

TIMSS 2 elaboriert

**Eine didaktische Analyse von Schülerarbeiten
im Ländervergleich Schweiz / Deutschland**

Dem Fachbereich Mathematik der
Universität Duisburg-Essen,
Campus Duisburg,

zur Erlangung des akademischen Grades eines
Dr. paed.
eingereichte Dissertation,

vorgelegt von
Reinhard Woschek
aus Mönchengladbach

Referent: Prof. Dr. Günter Törner

Korreferent: Prof. Dr. Manfred Leppig

Tag der mündlichen Prüfung: 18. Juli 2005

Inhaltsüberblick

Abstract	I
Kurzfassung	II
Inhaltsverzeichnis	III
Verzeichnis der Abbildungen und Tabellen	VII
Zur Genese	1
Entwurf der Studie	37
Realisierung und Datenpräparation	58
Detailanalysen zur Bearbeitungsqualität	67
Detailanalysen zur Geometrie	92
Ergebnisse	119
Ausblick und weitere Forschungsfragen	124
Literaturverzeichnis	126
Aspekte des Schulsystems in der Schweiz	135
Anhang: Ergänzende Daten zur Studie	160

Abstract

It was the major aim of the study to find out why the Swiss ranking in TIMSS 2 was so much better than the German ranking. Therefore a test population of 84 pupils in Switzerland and 107 pupils in Germany of comprehensive secondary schools had to work on twelve assorted TIMSS-2-items.

The main differences to the TIMSS test-design were that they had to elaborate their answers instead of just ticking marks, that they were given plenty of time, and that they had to comment their processings explicitly. Further on they had to comment curricular circumstances, even if they just guessed an answer. Reasons for guessing were raised, too. By the process of evaluating the elaborated solutions it was possible to discriminate different kinds of mistakes, and to find out whether special kinds of mistakes were due to special attributes of the questions posed.

One of the main results of the study is that students in Switzerland worked accurately and perseveringly even in topics that were not curricularly valid. This can be proved especially in geometry where the Swiss students handled topics that had not yet been subjects of tuition. They usually commented issue-related, while German pupils tended to skip problems, or in their annotations, shifted their blames.

Seemingly by the lot of time given geometry achievements became noticeably better.

Curricular mismatches don't seem to be an essential reason for the ranking difference considered.

Accepting little or additionally severe mistakes does not change the distance in ranking. Instead, it is shifted evenly.

Guessing has negligible effects and success is statistically even distributed.

Summarizing, the ranking difference between Switzerland and Germany is not merely an instructional effect, but might be due to educational backgrounds, let it be sociocultural effects or else.

Another substantial result of the study at hand is that especially in geometry, a lot of errors may be explainable by the type of questions given that led the students to the wrong trail. Might be this was an intention of the TIMSS-designers to test the students' stability. But this, naturally, can't explain the ranking differences considered.

Kurzfassung

Wesentliches Ziel der Studie war es, wissenschaftlich zu hinterfragen, warum die Schweizer Ranking-Ergebnisse bei TIMSS 2 so deutlich besser waren als die der deutschen Schüler. Dazu bearbeitete eine Testpopulation von 84 gymnasialen Schülern aus der Schweiz und 107 Schülern aus Deutschland jeweils zwölf gemischte TIMSS-2-Aufgaben.

Die Unterschiede zum TIMSS-Testdesign waren: die Schüler sollten ihre Aufgabenbearbeitungen elaboriert darlegen, statt nur Multiple-Choice-Antwortvorgaben anzukreuzen. Sie hatten dazu ausreichend Zeit, und sie mußten ihre Lösungswege kommentieren. Zusätzlich sollten sie curriculare Gegebenheiten kommentieren, selbst wenn Antworten nur geraten wurden. Auch für die Gründe des Ratens wurden Fragen erhoben.

Durch die Auswertung dieser elaborierten Aufgabenbearbeitungen wurde es möglich, verschiedene Fehlerarten zu isolieren und zu analysieren, ob spezielle Fehlertypen den Aufgabenstellungen zuzuschreiben waren.

Eines der Hauptergebnisse ist, dass die Schüler in der Schweiz sauberer und ernsthafter arbeiteten; sogar in Teilbereichen, die nicht curriculum valide waren. Dies gilt speziell für den Teilbereich Geometrie, wo sie Aufgaben bearbeiteten, die noch nicht im Lehrplan standen. Sie gaben meist sachbezogene Kommentare. Deutsche Schüler neigten allgemein dazu, Aufgaben auszulassen oder in ihren Kommentaren die Schuld für die Nichtbearbeitung abzuweisen.

Anscheinend durch das größere Zeitfenster bedingt verbesserten sich die Geometrieleistungen gegenüber TIMSS teilweise deutlich.

Curriculare Gründe schienen insgesamt keine wesentliche Bedeutung zu haben.

Der Rankingabstand aus TIMSS2 wurde stabil reproduziert. Eine leichte oder auch weitergehende Fehlertoleranz verschob den Abstand gleichmäßig.

Raten hatte vernachlässigbaren Einfluß und war im Erfolg statistisch gleichmäßig verteilt.

Ein weiteres wesentliches Ergebnis ist, dass speziell in der Geometrie eine Reihe von Fehlern durch die Fragestellung hervorgerufen wurde, die die Schüler auf die falsche Fährte führten. Es mag sein, dass dies in der Absicht der Designer von TIMSS lag.

Allerdings kann dies die Ranking-Unterschiede aus der Studie nicht erklären.

Inhaltsverzeichnis

1 ZUR GENESE	1
1.1 Einleitung	1
1.2 Forschungsmotivation	3
1.3 Struktur und Inhalt der Studie	4
1.4 Datenquellen	6
1.4.1 TIMSS-Datenquellen	6
1.4.2 Weitere Datenquellen und Zitierhinweise	7
1.5 TIMSS und die Folgen	7
1.6 TIMSS-Kritik	9
1.6.1 Aufgabekritik	10
1.6.2 Kritik der Testform	11
1.6.3 Testkritik	12
1.6.4 Verfahrenskritik	14
1.6.5 Weitere Aspekte der Kritik	15
1.6.6 Zusammenfassung	16
1.7 Bildung, Standards, Kompetenzen	17
1.7.1 Bildungsbegriff und Bildungsinhalte	17
1.7.2 Historischer Rückblick	19
1.7.3 Zur Standards-Diskussion	20
1.7.4 Mathematische Literacy	24
1.7.5 Standards und Bildungsstandards	27
1.7.6 Bildungspolitischer Diskurs nach TIMSS und PISA	28
1.8 Forschungsfragen	32
1.8.1 Zur Datengewinnung und Aufbereitung	33
1.8.2 Zur Interpretation der Ergebnisse	33
1.9 Planung der Untersuchung	34
2 ENTWURF DER STUDIE	37
2.1 Probenumfang	37
2.2 Datenerhebung	38
2.3 Testpopulation	39

2.3.1	Vereinbarungen zur Nomenklatur	40
2.3.2	TIMSS-Aufgabenverteilung	41
2.3.3	Kriterien zur Aufgabenauswahl	42
2.4	Nationale Adaption	45
2.5	Curriculumkonforme Zusatzaufgaben	47
2.6	Testhefte	47
2.6.1	Didaktische Funktion	47
2.6.2	Homogene Aufgabenverteilung	48
2.6.3	Schülerinformationsblatt	50
2.6.4	Schülerlösungsblatt	51
2.6.5	Schülerfragebögen	52
2.7	Lehrerbegleitbuch	53
2.7.1	Lehrerinformationsblatt	53
2.7.2	Lehrerfragebogen	53
2.8	Probelauf in Deutschland	55
2.9	Präparation des Testmaterials	57
3	REALISIERUNG UND DATENPRÄPARATION	58
3.1	Zeitplan	58
3.2	Schulbeschreibungen	58
3.3	Klassenräume	59
3.4	Testverlaufsprotokolle	60
3.5	Bewertungsschlüssel	62
3.6	Datenpool	65
4	DETAILANALYSEN ZUR BEARBEITUNGSQUALITÄT	67
4.1	Vergleich der Altersstrukturen	68
4.2	Schülerleistungsvergleich nach TIMSS-Maßstab	69
4.3	Mängeltoleranter Schülerleistungsvergleich	71
4.4	Ratehäufigkeit und Rateerfolg	73
4.5	Bearbeitungsintensität und Bearbeitungsqualität	76

