der insgesamt zu langen akademischen Ausbildung in Deutschland. In dieser historischen Konstellation gewann eine an sich bekannte Argumentationsfigur neues Gewicht.

Der Beschluss der Kultusminister über die wechselseitige Anerkennung des Abiturs änderte trotz Schulzeitverkürzung an der gymnasialen Oberstufe strukturell nichts. Die Einführungsphase wurde in die Jahrgangsstufe 10 vorverlegt und die viersemestrige Kursphase blieb unangetastet. In Kombination mit der Einigung auf ein gemeinsames Gesamtstundenvolumen für die Sekundarstufe I und II folgte daraus in erster Linie eine zeitliche und stoffliche Komprimierung der Mittelstufe und je nach landesspezifischen Vorschriften auch eine stärkere Stundenbelastung der Jahrgangsstufen 10 bis 12. Ob eine solche Stauchung des gymnasialen Bildungsgangs ohne Kosten und unerwünschte Nebenwirkungen realisiert werden konnte, ist eine bislang nicht systematisch untersuchte oder gar beantwortete Frage.

In einer Zusammenfassung unserer Analysen zur Schulbesuchsdauer sind folgende Punkte herauszuheben:

- Eine Verkürzung der Schulbesuchsdauer bis zum Abitur um ein Jahr bei gleichzeitiger Gewährleistung eines gemeinsamen Mindeststundenvolumens für die Sekundarstufen I und II muss nicht mit Abstrichen am erreichbaren Wissensstand erkaufte werden. Dies gilt sowohl für das durchschnittliche Wissensniveau als auch für die Verständnistiefe der vermittelten Sachverhalte – jedenfalls für das Fach Mathematik und den physikalischen Grundkurs.
- Befunde zum Leistungskurs Physik warnen allerdings vor einer Generalisierung. Im Physikleistungskurs werden im 13. Jahrgang neue und anspruchsvolle Stoffe und Konzepte erarbeitet, die physikalische Sachverhalte in eine theoretisch konsistentere Perspektive rücken. Die TIMSS/III-Ergebnisse lassen hier eine spezifische Stärke des 13-jährigen Bildungsgangs sowohl hinsichtlich des durchschnittlichen Wissensniveaus als auch hinsichtlich der erreichten Verständnistiefe erkennen. Möglicherweise sind die vergleichbaren oder sogar besseren Leistungsergebnisse, die in den Ländern mit 12-jähriger Schulbesuchsdauer im Fach Mathematik oder im Grundkurs Physik erzielt werden, auf eine suboptimale Zeitzunutzung im 13-jährigen Bildungsgang zurückzuführen.

- Es gibt keine Hinweise, dass eine Schulzeitverkürzung zu oberflächlicherem Lernen führte. Im Gegenteil: Schülerinnen und Schüler im verkürzten Bildungsgang scheinen eher von verständnisorientierten Lernstrategien Gebrauch zu machen.
- Negative motivationale Rückwirkungen der Schulzeitverkürzung auf das Lernen im Fach sind nicht nachweisbar, obwohl eine curriculare Komprimierung am ehesten im Fach Auswirkungen zeigen sollte. Im Mathematikunterricht sind tendenziell erwartungswidrige Vorteile des kürzeren Bildungsgangs zu erkennen.
- Die allgemeine psychosoziale Situation von Schülerinnen und Schülern ist im 13-jährigen Gymnasium durchweg günstiger. Dieser Effekt ist jedoch mit allgemeinen Ost-West-Unterschieden konfundiert.

14. Studienfachwünsche und Berufsorientierungen in der gymnasialen Oberstufe

Aus diesem Bereich werden zwei Fragestellungen thematisiert: Zum einen geht es darum, wesentliche Einflussfaktoren der Neigung, nach dem Abitur ein Studium aufzunehmen, zu identifizieren. Zum anderen wird der Frage nachgegangen, in welchem Ausmaß die Studienfachpräferenz am Ende der Schullaufbahn durch die Leistungsfachwahl in der gymnasialen Oberstufe vorstrukturiert wird.

Hinsichtlich der ersten Fragestellung lässt sich zunächst feststellen, dass sich das Hochschulstudium – nach einem Einbruch in den 1980er Jahren – wieder großer Beliebtheit erfreut: Rund vier Fünftel der deutschen Abiturienten planen kurz- oder mittelfristig die Aufnahme eines Hochschulstudiums. Die TIMSS-Befragungsdaten stehen hier im Einklang mit Populationsdaten, wonach sich der Trend zum Übergang auf die Universitäten und Hochschulen in den vergangenen zehn Jahren wieder spürbar verstärkt hat.