4.6 Lehrer- und Schülerquestionnaires	79
4.6.1 Informationsgehalt und Forschungsfragen	79
4.6.2 Synchrone Kodierung	82
4.7 Zu den Lehrerquestionnaires	84
4.7.1 Präferenzen bei der Kommentierung	85
4.7.2 Curriculumvalidität aus Lehrerperspektive	86
4.7.3 Curriculare Schwerpunkte im Ländervergleich	88
4.8 Zu den Schülerquestionnaires	89
5 DETAILANALYSEN ZUR GEOMETRIE	92
5.1 Globale TIMSS - Leistungsprofile	92
5.2 Exemplarische Detailanalysen von Geometrieaufgaben	94
5.2.1 Auswahlkriterien	94
5.2.2 Winkelmaß (TIMSS-Aufgabe M07)	96
5.2.3 Kongruente Dreiecke (TIMSS-Aufgabe K08)	103
5.2.4 Koordinaten 1 (TIMSS-Aufgabe J16)	106
5.2.5 Koordinaten 2 (TIMSS-Aufgabe I08)	109
5.2.6 Ähnlichkeit (TIMSS-Aufgabe P09)	113
6 ERGEBNISSE	119
6.1 Globale Befunde	119
6.2 Einzelbefunde	120
7 AUSBLICK UND WEITERE FORSCHUNGSFRAGEN	124
8 LITERATURVERZEICHNIS	126
9 ASPEKTE DES SCHULSYSTEMS IN DER SCHWEIZ	135
9.1 Vorbemerkungen	135
9.2 Ausgewählte demographische Daten zur Schweiz	140
9.3 Verfassungsmäßige Verankerung	140
9.4 Interkantonale Zusammenarbeit	141
9.5 Verwaltung	142
9.6 Schulaufsicht	143

9.7	Finanzierung	144
9.8	Beratende Organe	145
9.9	Obligatorische Schule	145
9.10	Übersicht Bildungswesen	146
9.11	Primarschule	146
9.12	Sekundarstufe I	148
9.13	Übergänge, Abschlüsse	149
9.14	Curriculum	150
9.15	Evaluation	151
9.16	Lehrkräfte	152
9.17	Statistik Sekundarstufe I	152
9.18	Lehrplanweiterentwicklung	153
9.19	Beispiele aktueller Forschungsfragen	153
9.20	Internet-Links zu Bildungsorganisationen in der Schweiz	156
10	ANHANG: ERGÄNZENDE DATEN ZUR STUDIE	160
10.1	Verwendete TIMSS-Items	160
10.2	Testheft A: Schülerinformationsblatt	165
10.3	Lehrerbegleittext	167
10.4	Lösungsblatt und Kodierschlüssel zu Testheft A	168
10.5	Transskripte der Lehrerquestionnaires	170
10.6	Transskripte der Schülerlösungswege bei Aufgabe M 07	171
10.7	Beispielhafte Transskripte von Schülerkommentierungen	173
10.8	Zum Themenkomplex „testwiseness“	181
10.9	Details zur „mathematical literacy“	181

Verzeichnis der Abbildungen und Tabellen

Abbildung 1: Anzahl nicht bearbeiteter Aufgaben im Vergleich GER/CH	78
Abbildung 2: Linearität (1)	123
Abbildung 3: Geradenproblem.....	112
Abbildung 4: Linearität (2)	127
Abbildung 5: Zur Ähnlichkeit.....	116
Abbildung 6: Verhältnisberechnung.....	117
Abbildung 7: Weg zur Verhältnisberechnung	118
Abbildung 8: Proportionale Analogie	130
Abbildung 9: graphische Übersicht Bildungswesen Schweiz	146
Tabelle 1: Am Test partizipierende Schweizer Schulen.....	40
Tabelle 2: TIMSS-Aufgabenverteilung nach inhaltlichen Bereichen	42
Tabelle 3: TIMSS-Aufgabenverteilung nach Anspruchsbereichen	42
Tabelle 4: Beispiele mathematischer Mikrobereiche in TIMSS-Aufgaben und zur Lösung erforderliche(s) Wissen/Fähigkeiten/Fertigkeiten ...	43
Tabelle 5: Verteilung der Inhaltbereiche und TIMSS-Aufgaben auf die Testhefte A bis E	48
Tabelle 6: Schwierigkeitsgrad der Aufgaben	49
Tabelle 7: Lösungsblatt zu Testheft A	63
Tabelle 8: Ergebniskodierung für Testheft A	64
Tabelle 9: Altersverteilung der Schüler nach TIMSS und der vorliegenden Studie	68
Tabelle 10: Leistungsvergleich nach TIMSS-Maßstab	70
Tabelle 11: mängeltoleranter Schülerleistungsvergleich (1).....	72
Tabelle 12: mängeltoleranter Schülerleistungsvergleich (2).....	72
Tabelle 13: geratene und abgebrochene Aufgaben	73
Tabelle 14: Absolute Zahl geratener und abgebrochener Aufgaben nach Ländern	75
Tabelle 15: Relativer Anteil geratener Aufgaben nach Ländern	75
Tabelle 16: Anzahl <u>nicht</u> bearbeiteter Aufgaben nach Aufgabentyp und Position im Testheft	77
Tabelle 17: Zusatzinformationen Lehrerquestionnaire	79
Tabelle 18: Zusatzinformationen Schülerquestionnaire.....	80
Tabelle 19: Komponenten des Kodiergektors – Lehrerquestionnaires	82
Tabelle 20: Komponenten des Kodiergektors – Schülerquestionnaires	84
Tabelle 21: Abweichungen von der Curriculumvalidität nach Lehrerquestionnaires	86
Tabelle 22: Ländervergleich gelöster curriculumvalider Aufgaben	87
Tabelle 23: Anzahl freier Kommentierungen.....	89
Tabelle 24: Nationale Leistungsprofile nach mathematischen Sachgebieten....	92
Tabelle 25: Schülerleistungen in der Geometrie	94
Tabelle 26: TIMSS-Lösungshäufigkeiten des Koordinatenproblems (1)	107
Tabelle 27: Lösungsstrategien beim Koordinatenproblem (1).....	108

Tabelle 28: TIMSS-Lösungshäufigkeiten des Koordinatenproblems (2)	110
Tabelle 29: Lösungshäufigkeit bei der Ähnlichkeitsaufgabe.....	114
Tabelle 30: Wahl der Antwortalternativen zum Ähnlichkeitsproblem bei TIMSS.....	114
Tabelle 31: Schulforum Schweiz	136
Tabelle 32: Zuteilung zu den Schultypen Schweiz	149
Tabelle 33: Statistik Schweiz Sekundarstufe	153
Tabelle 34: Bildungsorganisationen in der Schweiz.....	156
Tabelle 35: Gehalt der verwendeten TIMSS-Aufgaben.....	160
Tabelle 37: Lösungsblatt zu Textheft A	169
Tabelle 38: Ergebniskodierung zu Testheft A	169
Tabelle 39: Transskripte der Lehrerquestionnaires	170
Tabelle 40: Transskripte zu Aufgabe M07 GER Kl.7	171
Tabelle 41: Transskripte zu Aufgabe M07 CH Kl.7	172
Tabelle 42: Transskripte zu Aufgabe M07 GER Kl.8	173
Tabelle 43: Allgemeine Transskripte der Schülerkommentare	173

1 Zur Genese

1.1 Einleitung

Mit der „Dritten Internationalen Mathematik- und Naturwissenschaftsstudie“, kurz TIMSS¹ genannt, wurden

„seit vielen Jahren zum ersten Mal wieder Daten verfügbar, die für die mathematisch-naturwissenschaftlichen Fächer Auskunft über das in ausgewählten Jahrgangsstufen von deutschen Schülerinnen und Schülern erreichte Wissens- und Fähigkeitsniveau geben.“ (Baumert et al., 1997, S.17)

Danach liegen die Leistungen der deutschen Schülerinnen und Schüler „in einem breiten internationalen Mittelfeld“ (Baumert et al., 1997, S.18). Bildungssysteme anderer Länder erzielten im Bereich der Mathematik zum Teil deutlich bessere Rankings als Deutschland.

Dies wurde durch die nachfolgende PISA und PISA-E²-Studie, bei der die Mathematikleistungen der deutschen Schüler (fünfzehnjährig) nicht einmal mittelmäßig waren, im wesentlichen bestätigt.

Für die mathematikdidaktische Forschung stellt sich somit die Aufgabe, Indikatoren und bedingende Faktoren für die Diskrepanzen der gemessenen Mathematikleistungen in den nationalen Unterrichtssystemen zu ermitteln. Dazu soll die hier vorliegende Arbeit beitragen.

In Deutschland dürfte die Klassenarbeit mit zu berechnenden Aufgaben nach vorgegebenen Verfahren und mit eindeutigen Lösungen Standard für Leistungsmessungen im schriftlichen Bereich sein. TIMSS und PISA verwenden dagegen durchgängig Testmethoden, die deutschen Schülern weniger geläufig sind: bei TIMSS multiple-choice-Aufgaben, bei PISA offene Aufgaben, die Kompetenzstufen abfragen. Daraus resultiert z.B. die Frage, ob die gemessenen Schülerleistungen möglicherweise durch die Testform bei TIMSS oder PISA mit bedingt wurden.

Die Ergebnisse der TIMSS-Studie für Deutschland werden in dem oben zitierten Werk von Jürgen Baumert, Rainer Lehmann et al. „*TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich: Deskriptive Befunde*“ detailliert dargelegt. Es dient daher für die vorliegende Arbeit als Referenzwerk.

¹ „TIMSS“: Third International Mathematics and Science Study, TIMSS 1999 International Mathematics Report. Getestet wurden dreizehnjährige SchülerInnen. Vorausgegangen waren die beiden Studien F(irst)IMSS und S(econd)IMSS.

² PISA: Program for International Students Assessment, PISA-E (=Ergänzung) 2000

Um Schülerleistungen detailliert analysieren zu können, bedarf es eines gegenüber TIMSS geänderten Studiendesigns. Multiple-choice-Aufgaben können wenig Rückschlüsse auf die eingeschlagenen Lösungswege und die Art von Fehlern geben. Hier soll daher anhand elaborierter, also ausführlich niedergeschriebener und zusätzlich kommentierter Aufgabenbearbeitungen eine Detailanalyse von Schülerlösungs wegen erstellt werden, um daraus möglicherweise typische Fehlerquellen, Hinweise auf die Bearbeitungsqualität oder sonstige didaktische Erkenntnisse zu extrahieren.

Dazu wurden freigegebene Aufgaben aus TIMSS, Population 2 (i.w. dreizehnjährige Schüler) herangezogen und im Ländervergleich Schweiz zu Deutschland ausgewertet.

Wesentliche Modifikationen am TIMSS-Studiendesign sind folgende Aspekte:

- Aspekt der Motivation:
Die Schüler erhielten vorab umfassende Informationen über den Sinn des Tests.
- Aspekt der Zeit:
Es wurde gegenüber TIMSS erheblich mehr Bearbeitungszeit pro Aufgabe zur Verfügung gestellt.
- Aspekt der Lösungsanalyse:
Die Aufgabenbearbeitungen liegen in elaborierter Form vor, wodurch Strategien der Schülerlösungen transparent und analysierbar werden.
- Aspekt der graduellen Qualität:
Bei der Auswertung wurde ein kleinschrittig gestuftes Bewertungssystem verwendet, welches Aussagen über Fehlerarten und Fehlertiefe ermöglicht. Zusätzlich können Aussagen über die Bearbeitungsintensität extrahiert werden.
- Aspekt der Schülerperspektive:
Die Schüler erhielten zu jeder Aufgabe die Möglichkeit freier Kommentare. Damit wurden subjektive Angaben wie z.B. Aktualität oder Bekanntheitsgrad der Aufgaben erfasst, somit also das realisierte Curriculum *aus Schülersicht*.
- Aspekt der vergleichenden Curriculumsichten:
Neben den Schülern bewerteten auch die betreuenden Lehrer die Curriculumvalidität der Aufgaben. Dies ermöglicht Vergleiche zur unterschiedlichen Wahrnehmung von Curriculumvalidität.

Die Studie widmet im Ländervergleich Schweiz/Deutschland folgenden Bereichen ein besonderes Augenmerk:

- fehlertolerante Beurteilung von Aufgabenbearbeitungen, um die Auswirkung von Fehlern auf den Lösungserfolg abschätzen zu können,
- bedingende Faktoren außermathematischer Genese, da Anzeichen darauf hindeuten, dass Bearbeitungsqualität und -intensität länderspezifisch unterschiedlich sind,
- Vergleich von Schülerleistungen im Bereich der Geometrie, da sich hier die Unterrichtsinhalte und somit die curriculare Validität zum Testzeitpunkt länderspezifisch erheblich unterschieden,
- Vergleich der curricularen Validität aus Schüler- und Lehrersicht, da diese beiden Sichten nicht zwingend identisch sind,
- Einfluss der curricularen Validität auf die Lösungsqualität.

Außerdem wird an Beispielen untersucht, ob möglicherweise soziale Kontexte oder kulturelle Skripte in die Schülerleistungen einfließen.

Soweit nicht ausdrücklich anders angegeben beziehen sich die Ausführungen in dieser Arbeit auf die TIMS-Studie Mathematik für die Population 2. Die einzelnen Aufgaben werden anhand ihrer originalen TIMSS-Aufgabenummern wie etwa K04 oder R09 zitiert. Termini wie Schüler, Lehrer etc. stehen für beide Geschlechter. Wortkombinationen mit ‚multiple-choice‘ werden literaturüblich mit ‚MC‘ abgekürzt wie etwa MC-Items, MC-Tests etc. Transskripte zu Aufgabenbearbeitungen werden einschließlich der Rechtschreibfehler wiedergegeben. Englischsprachige Werke werden in Englisch zitiert. Englischsprachige, in den Publikationen von TIMSS verwendete Fachtermini werden bei ersten Auftreten in Deutsch und Englisch angeführt. Die tabellarischen Abkürzungen basieren jedoch aus Gründen der besseren Vergleichbarkeit mit den Originaltexten auf den englischen Termini.

1.2 Forschungsmotivation

Das Ranking deutscher Schüler anhand ihrer TIMSS-Leistungen entspricht nicht dem, was bisher nicht nur in Deutschland als gesicherter Wert deutscher Schulbildung angesehen wurde. Insbesondere werden nach TIMSS deutschen Schülern Qualitäten lediglich in bezug auf den mechanischen Kalkül, nicht aber hinsichtlich kreativer Aufgabenlösungen zugesprochen. Andere sprachlich ver-

wandte Nationen mit vergleichbarem kulturellem Anspruch wie die Niederlande oder die Schweiz erzielen im Ranking deutlich bessere oder sogar Spaltenplätze unter den partizipierenden Nationen

Damit steht neben dem Mathematikunterricht auch das deutsche Schulsystem auf dem Prüfstand. Es bleibt zu hinterfragen, ob die Einordnung im Mittelfeld primär als Leistungsdefizit im Vergleich zu anderen Nationen des europäischen oder des globalen Testraumes zu sehen ist. Möglicherweise ist die Positionierung in der Ranking-Skala eher sekundären Effekten wie Curriculuminvaliditäten, fremden Testformen oder ähnlichen Effekten zuzurechnen. Denkbar ist auch, dass das deutsche Schulsystem prinzipiell nicht darauf ausgelegt ist, den durch TIMSS gesetzten Standards und Methoden zu entsprechen.

Dieser letzte Einwand gegen internationale Leistungsvergleiche klingt eingängig. Dadurch werden jedoch die Unterrichts- und Erziehungsziele in einem gewissen Grade der Beliebigkeit ausgesetzt. Auch ignoriert er die Frage, welche Bedingungsfaktoren den unterschiedlichen Schülerleistungen verschiedener Nationen bei gleicher Testform zugrunde liegen.

Somit eröffnet sich schon vor der Problematik vergleichbarer „Leistungs“-Standards die Frage, ob in nationale Kontexte des jeweiligen Bildungssystems eingebettete Schülerleistungen objektiv vergleichbar sein können.

Didaktische Vorschläge zur Verbesserung von Mathematikunterricht – welche Zielrichtung auch immer darunter verstanden wird – erfordern Basiswissen über den Ist-Zustand. Dazu können vergleichende Analysen elaboriert vorliegender Schülerarbeiten einen Beitrag leisten.

1.3 Struktur und Inhalt der Studie

Kapitel 1 befasst sich nach Einleitung und Darlegung der Forschungsmotivation mit der Kritik, die der TIMSS-Studie entgegengehalten wird. Nach einem Exkurs über historische und aktuelle Zielvorstellungen des Mathematikunterrichts werden die Forschungsfragen für diese Arbeit dargelegt.

Kapitel 2 erläutert die technischen Details und begründet die getroffenen Entscheidungen zu Population, Methodik und Aufgabenselektion.

Kapitel 3 skizziert die Realisierung der Studie in der Schweiz und in Deutschland und beschreibt das für die Detailanalyse der Schülerarbeiten entwickelte Schlüsselsystem.