Dies gilt allerdings nur für die alten Bundesländer. Für die neuen Bundesländer belegen die TIMSS-Daten deutliche Folgen des Transformationsprozesses im Bildungswesen, die zum Zeitpunkt der Erhebung dazu führten, dass die beruflichen Zukunftspläne ostdeutscher Oberstufenschüler im Durchschnitt weniger auf eine akademische Ausbildung ausgerichtet waren als in den alten Bundesländern. Wie die Zusammenstellung von Befunden aus TIMSS und der BIJU-Studie nahelegt, sind in der Phase nach der Umstellung auf ein gegliedertes Schulsystem verstärkt auch Teile der Schülerschaft auf das Gymnasium und in der Folge auf die Oberstufe

gegangen, die in den westlichen Bundesländern eher Realschulklientel darstellen. Die auffallend starke Unsicherheit über ihre berufliche Zukunft unter ostdeutschen Gymnasiasten verweist auf eine im Rückblick durchaus rationale Handlungsstrategie: In Zeiten besonders unsicherer Arbeits- und Ausbildungsmärkte, wie dies zweifelsohne für die Jahre vor der TIMSS-Haupterhebung in den neuen Bundesländern zutraf, ist ein Verbleib im allgemeinbildenden Schulwesen mit dem Ziel einer Höherqualifizierung sinnvoll, weil sie die spätere Entscheidung für eine nicht akademische berufliche Ausbildung offenhält.

Die in Ostdeutschland geringere Studienneigung hängt indirekt auch mit dem dort höheren Frauenanteil zusammen. Wie sich zeigen lässt, ist die Absicht, ein Studium aufzunehmen, bei jungen Frauen nach wie vor geringer als bei jungen Männern. Diese Differenz erfährt eine zusätzliche Verschärfung, wenn sich die Abiturientinnen vor allem an einer traditionellen Geschlechtsrolle, zu der die Jugendlichen in West wie Ost trotz unterschiedlicher Sozialisationsbedingungen gleichermaßen neigen, orientieren: In diesem Fall nimmt die ohnehin geringere Studienneigung – unabhängig von der gezeigten Schulleistung – noch weiter ab.

Trotz dieser geschlechtsspezifischen Einflüsse spielen andere Aspekte wie die erreichten Schulleistungen oder der soziokulturelle Hintergrund der Abiturienten eine weit größere Rolle für die Studienneigung: Die Wahrscheinlichkeit, ein Studium zu beginnen, ist bei sehr guten Leistungen in den Leistungskursen doppelt so hoch wie bei nur ausreichenden Leistungen. Die Aufnahme eines Studiums ist zudem bei jenen Abiturienten wahrscheinlicher, deren Eltern einer höheren Bildungsschicht angehören. Vermittelt werden diese Einflüsse zu einem großen Teil über entsprechende Erwartungen der Eltern. Diese Erwartungen an die weitere Bildungslaufbahn ihres Kindes sind aber nicht nur durch die Zugehörigkeit zu einer bestimmten Bildungsschicht bedingt, sondern wiederum durch die in der Oberstufe erbrachten Leistungen ihres Sohnes/ihrer Tochter mitverursacht.

Im Hinblick auf die zweite Fragestellung lässt sich anhand der TIMSS-Daten das ungebrochen starke geschlechterstereotype akademische Wahlverhalten sowohl in der Schule bei der Wahl der Leistungskurse als auch in der Studienplanung bezüglich des angestrebten Studienfachs nachweisen: Junge Frauen wählen häufiger als junge Männer sprachliche und gesellschaftswissenschaftliche Leistungsfächer und präferieren Studienfächer aus dem erziehungs- oder sprachwissenschaftlichen Bereich. Zwischen der Wahl von Leistungsfächern bzw. -kombinationen einerseits und Studienpräferenzen andererseits bestehen – zumindest bei jenen Studienfächern, die eine Korrespondenz zu den schulischen Fächern aufweisen – deutliche Übereinstimmungen. Der Studienwunsch wird offensichtlich durch die Wahl eines bestimmten

Leistungsfachs angebahnt, wobei allerdings offen bleiben muss, inwieweit nicht auch die umgekehrte Einflussrichtung für diese Korrespondenz verantwortlich ist.