Kapitel 4 versucht, Schülerleistungen aus einem anderen Blickwinkel als dem der reinen Leistungsmessung zu ergründen. Dazu zählen Faktoren, die sich einer einfachen Messbarkeit entziehen. In der didaktischen Literatur werden sie auch als extrinsische oder auch als soziokulturelle Faktoren angesprochen. Dazu gehören – hier immer im Vergleich Schweiz/Deutschland – Analysen zu Schülerleistungen bei gestuft fehlertoleranter Aufgabenbewertung, Untersuchungen zu Ratehäufigkeit und Rateerfolg wie auch zur Intensität der Aufgabenbearbeitungen. Letztere legt einen Indikator für die Leistungsbereitschaft der partizipierenden Schüler nahe. Dieser Faktor ist ein nicht zu vernachlässigendes Maß bei der vergleichenden Interpretation der Schülerleistungen.

Kapitel 5 befasst sich mit ausgewählten Geometrieaufgaben. Die didaktische Detailanalyse der Schülerlösungswege soll Ursachen für die in TIMSS zum Teil systematisch falschen Lösungen ergründen³. Der Vergleich der Aufgabenbearbeitungen ist hier besonders aufschlußreich, da die Aufgaben für die Schweizer Schüler in weiten Bereichen nicht curriculum valide waren.

Kapitel 6 gibt einen zusammenfassenden Überblick über globale und Einzelbefunde der Arbeit .

Kapitel 7 gibt eine Ausblick auf weitere Forschungsfragen, die sich daraus entwickeln.

Kapitel 8 enthält das Literaturverzeichnis.

Kapitel 9 gibt dem interessierten Leser einen einführenden Überblick über das Schulsystem in der Schweiz. Allein schon auf Grund der Mehrsprachigkeit kann man deutliche Unterschiede zur Struktur des Bildungssystems in Deutschland erwarten. Diese erstrecken sich auf weite Bereiche wie etwa die Schulbuchentwicklung, die gegenseitige Anerkennung der Schulabschlüsse der Kantone etc.

Das letzte Kapitel 10 beinhaltet Daten und Ergänzungen zur Studie, die aus Gründen besserer Lesbarkeit nicht in den Hauptteil aufgenommen wurden. Sie sind für einen ersten Überblick nicht zwingend erforderlich.

Die Untergliederung der Kapitel erfolgt alphanumerisch (reine Dezimalklassifikation).

³ So wird etwa bei der TIMSS-Aufgabe K08 länderunabhängig zu etwa 40 bis 60 Prozent die falsche Antwort „A“ gewählt. Details dazu folgen im Kapitel 5.2.3 .

1.4 Datenquellen

1.4.1 TIMSS-Datenquellen

Die internationale TIMS-Studie „Third International Mathematics and Science Study“ der International Association for the Evaluation of Educational Achievement IEA wurde in drei Altersstufen oder Populationen durchgeführt: Population 1 (Grundschule Klassen 3 und 4: Mathematics Achievement in the Primary School Years, third and fourth grade, in Deutschland nicht durchgeführt), Population 2 (weiterführende Schulen Klassen 7 und 8: Mathematics Achievement in the Middle School Years, seventh and eighth graders) und Population 3 (Oberstufe: Mathematics Achievement in the Final Year of Secondary School).

Die TIMSS-Publikationen, die freigegebenen Aufgaben, das Daten- und Beileitmaterial sowie detaillierte Testresultate stehen online zur Verfügung (online, TIMSS, 1998). Die hier verwendeten Aufgaben und Daten zu TIMSS 2 wurden über diese Startadresse abgerufen. Unter anderem finden sich dort: Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study, TIMSS Technical Report, Volume I: Design and Development und TIMSS Mathematical Items: Released Set for Population 2 (Seventh and Eighth Grade).

In (Schmidt, 1996a) finden sich Ergebnisse einer Hintergrundstudie zur Entwicklung des TIMSS-Instrumentariums unter Einbeziehung der unterschiedlichen Curricula, Lehrpläne und Rahmenbedingungen. (Schmidt, 1996b) enthält umfangreiches Datenmaterial zur Strukturierung der länderspezifischen Curricula. Er analysiert und vergleicht die Unterrichtsgegenstände („topics“). Aspekte des Vergleichs sind unter anderem

- sehr frühe Einführung („very early introduced topics“),
- Zeit der unterrichtlichen Behandlung („topic duration“),
- Anzahl pro Klasse („topics per grade“),
- hervorgehobene Themen („emphasized curriculum topics“),
- Erhaltungszeit über Klassenstufen („topics duration“),
- Gewicht, welches auf ausgewählte Unterrichtsinhalte gelegt wird („curricular attention“).

In der gleichen Quelle findet sich auch eine tabellarische Übersicht der nationalen Projektkoordinatoren sowie der international zur Analyse der Curriculumvalidität der TIMSS-Aufgaben herangezogenen Schulbücher.

Über die TIMS-Studie und ihre Ergebnisse sind mittlerweile eine Reihe weiterer Publikationen erschienen. Zum Vergleich mit den in der vorliegenden Studie erhobenen Daten dient das bereits erwähnte Standardwerk „TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich, deskripti-

ve Befunde“ (Baumert, J. et al., 1997). Es enthält neben Informationen über Anlage und Durchführung von TIMSS auch Einblicke in die Fachleistungen in den mathematisch-naturwissenschaftlichen Fächern, Curricula, Motivation und einen Vergleich des Mathematikunterrichtes in Deutschland, Japan und den USA. Das Literaturverzeichnis verweist auf weitere Fundorte zu Vorbereitung, Realisierung und Auswertung von TIMSS.

Eine Evaluation für die Schweiz findet sich in (Moser et al., 1997).

1.4.2 Weitere Datenquellen und Zitierhinweise

Die wissenschaftliche Literatur in Buchform wird nach der Harvard-Notation zitiert. Daneben wurde auch das Internet für die TIMSS-Dateien und weitere Zitate als Datenquelle verwendet.

Die Problematik des Zitierens von Internetquellen wird ausführlich diskutiert in (online, Bleuel, 1996). Auf ihn verweist auch das Standardwerk „Die Form der wissenschaftlichen Arbeit“ (Standop, 1990). Bleuel regt die Kennzeichnung mit „online“ an. Diesem Vorschlag wurde hier gefolgt.⁴ Er entspricht internationalen Standards wie z.B. in „MLA Citation Form Guide“ (online, MLA, 2002).

Der formale Aufbau der Arbeit folgt streng den Vorgaben von Winter: „Wissenschaftliche Arbeiten schreiben“ (Winter, 2004, S. 83 ff).

Auf Grund der Sprachanbindung an die TIMSS-Kernsprache Englisch sowie der besseren Vergleichbarkeit mit den TIMSS-files werden zeitweise die originalen TIMSS-Bezeichnungen (auch mit den entsprechenden englischen Abkürzungen) verwendet.

1.5 TIMSS und die Folgen

Die Ergebnisse internationaler Schülerleistungsvergleiche wie z.B. TIMSS und das daraus abgeleitete Ranking der bewerteten Länder wurden nicht nur in den Fachwissenschaften, sondern auch in Politik und Öffentlichkeit fortdauernd diskutiert, und sie haben schon jetzt weit reichende Konsequenzen gezeitigt. In der „Stellungnahme der Deutschen Mathematiker-Vereinigung im Rahmen der Anhörung zu TIMSS bei der Kultusministerkonferenz am 26./27.6.1977 in

⁴ Danach wird nach dem „online“-Vermerk jeweils das Datum des downloads angegeben. So weit ein Datum der Erstpublikation angegeben ist, wird dieses im Literaturverzeichnis vermerkt. Weitere Hinweise finden sich in (Bleuel, 2001).

Bonn“ (online, Törner, 1998) wird betont, dass die Mathematik mittlerweile den Rang einer Schlüsseltechnologie einnimmt, Hochtechnologie mehr und mehr mathematische Hochtechnologie wird und daraus ein Bedarf an breiter mathematischer Grundausbildung resultiert. Einen Hinweis auf die „unzulängliche mathematische Vorbildung der meisten Studienanfänger“ veröffentlichte die DMV bereits 1977. Allerdings wird dort auch betont: „Die Ausbildung einer mathematischen Kultur bei unseren Kindern wird, wie auch TIMSS hervorhebt, noch von anderen Faktoren als der Schule allein geformt“ (Bauer, 1998).

In der „Erklärung der Fachverbände DMV, GDM, MNU vom 19.02.1997 werden diese Erkenntnisse wie folgt erörtert:

„Einige der Gründe liegen sicher außerhalb der Mathematik in der gesamtgesellschaftlichen Sicht von Schule, etwa in einer offenbar immer mehr zurückgehenden öffentlichen Wertschätzung für schulisches Lernen überhaupt und damit einhergehend Verminderung von verbindlichen Leistungsanforderungen, was natürlich im Mathematikunterricht negative Auswirkungen zeigt.“ (online, Törner, G. et al., 1998).

Weiter heisst es:

„Unter mathematikspezifischen Aspekten fällt auf, dass die Schülerinnen und Schüler aus Deutschland bei reinen Routineaufgaben aus Arithmetik und Algebra meist besser abgeschnitten haben als bei geometrischen Problemstellungen und dass sie vor allem bei Aufgaben, in denen ein inhaltliches Eingehen auf gegebene Problemsituationen oder ein selbständiges Anwenden von mathematischen Erfahrungen erforderlich waren, enttäuscht haben. ... Zu kurz kommen insbesondere das selbständige, aktive Problemlösen, das inhaltliche, nichtstandardisierte Argumentieren sowie das Herstellen von Verbindungen mathematischer Begriffe mit Situationen aus Alltag und Umwelt.“ (online, Törner, G. et al., 1998).

Weitere Hinweise zur TIMSS-Testdurchführung und Ergebnissen sowie Überlegungen zu Konsequenzen und Folgerungen, allerdings auch kritische Einwände zur TIMS-Studie finden sich z.B. bei Weigand in (Weigand, 1998).

Die im Dezember 1997 erschienene Expertise „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ gab mit „Probleme[n] und Prinzipien des Lehrens und Lernens in der Schule“ (BLK, 1997, S. 16 ff.) Leitlinien und Module zur Unterrichtsverbesserung vor. Landesweit werden Maßnahmen zur Qualitätssicherung und vorgelagert zur Entwicklung von Leistungsvergleichsverfahren ergriffen. Exemplarisch seien erwähnt das Projekt ‚Weiterentwicklung der Unterrichtskultur im Fach Mathematik‘ des Kultusministeriums in-Württemberg⁵, der hessische BLK-Modellversuch ‚Kooperative Modelle zur Entwicklung einer guten Unterrichtspraxis im Mathematikunterricht‘, der nordrhein-westfälische Ansatz zur Durchführung von Analysis-Vergleichsklausuren

⁵ Im Ansatz geht es dabei nicht primär um Änderungen mathematischer Inhalte, sondern um einen veränderten Unterrichtsstil.

in Grundkursen der Jahrgangsstufe 12 im Regierungsbezirk Düsseldorf (Bezirksregierung Düsseldorf, 1999), die Vergleichsarbeiten im Fach Mathematik in der S I in Mecklenburg/Vorpommern, der Mathematik-Test in der Jahrgangsstufe 8 in Rheinland-Pfalz, der bayerische Mathematik-Test in Realschulen und Gymnasien oder die Fortschreibung der Lehrausgangslagenuntersuchung in Hamburg in der Jahrgangsstufe 7. Eine detaillierte Darstellung aktueller nationaler Studien findet sich in (Kaiser, G., 2001).

Da die mathematischen und naturwissenschaftlichen Fachleistungen unmittelbaren Einfluss auf die beruflichen Einsatzmöglichkeiten der Schüler haben, wurden auch von den deutschen Arbeitgebern entsprechende Forderungen erhoben und Initiativen gestartet:

„Wir brauchen in Deutschland eine neue Bildungsoffensive, um einen Qualitätsprung in der Schule zu erreichen. Die Arbeitgeber haben eine solche Offensive gestartet: BDA@Bildung.de –eine Initiative der Arbeitgeber.“ (Hundt, 2000, S. 5)

So entstanden Initiativen zur Stärkung der MINT-Fächer (MINT = Mathematik, Informatik, Naturwissenschaften, Technik). Beispielsweise seien erwähnt die Organisation von ‚Excellence Centres‘ an Schulen oder die Durchführung auch für Schüler offener Kongresse wie ‚Mathe ist TOP‘ an der Gerhard-Mercator-Universität Duisburg unter der Organisation und Federführung von Prof. Dr. G. Törner⁶. Aus der Fußnote ist ersichtlich, dass sich eine breite Front von Institutionen in Folge von TIMSS um eine Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts bemüht.

Letztlich ist als Konsequenz der „schockierenden Verlautbarungen aus der TIMS-Studie“ (Hundt, 2000, S. 5) zu beobachten, dass das Prinzip Leistung wieder öffentlich diskutiert und eingefordert wird, nachdem die Bildungsschaft lange eher durch andere Aspekte und Gesichtspunkte geformt wurde.

1.6 TIMSS-Kritik

Schon bald nach ihrer Publikation wurden die TIMSS-Testmethodologie und auch die einzelnen Testaufgaben kritisch gewürdigt. Im folgenden werden einige dieser Aspekte der Kritik dargelegt.

⁶ „Mathe ist TOP“, DMV-Initiativ-Kongress zur Förderung von Hochinteressierten und Hochbegabten zum Internationalen Mathematischen Jahr 2000, 25.-27.9.2000, unterstützt durch das Bundesministerium für Bildung und Forschung, Berlin, Ministerium für Schule und Weiterbildung, Wissenschaft und Forschung, Düsseldorf, Siemens AG, München, Münchener Rückversicherungs-Gesellschaft AG und weitere Institutionen

Da die TIMSS-Ergebnisse im Original in Englisch vorliegen, werden zum besseren Verständnis der dort verwendeten Abkürzungen die englischen Fachbegriffe und ihre Abkürzungen übernommen. Dies ist nicht immer unproblematisch, wie eines der nachfolgenden Aufgabenbeispiele zeigen wird.

Die TIMSS-Aufgaben sind klassifiziert nach sechs inhaltlichen Kategorien („content categories“), drei Aufgabentypen und vier Schwierigkeitsstufen („performance expectations“). Details dazu finden sich unter Abschnitt 2.4.1.

1.6.1 Aufgabenkritik

Die Aufgabe K02 lautet:

K02. Ein Chemiker mischt 3,75 Milliliter von Lösung A mit 5,625 Milliliter von Lösung B zu einer neuen Lösung. Wieviel Milliliter neue Lösung sind das?

Diese Aufgabe wird nach TIMSS in die Kategorie „solving problems“ eingestuft, obwohl lediglich eine Addition von Dezimalzahlen gefordert ist.⁷ Der Terminus „problem solving“ ist hier nicht mit „Probleme lösen“, sondern mit „Bearbeitung anwendungsbezogener Aufgaben“ zu übersetzen. Andererseits wurde die Kategorie „solving problems“ ebenso komplexen Aufgaben wie z.B. dem Vergleich von zwei Mietangeboten zugeordnet, die auch nach deutschem Verständnis Elemente des Problemlösens enthalten.

Dagegen wurde eine Aufgabe, in der eine Funktion anhand einer tabellarischen Zuordnung erkannt werden musste, als „performing routine procedures“ eingestuft. Die Aufgabe lautet:

J18. Welches ist die fehlende Zahl in folgender Tabelle: 1|1, 2|?, 4|7, 7|13 ?

Als Antworten wurden vorgegeben 2,3,4,5 und 6. Es ist wohl kaum eine Routineaufgabe, die dahinter stehende Regel ($x|2x-1$) zu erkennen. Dies wird bestätigt durch den in TIMSS länderübergreifend ermittelten Schwierigkeitsgrad („international difficulty index“) von 594 (bei einem Indexbereich von etwa 350 bis 700 für die siebten Klassen⁸).

Weiter sind Verzerrungen bei der Übertragung in die Sprachen der beteiligten Länder denkbar, obgleich die Designer von TIMSS dieser Problematik große Aufmerksamkeit gewidmet haben.

⁷ Die Originalaufgabe lautet nach file BMItems.pdf: „Find total volume of mixture using rational number addition.“ (online, TIMSS 1998b). Eine Abnahme des Mischungsvolumens bei der Mischung unterschiedlicher Molekülgrößen bleibt damit unberücksichtigt.

⁸ Es gibt zwei Ausnahmewerte von 737 und 817.

1.6.2 Kritik der Testform

In der Kritik der Methodologie wurde TIMSS auch das gewählte Testformat vorgehalten. Nach TIMSS war “about one fourth of the questions (...) in free responses format” (online, TIMSS 1998a, S.11) danach also etwa 75% aller Aufgaben im MC-Form.