In der Stärke dieses Zusammenhangs lassen sich keine Unterschiede zwischen männlichen und weiblichen Oberstufenschülern erkennen. Dies ist ein zentraler Befund der hier gezeigten Analysen, da es den Schluss zulässt, dass die berufliche Auseinanderorientierung von Männern und Frauen bereits in der Oberstufe anhand der Leistungsfachwahlen klar abzulesen ist und sich an der Übergangschwelle zur Hochschule nicht durch antizipierte Geschlechterdiskriminierung in einzelnen Studienfächern weiter verschärft. Dies gilt zumindest für die Studienabsichten, für die man folglich die begründete Erwartung ableiten kann, dass deren starke Orientierung am traditionellen Geschlechterstereotyp durch eine Erhöhung der Frauenanteile in den eher untypischen Leistungskursfächern abgebaut würde. Inwieweit dies ohne flankierende Maßnahmen bereits in der Mittelstufe möglich ist, wird man aufgrund der Befunde für die TIMSS Population II allerdings eher skeptisch beurteilen müssen.

15. Geschlechtsdifferenzen in den mathematisch-naturwissenschaftlichen Leistungen

Leistungsunterschiede zwischen Mädchen und Jungen bzw. Frauen und Männern im mathematisch-naturwissenschaftlichen Unterricht stellen nach wie vor eine große pädagogische Herausforderung dar. Insbesondere die etwas geringeren Leistungen der Mädchen im Mathematikunterricht sind Gegenstand eines kaum noch überschaubaren Berges an Literatur.

Die praktische Relevanz dieses Themas spiegelt sich in unterschiedlichen Kurswahlen in der gymnasialen Oberstufe sowie in geschlechtsspezifischen Berufs- und Studienwahlen wider. Wie zum bloßen Phänomen der Leistungsunterschiede zwischen Jungen und Mädchen in Mathematik gibt es auch eine Vielzahl von theoretischen und empirischen Arbeiten, die potentielle Ursachen der Differenzen beleuchten. Als Ursachen werden genetische, sozialisationsbedingte, unterrichtsbedingte, motivationale und viele andere Gründe genannt.

Im naturwissenschaftlichen Bereich zeigt ein Überblick über rezente Studien, dass substantielle Geschlechtsdifferenzen zu Gunsten der Jungen im physikalischen und chemischen Bereich bereits im Grundschulalter bestehen, in Biologie werden die Unterschiede erst im Sekundarbereich I bedeutsam.

Theorien zur Erklärung der Geschlechtsunterschiede in den Naturwissenschaften, insbesondere in Physik, entsprechen weitgehend den Ansätzen der Mathematik, zum Teil werden aber auch andere Aspekte betont. Ergänzende Modelle rekurrieren sehr stark auf unterschiedliche Aktivitäten bei schulischer Laborarbeit (schulischen Experimenten) und *Out-of-School Activities* von Jungen und Mädchen. So ließ sich zeigen, dass Jungen größere Erfahrung im physikalischen Experimentieren hatten und sich deutlich stärker mit naturwissenschaftlichen Inhalten in ihrer Freizeit auseinandersetzten. Die Benachteiligung der Mädchen beim schulischen Experimentieren und die damit verbundenen Minderleistungen haben zu Forderungen nach einer zeitweisen Aufhebung der Koedukation im naturwissenschaftlichen Unterricht geführt.

Die TIMSS-Oberstufenkohorte bot die einmalige Möglichkeit, für ein national repräsentatives Sample dieser Altersstufe Geschlechtsdifferenzen im mathematischen und naturwissenschaftlichen Leistungsbereich zu untersuchen. In Einklang mit früheren nationalen und internationalen Befunden erwarten wir eine Bestätigung dahingehend, dass junge Männer sowohl in Mathematik als auch in den Naturwissenschaften im Vergleich zu den Frauen höhere Leistungen zeigen sollten. Die Unterschiede in der Physik sollten dabei stärker als in der Mathematik sein. Im Vergleich zur TIMSS-Mittelstufenkohorte sollten die Abstände größer geworden sein.

Die Analysen bestätigten im Wesentlichen unsere Erwartungen. Junge Männer zeigten höhere Leistungen im mathematischen und naturwissenschaftlichen Bereich, allerdings mit bemerkenswerten Schwankungen:

- Voruniversitäre Mathematik weist im Vergleich zur Physik und zur Grundbildung die geringsten Geschlechtsdifferenzen auf.
- Dies gilt insbesondere für die voruniversitären Mathematikleistungen im Grundkurs, in dem die Differenzen beinahe durchgängig trivial sind.
- Bezogen auf die drei Sachgebiete in Mathematik zeigen sich wie schon in TIMSS-Mittelstufenanalysen in der Geometrie die geringsten Geschlechtsdifferenzen. Viele der Geometrieaufgaben thematisieren eher Mittelstufenstoff.
- Hinsichtlich der kognitiven Anforderungsdimensionen der voruniversitären Mathematikaufgaben ergibt sich im Einklang mit der nationalen und internationalen Literatur, dass Aufgaben, die einfache mathematische Operationen zu ihrer Lösung erfordern, kleine oder keine Geschlechtsdifferenzen aufweisen.