Hier könnten Länder im Vorteil sein, bei denen solche Testformen zum üblichen Unterrichtsprocedere gehören und die Schüler daher Strategien zur effektiven Bearbeitung entwickelt haben („testwiseness“-Strategien⁹). Ein dafür typisches Beispiel ist die Aufgabe Q08:

Q08. *In welcher Liste sind die Zahlen sortiert von der kleinsten zur größten?*

A. $0,345$ $0,19$ $0,8$ $\frac{1}{5}$

B. $0,19$ $\frac{1}{5}$ $0,345$ $0,8$

C. $0,8$ $0,19$ $\frac{1}{5}$ $0,345$

D. $\frac{1}{5}$ $0,8$ $0,345$ $0,19$

Auf den ersten Blick mag es scheinen, als wäre es zur Lösung der Aufgabe erforderlich, den Bruch in eine Dezimalzahl umzuwandeln und erst dann auf steigende Anordnung zu überprüfen. Da allerdings die Antworten A, C und D jeweils direkt erkennbar eine größere vor einer kleineren Dezimalzahl haben und die Antwort „keine von allen“ nicht vorgesehen ist, kann die korrekte Antwort B auch ohne Umwandlung Bruch zu Dezimalzahl gefunden werden. Die korrekt angekreuzte Lösung lässt daher keinen Rückschluss darauf zu, welchen Lösungsweg der Schüler eingeschlagen hat oder wie zeitaufwendig die Lösung für ihn war. Natürlich wird dabei angenommen, dass Schüler Dezimalzahlen nach ihrer Größe anordnen können.

Eine Aufgabe wie P14.:

P14. *Janis, Maija und ihre Mutter aßen einen Kuchen.*

*Janis aß $\frac{1}{2}$ des Kuchens. Maija aß $\frac{1}{4}$ des Kuchens. Ihre Mutter aß $\frac{1}{4}$ des Kuchens.
Wieviel ist vom Kuchen noch übrig? (Originaltext: „.... ist left“?)*

A. $\frac{3}{4}$ **B.** $\frac{1}{2}$ **C.** $\frac{1}{4}$ **D.** Nichts

⁹ Eine Definition von „testwiseness“ sowie ein Beispiel der eingesetzten Strategien findet sich in Anhang 10.9 .

erlaubt ebenfalls keine Rückschlüsse auf Schülerlösungswege. Sie lässt sich sicher mit den Methoden der Bruchrechnung angehen, aber auch von Grundschulkindern durch eine bildliche Vorstellung des Aufwands, also Wegnehmens, lösen. Dazu ist lediglich eine elementare Bruchvorstellung erforderlich, Bruchrechnen wird dann nicht benötigt. Auch hier lässt also die korrekte Lösung keine zwingenden Rückschlüsse auf eingesetzte Strategien zu. Eher werden aus dem Lernkontext entfernte und isolierte Fertigkeiten gemessen, wobei sich erst im Einzelfall durch didaktische Analyse elaborierter Aufgabenbearbeitungen erschließt, welche mathematischen Fähigkeiten oder Fertigkeiten zur Bewältigung der Aufgaben eingesetzt wurden.

Weitere Aufgabenbeispiele sowie ein „Interpretationsansatz auf der Basis stoff-didaktischer Analysen“ finden sich in „Wie kommen die deutschen TIMSS-Ergebnisse zustande?“ (Blum, Wiegand, 1998, S 28 ff.)

1.6.3 Testkritik

Die Möglichkeiten und auch die Grenzen internationaler Vergleichsuntersuchungen zum Mathematikunterricht wie TIMSS oder PISA¹⁰ sind weiterhin Gegenstand wissenschaftlicher Diskussion. So werden diskutiert die testtheoretischen Grundlagen (z.B. Lind, 1999 über die in TIMSS verwendeten Testmodelle), der Einsatz des Rasch-Modells und die Rateproblematik bei MC-Items (z.B. Meyerhöfer, 2004 und Lind, 2004), die Kalibrierung von Aufgaben mit freier Beantwortung oder auch die Interpretationsmöglichkeiten der Testaussagen unter Einbeziehung methodologischer Aspekte sowie der Rankingproblematik (Knoche, 1998). Wolfe untersucht Abhängigkeiten der Ergebnisse von verwendeten Items und Themengebieten (internationaler Meßfehler und internationales Vorurteil), (Wolfe, 1999), da denkbar ist, dass die Testergebnisse umso besser ausfallen, je näher sie am realisierten Curriculum¹¹ liegen.

Nicht nur international, sondern auch im nationalen Bereich wurden zunehmend Vergleichsuntersuchungen auf Länderebene etabliert. Das damit verbundene weite Spektrum offener Forschungsfragen führte zur Gründung eines eigenen Arbeitskreises „Vergleichsuntersuchungen im Mathematikunterricht“ bei der deutschen Gesellschaft für Didaktik der Mathematik (GDM).¹² Diskussionsbe-

¹⁰ PISA steht für ‚Programme for International Student Assessment‘. Weitere Ausführungen folgen in Abschnitt 1.7.4

¹¹ Zur Unterscheidung zwischen intendiertem (intended), implementiertem (implemented) und realisiertem (attained) Curriculum vgl. u.a. (Schmidt, 1996).

¹² Die Gründung erfolgte auf der 33. GDM Tagung in Bern 1999. In (Kaiser et al., 2001) findet sich ein Grundsatzreferat über die bildungspolitische Bedeutung von Schulleistungsvergleichen und eine Übersicht zu Vergleichsuntersuchungen in den Bundesländern.

darf besteht zum Beispiel, ob das Problem der adäquaten Berücksichtigung der Curricula so vieler verschiedener Länder aus den unterschiedlichsten Kulturkreisen überhaupt gelöst werden kann.“ (Kaiser, 1999, S. 265 ff).

Weiter stellt sich die Frage, inwieweit eine nationale Adaption der TIMSS-Aufgaben erfolgte. In der ersten Entwicklungskomponente zu TIMSS, einer international vergleichenden Curriculum-Studie, die Analysen von Lehrplänen und Schulbüchern sowie die Erfassung des tatsächlich unterrichteten Stoffes umfasste, wurden für Deutschland und die Schweiz eine sehr unterschiedliche Anzahl von Quellen ausgewertet (Schmidt, 1996b, S. 219, S 227,S 228).

Danach wurden für den Bereich Deutschland zur Analyse des intendierten Curriculums lediglich die Richtlinien des Landes Nordrhein-Westfalen herangezogen, zur Beurteilung der Curriculumvalidität nur zwei Lehrbücher eines Bundeslandes (Bayern), welche nach Curriculum, Unterrichtsorganisation und Abschlussprüfungen als repräsentativ für die gesamte Bundesrepublik angesehen wurden. Im Vergleich dazu wurden allein für die deutschsprachige Schweiz acht „curriculum guides“ und achtzehn „textbooks“ ausgewertet. Aus dieser Sicht sind Zweifel an der bundesweiten Curriculumvalidität der TIMSS-Aufgaben in Deutschland erlaubt. Keitel-Kreidt spricht vom „idealisierten internationalen Curriculum (...) als Standardmaß“ (Keitel-Kreidt, 1998). De Lange verweist am Beispiel USA auf ein grundlegendes Strukturproblem bei der Setzung eines solchen idealisierten internationalen Curriculums:

„The Netherlands concluded that TIMSS did **not** fit in a satisfying way to the intended/perceived/implemented curriculum. And that the format did not fit to the actual (familiar to the Dutch students) assessment formats. For those reasons a ‚national option‘ was discussed, a plan designed, money raised, and items designed...“ und resümiert „The actual TIMSS score at present for grade 8 stands at 541. This score within the national option would be around 585.“ (de Lange, 1998).

Als Erklärung für diese erhebliche Ranking-Differenz führt er an:

„Sofar the only explanation we have brought up is that the national context made children feel better and more motivated. Or put slightly different: it might be that the fact that the anchor items were culturally correct embedded made students far better.“ (de Lange, 1998).

Hier wird die Bedeutung der nationalen Adaption der Testitems hervorgehoben. Im Tenor wird vermutet, dass die Testergebnisse umso besser ausfallen, je näher die Tests am *realisierten* Curriculum sind. Damit entsteht jedoch ein strukturelles Problem des Aufgabendesigns, denn danach wäre auch eine Adaption der Testitems auf Ebene der Bundesländer erforderlich und nicht nur im nationalen Rahmen. Für mehrsprachige Länder ergäben sich zusätzlich multiple sprachbedingte Adoptionsanforderungen.

De Lange weist weiter darauf hin, dass:

„TIMSS makes clear, as does Pursuing Excellence and Splintered Vision, that there exist not one curriculum in the US, but thousands. This diversity itself is a major problem.“ (de Lange, 1998).

Diese von ihm ‚unfocused curricula‘ genannten Unterrichtskomplexe führen seiner Ansicht nach dazu, dass Lehrer verschiedene Curricula lehren müssen: einmal orientiert an Richtlinien oder Lehrplänen (‘textbook curriculum’), zum anderen an genormten Testvorbereitungsverfahren (‘standardized test preparation curriculum’).

Die im angelsächsischen Sprachraum verbreiteten detaillierten Standard-Wissensziele in einem Unterrichtsfach, – standards genannt – werden in 1.7.3 in Genese und Realisierung diskutiert.

Es ist denkbar, dass auch in Deutschland durch den beschränkten Umfang der für TIMSS analysierten Curricula und Schulbücher der föderalen Vielfalt in der Bundesrepublik nicht ausreichend Rechnung getragen wurde.

1.6.4 Verfahrenskritik

Die Durchführung des TIMSS-Tests erfolgte nicht unter den Bedingungen, unter denen in Deutschland regelmäßig schulische Leistungen eingefordert werden. Haenisch schreibt dazu:

„Man braucht sich dabei nur einmal die Situation zu vergegenwärtigen, in der die Schüler/innen bei diesem Test konfrontiert waren: Sie erhielten ein Testheft mit 70 Aufgaben, die sie in 90 Minuten zu lösen hatten. Sie bekamen dafür keine Noten, noch nicht einmal eine Rückmeldung über die von ihnen gelösten und nicht gelösten Aufgaben. (...) Wer glaubt denn wirklich daran, dass bei diesen Bedingungen die Schüler/innen 90 Minuten problemlösend ohne Unterlass ihre gesamte Kompetenz und Anstrengung hineingelegt haben, um am Ende möglichst viele Aufgaben zu lösen? (...) Ist es denn wirklich so abwegig anzunehmen, dass bei den oben beschriebenen Durchführungsbedingungen (keine Noten, keine Rückmeldung, keine Einsicht in den Sinn des Ganzen usw.¹³) die deutschen Schüler/innen die ganze Sache vielleicht nicht ganz so wichtig genommen haben, dass sie sich zwar damit beschäftigt haben, dass sie aber immer dann, wenn die Anforderung nicht mehr mit oberflächlicher Routine und Drübergehen bewältigt werden konnte, sie ihre Konzentration und Anstrengung nicht überbeansprucht haben.“ (Haenisch, 1998)

Ähnlich weisen auch die OECD/CERI-Studie „What Works in Innovation“ (OECD/CERI, 1994) und die Analyse von Keitel-Kreidt/Kilpatrick (Keitel-

¹³ Die Funktion der Rückkopplung betrachtet z.B. Jacobs in (Jacobs, 1998).

Kreidt, Kilpatrick, 1998) auf das Faktum der unterschiedlichen Rahmenbedingungen von Leistungsabfragen in verschiedenen Ländern bzw. Kulturen hin.

Haenisch (Hänisch a.a.O.) bedenkt hier soziale Skripte, die Einfluss auf die Testergebnisse haben können. Wie auch andere Autoren verweist er auf die Unerfahrenheit deutscher Schüler mit MC-Tests, „Testfallen“ mit attraktiven Falschaussagen oder auch auf das Problem der Bearbeitungszeit. Ein Test, der in 90 Minuten 70 Aufgaben bearbeitet haben will, also gerade 1 Minute 17 Sekunden pro Aufgabe zugesteht, ist sicher zeitknapp gehalten. Hier dürften testerfahrene Schülerpopulationen im Vorteil sein.

1.6.5 Weitere Aspekte der Kritik

De Lange weist darauf hin, dass es der Zweck eines internationalen Vergleichstests („alien“ test) ist, Unterrichtsverbesserung auf den nationalen Ebenen zu erreichen. Er sieht darin auch Motivationsprobleme:

„As the fit between TIMSS and any national curriculum cannot be quantified and the fact that we cannot estimate the effect on students motivation on getting involved with an ‚alien‘ test the representativity of TIMSS is, to some extend, questionable. This observation is relevant ... as we try to improve teaching at a national level, based on an international test. We should consider for future international comparative studies how to integrate ‚national‘ or ‚regional‘ options as a kernel.“ (de Lange, 1998).

Keitel-Kreidt betrachtet die Passung der nationalen Curricula mit TIMSS und betont:

„Der Test misst nur, wie gut das ‚nationale Curriculum‘ in das ‚internationale Curriculum‘ paßt! Aber wie gut lernt ein Schüler oder eine Schülerin innerhalb eines Systems dasjenige Curriculum, welches das System ihm tatsächlich anbietet?“ (Keitel-Kreidt, 1998).

Wulf lange weist darauf hin, dass TIMSS nur einen Bruchteil dessen ermittelt, was Schule leisten muss und leistet, und dass auch aus dieser Vergleichsuntersuchung keine Kausalbeziehungen zwischen Unterricht und Schülerleistungen deduziert werden können:

„Das so einprägsame und publikumswirksame Ranking in der TIMSS basiert ausschließlich auf der Fähigkeit der Schüler, die in der TIMSS ausgewählten und zum Einsatz gekommenen Aufgaben richtig bearbeiten zu können. ... dass die tatsächlich eingesetzten Aufgaben Fähigkeiten überprüfen, die sicher nur einen Ausschnitt und nicht das ganze Spektrum dessen berücksichtigen, was der Unterricht in Mathematik und in den Naturwissenschaften in einem allgemeinbildenden Schulwesen leisten muss. ...Die TIMSS-Untersuchung kann von ihrer Anlage her ausschließlich deskriptive Befunde ermitteln.“ (Wulf lange, 1998).

Kaiser bemerkt dazu in einem Aufsatz zu „Möglichkeiten und Grenzen internationaler Vergleichsuntersuchungen“

„Das Anregungspotential internationaler Vergleichsuntersuchungen steht außer Frage; allerdings erscheint eine Trendwende bei internationalen Vergleichsuntersuchungen nötig weg von immer größeren, komplexeren, stark methodologisch dominierten Vergleichsuntersuchungen hin zu regionalisierten Studien, die sowohl qualitative wie quantitative Komponenten beinhalten.“ (Kaiser, 1999).

Dazu setzt sie sich in (Kaiser, 2000) auch mit der Problematik qualitativer Begleituntersuchungen auseinander.

1.6.6 Zusammenfassung

Zu TIMSS 2 wurden kritisiert und diskutiert

- die für deutsche Schüler ungewohnte Testform,
- der mangelnde Informationsgrad der Probanden, wobei Auswirkungen auf die Motivation und damit auch auf die Qualität und Intensität der Aufgabenbearbeitungen vermutet werden,
- die Aussagekraft der Aufgabenbearbeitungen auch unter den vorgenannten Aspekten,
- die mangelnde Berücksichtigung sozialer Skripte,
- die nationale Adaption im Ländervergleich,
- das realisierte Testverfahren unter den nationalen Rahmenbedingungen,
- die möglicherweise nicht ausreichende Berücksichtigung der Diversität der Curricula in den einzelnen Bundesländern,
- die Aussagekraft einer reinen Rankingliste hinsichtlich des tatsächlichen Leistungsabstandes der partizipierenden Staaten,
- die Übertragbarkeit von TIMSS-Leistungen auf sonstige Fachleistungen,
- die aufgrund der rein deskriptiven Natur der Studie fehlende Möglichkeit zur Ableitung von Kausalbeziehungen.

Aus TIMSS lässt sich letztlich nur ableiten, dass unter dem vorgegebenen Testverfahren Länder mehr oder weniger erfolgreich waren in dem, was durch TIMSS erhoben wurde. Zusammenhänge mit einer Qualität von Mathematikunterricht oder mit Schülerleistungen in Mathematik sind daraus nicht zwingend konstruierbar.

Baumert et al. formulieren in ihrer Zusammenfassung deskriptiver Ergebnisse zu Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich (TIMSS 3) explizit:

Internationale Leistungsvergleiche geben deskriptive Auskünfte über Leistungsverteilungen in den an der Untersuchung teilnehmenden Ländern. Sie liefern je-

doch kein Tableau normativer Zielvorgaben. Aus dem Befund hervorragender Fachleistungen in einem Land ergibt sich keineswegs der Schluss, dass diese Exzellenz auch Maßstab in einem anderen Land sein müsse. Die internationalen Ergebnisse bedürfen der Interpretation im Kontext jeweils nationaler Bildungsvorstellungen. (Baumert et al., 1999, S. 89)

1.7 Bildung, Standards, Kompetenzen

Ergebnisse und Rankings internationaler Schulleistungsvergleiche wie TIMSS oder PISA sollten im Kontext deutscher und internationaler Bildungsvorstellungen gesehen und interpretiert werden.