- Innerhalb der voruniversitären Physik schwanken die Geschlechtsdifferenzen beträchtlich, allerdings durchgängig zu Gunsten der Jungen.
- Im Gegensatz zur Mathematik sind die Geschlechtsdifferenzen in der voruniversitären Physik in Grund- und Leistungskursen in ähnlicher Höhe.
- Verblüffend sind die im Vergleich zur voruniversitären Mathematik deutlich größeren Geschlechtsdifferenzen in der mathematischen Grundbildung. Im Gegensatz zu den Aufgaben zur voruniversitären Mathematik beanspruchen die Items zur Grundbildung explizit keine curriculare Validität. Vielmehr soll hier das im Laufe der Sekundarstufe I erworbene Wissen auf alltagsnahe Problemstellungen übertragen bzw. angewandt werden. Aus der Forschung zum Wissenstransfer ist das Phänomen des „trägen Wissens“ bekannt, das heißt, Personen verfügen im Prinzip über das zur Lösung notwendige Wissen, erkennen aber nicht, dass es auf die konkrete Aufgabe angewendet werden kann. Träges Wissen, auf Seiten der jungen Frauen stärker ausgeprägt, könnte eine Erklärung für die doch deutlich größeren Nachteile im Grundbildungstest sein.
- Die insgesamt sehr großen Leistungsnachteile der Frauen in Physik machen deutlich, dass es dem Physikunterricht der gymnasialen Oberstufe nicht gelingt, Voraussetzungen zu schaffen, die Frauen und Männern zu gleichen Lernerfolgen führen.

16. Eine Nachbemerkung zum internationalen Vergleich

Im internationalen Vergleich liegen die deutschen Schülerinnen und Schüler in den hier untersuchten Fächern im Mittelfeld. Dies gilt sowohl für die beruflichen Schulen als auch für die gymnasiale Oberstufe, für die mathematisch-naturwissenschaftliche Grundbildung ebenso wie für die Leistungen in voruniversitärer Mathematik und Physik. Was sich in der Untersuchung zur Sekundarstufe I abgezeichnet hatte, setzt sich in der Sekundarstufe II fort. Die Abstände zu den führenden europäischen Nachbarstaaten werden eher größer als kleiner. Im Bereich der Spitzenleistungen sind die Unterschiede besonders auffällig. Aber auch der Anteil sehr testleistungsschwacher Schüler ist in Deutschland im Vergleich zu einigen Nachbarländern ungewöhnlich groß.

Die Befunde des internationalen Vergleichs geben zunächst nur Zustandsbilder. Aber schon in dieser Funktion sind sie eine Realitätskontrolle gegenüber gängigen

Vermutungen über vermeintliche Tatbestände und Zusammenhänge. TIMSS hat in der Öffentlichkeit Aufmerksamkeit gefunden und nicht wenige Protagonisten im Bildungssystem verärgert, weil deren Programme nicht durch die Ergebnisse gestützt wurden. Insofern liefert der internationale Vergleich gerade innerhalb Europas auch eine realitätsbezogene Messlatte für das, was in schulischen Bildungsprozessen bei vergleichbarer Investition von Zeit und Geld erreicht werden kann. Die Befunde lassen sich aber auch – und zwar in sehr differenzierter Weise – im Hinblick auf die normativen Vorgaben der Lehrpläne und – soweit es um Mindeststandards geht – im Hinblick auf die Erwartungen von Abnehmern interpretieren. Auch in dieser Hinsicht ist TIMSS ein Beitrag zur Realitätsprüfung. Wenn die Befunde die gesellschaftlich und politisch gesetzten Normen nicht treffen, ist dies allerdings noch lange kein Menetekel für die wirtschaftliche Prosperität der Bundesrepublik, wie in der öffentlichen Rezeption von TIMSS immer wieder unterstellt wird. Der Zusammenhang von in der Schule erzielten Fachleistungen und wirtschaftlicher Entwicklung ist mehr als indirekt und in einer Kausalkette auch nicht nachweisbar. Eher lassen sich die TIMSS-Ergebnisse als Indikator für die gesellschaftliche Wertschätzung von Bildung und die Bereitschaft, Mittel, Zeit und Anstrengung in Lernprozesse zu investieren, deuten. Insofern mögen die Befunde zu Recht Anlass zur Nachdenklichkeit geben. Von einer Krise kann jedoch keine Rede sein. Die TIMS-Studie hat einen wichtigen Zweck erfüllt, wenn sie dazu anregt, den reichen Schatz des bereits verfügbaren Wissens über Lernprozesse für eine produktive Schul- und Unterrichtsentwicklung verstärkt zu nutzen und, wo Wissenslücken bestehen, systematisch nach Ursachen zu forschen. Die Diagnose hat nur dienende Funktion, und ohne Entwicklung ist sie nichts wert.