Daher befasst sich Kapitel 1.7 kurz mit – auch historischen – Vorstellungen zum Bildungsbegriff. Danach wird der Idee und historischen Entwicklung von Standards im amerikanischen Raum sowie curricularen Setzungen, wie sie auch beim Design der TIMSS-Items mit eingebracht wurden, nachgegangen. Unter diesem Blickwinkel werden Unterschiede zwischen dem Testverfahren bei TIMSS und deutschen Leistungsabfragen deutlich. Danach wird eine Abgrenzung zu dem bei der PISA-Studie eingebrachten Modell der Kompetenzstufen gezogen, welches wieder andere Leistungsstufen zu Grunde legt und einfordert. Dabei wird eine Differenzierung zwischen den Begriffen ‚mathematical literacy‘ und ‚mathematische Grundbildung‘ vorgestellt, obwohl beide in der didaktischen Literatur gelegentlich auch synonym verwendet werden. Abschließend folgt eine exemplarische Darstellung bildungspolitischer Beschlüsse und damit auch didaktischer Vorgaben der Kultusministerkonferenz der Bundesrepublik Deutschland, wie sie aus der Standards- und Kompetenzstufendiskussion erwachsen sind.

Insgesamt ergibt sich ein erweiterter Blickwinkel, unter dem die TIMSS-Ergebnisse gedeutet und diskutiert werden können.

1.7.1 Bildungsbegriff und Bildungsinhalte

TIMSS wird als Auslöser einer Besinnung darüber gesehen, dass die deutschen Schülerleistungen in Mathematik nur mittelmäßig sind; als Konsequenzen werden z.B. veränderte Unterrichtsformen, sogenannte ‚offene Aufgaben‘ und verbesserte Rahmenbedingungen für Mathematik- und naturwissenschaftlichen Unterricht eingefordert. Dazu bedarf es aber eines Blicks auf die Problematik, was Mathematikunterricht leisten soll und was er leisten kann.

Dabei ist Unterricht in Mathematik nach Form und Inhalten eingebettet in den allgemeinen Bildungsdiskurs. Bildung und ihre Verwirklichung können nicht als

Konstante angesehen werden. So notiert z.B. Schaak in seinen „Bemerkungen zur Bildungsdiskussion in Schweden“:

„Es wird immer wieder darauf hingewiesen, dass ‚bildung‘ kein eindeutiger Begriff sei und dass jede Epoche immer wieder bestimmen müsse, was Bildung beinhaltet und wie sie verwirklicht werden müsse. ... Veränderungen im Hinblick auf Struktur und Volumen der "Wissensmenge" (etwa das sich alle 5 Jahre verdopelnde Wissen im naturwissenschaftlichen Bereich) machen eine neue Auswahl der Inhalte notwendig, ...“ (online, Schaak, 2000)

Dies wird leicht nachvollziehbar bei Betrachtung der Aspekte, die in die aktuelle Bildungsdiskussion mit einfließen wie etwa die rapide fortschreitende Technisierung und Computerisierung, die Globalisierung, die Umweltproblematik und auch politische Umstrukturierungen. Unterliegen danach Deutung, Wert, Inhalt und Verwirklichung von Bildung dem zeitlichen Wandel, so gilt dies umso mehr für die von der Gesellschaft eingeforderten und von der Schule zu vermittelnden Bildungsinhalte. Sie werden in den Lehrplänen ausformuliert in allgemeinen Bildungsinhalten und konkreten Lernzielen.

Dabei ist die Normierung von Unterrichtsinhalten sicher nicht absolut, sondern abhängig von der jeweils unterlegten Bildungstheorie. In der Theorie der materialen Bildung werden andere Inhalte hervorgehoben und als wertvoll erachtet werden als in der formalen oder kategorialen Bildung. Andererseits beruhen bildungstheoretische Ansätze auf Auffassungen von Bildung, wie sie von verschiedenen Philosophen wie Heidegger, Herder, Humboldt oder auch Kant vertreten wurden.

Die abstrakten und möglicherweise auch ideologisch unterfütterten bildungstheoretischen Ansätze können jedoch weder konkrete Ziele noch Inhalte des Mathematikunterrichtes begründen. Die Konkretisierung der bildungstheoretischen Modelle bleibt weiterhin Aufgabe der Didaktik, von Klafki formuliert als ‚Prinzip der Didaktik‘ vor methodischen Entscheidungen.

Unterrichtsinhalte unterliegen langfristig den Strömungen der Zeit und folgen zeitversetzt dem aktuellen Stand der didaktischen Diskussion. So hat sich die Intensität bei der Behandlung der Kegelschnitte, sei es algebraisch oder vektoriell, im Lauf der Zeit geändert. Genauso war die Vektorrechnung nicht immer im Pflichtkanon¹⁴ der Oberstufe enthalten. Auch die Intensität des Themenbereiches Stochastik schwankt in der Stoffobligatorik. Der Anteil der Geometrie am mathematischen Wissenskanon und die Art ihrer Vermittlung hat starken historischen Schwankungen unterlegen. Die Einbindung der Mengenlehre sogar schon in den Unterricht der fünften Klassen und das damit verbundene neue Vokabular hat Schüler, ihre mathematischen Fähigkeiten und Fertigkeiten und auch ihr Bild von Mathematik anders geformt als Schüler, die ohne diese Sicht- und Sprach-

¹⁴ Der Pflichtkanon hat sich eher zum Auswahlkanon entwickelt.

weisen aufwuchsen. Lediglich der Strukturaspekt der Mathematik scheint historisch stabil als zu vermittelnder Unterrichtsinhalt gesehen worden zu sein, wohingegen der Wirklichkeitsbezug der Mathematik sich nicht in den Vordergrund drängte.

1.7.2 Historischer Rückblick

Mathematische Inhalte wurden in historischer Zeit eher als Aufgabensammlungen denn als ausformulierte Stoffbereichsvorgaben tradiert. Dagegen legt die Stoffobligatorik des Wilhelm von Humboldt¹⁵ detailliert Unterrichtsinhalte fest. Trotz der eher sprachlichen Ausrichtung der Abiturianforderungen – „Latein, Griechisch, Französisch, Deutsch einerseits, Geschichte, Geographie, Mathematik, Physik und Naturbeschreibung andererseits“ (online, Tennenbaum, 2000) – gab er detailliert zu behandelnde Stoffbereiche oder nach heutiger Terminologie Stoffverteilungspläne vor. Der Unterricht in Mathematik war wie folgt vorgegeben:

„In V beginnen Algebra und Geometrie, in IV Theorie der Gleichungen und Geometrie nach dem 6., 11., 12. Buch des Euklid, in III Logarithmen und analytische Geometrie; geometrische Konstruktion fällt als zu zeitraubend weg und kann unter Leitung des Lehrers Objekt häuslicher Beschäftigung werden; in II Lehre von den Reihen, ebene und sphärische Trigonometrie, Kegelschnitte; in I Gleichungen 3. und 4. Grades, Anfangsgründe der unbestimmten Analytik, Fortsetzung der Lehre von den Reihen, Wahrscheinlichkeitsrechnung; daneben in der Hälfte der Stunden angewandte Mathematik, besonders die mechanischen Wissenschaften.“ (online, Tennenbaum, 2000)

Verständlich wird diese präzise Festlegung durch die Aufgabe, die den Unterrichtsinhalten zugewiesen war:

„Die Unterrichtsgegenstände bildeten keine zufällige Ansammlung zusammengewürfelter, interessanter Themen, sondern eine "organische Einheit", die dem ‚Organismus der Wissenschaften selbst‘ entsprach. Die ‚harmonische Ausbildung des Geistes‘ war gesichert, wenn der Schüler gleichmäßig an allen Fächern teilnahm. Den Schulen wurde ausdrücklich aufgetragen, sorgsam zu verhüten, daß sie nicht die unharmonische Bildung und die Einseitigkeit derselben begünstigen, welche durch zu rasches Voreilen des Schülers in einigen Lieblingsobjekten und sein unverhältnismäßiges Zurückbleiben in anderen entsteht‘. Eine harmonische Ausbildung aller