Fragt man nach analytischen Erträgen von TIMSS, liegt die Stärke der Untersuchung vielleicht weniger in Evidenzen, die spezifische Annahmen stützen, als vielmehr in der Entkräftung gängiger bildungspolitisch motivierter Erklärungsmuster. Einer der kaum zu überschätzenden Vorzüge von TIMSS ist die große Zahl der teilnehmenden Länder. Damit wird die organisatorische Varianz in einer Weise erhöht, die man in großen, föderal organisierten Staaten nicht vorfindet. Auf dieser Basis legen die Befunde Schlussfolgerungen nahe, die auch in Deutschland helfen können, Problemzonen neu zu lokalisieren. So zeigt der internationale Vergleich, dass die Organisationsstruktur von Schulsystemen und die mathematisch-naturwissenschaftlichen Leistungsergebnisse praktisch unabhängig voneinander variieren. Ob ein Schulsystem zentral oder dezentral verwaltet wird, ob es die Ganz- oder Halbtagschule präferiert, ob es gegliedert oder integriert organisiert ist, hat offenbar für die Ertragslage des Unterrichts keine eigenständige Bedeutung. Zukunftsaufgabe ist die Weiterentwicklung von Schule und Unterricht in situationsangemessener Form.

Ein weiterer Befund, der sich gut in das Ergebnismuster der Schul- und Unterrichtsforschung einfügt, ist das Resultat, dass sich kein einzelner Faktor identifizieren lässt, von dem man eine durchschlagende Verbesserung der mathematisch-naturwissenschaftlichen Leistungsergebnisse erwarten könnte. Die Ergebnisse von Bildungsprozessen hängen von komplexen Bedingungsgefügen ab, deren Elemente oftmals wechselseitig, allerdings nicht beliebig, substituiert werden können. Es gibt keinen Königsweg zum Erfolg, sondern vielfältige Zugänge. Die optimale Lösung ist nur im Zusammenspiel sehr unterschiedlicher Faktoren – dazu gehören Kultur, Geschichte, Personen, Sache und Situation – zu ermitteln. Diese Vorstellung widerspricht jedem pädagogischen Dogmatismus. Sie hat aber den Vorzug, durch Ergebnisse der Schul- und Unterrichtsforschung gut belegt zu sein.

Versucht man, aus den Befunden von TIMSS konstruktive Schlussfolgerungen zu ziehen, so sprechen die deskriptiven Ergebnismuster dafür, systematische Erklärungen für Leistungsunterschiede sowohl im Unterricht selbst als auch in den subsidiären Systemen zu suchen. Unterricht funktioniert nicht voraussetzungslos. Er ist auf den Rückhalt in unterschiedlichen Stützsystemen angewiesen, deren Zubringerleistungen den Unterrichtserfolg nicht unwesentlich mitbestimmen. Zu diesen Stützsystemen gehören zunächst der schulische Kontext selbst, der hinsichtlich der Professionalität des Lehrkörpers, der akademischen Kultur und Wertschätzung des jeweiligen Fachunterrichts sehr variieren kann. Dann das Elternhaus, das die schulische Arbeit in unterschiedlicher Weise mitträgt, und nicht zuletzt die Bildungsnähe und Schulkultur einer Gesellschaft insgesamt, die den breiten normativen Rahmen definieren, in den das Lernen in der Schule und der Fachunterricht eingebettet sind. Sie bestimmen die generelle Wertschätzung schulischen Lernens, die Bereitschaft zur Anstrengung und Ausdauer und nicht zuletzt die Qualitätserwartungen. Will man Unterricht optimieren, muss man den systemischen Gesamtzusammenhang im Auge behalten, auch wenn man die Aufmerksamkeit zunächst auf einzelne Problemzonen konzentriert.

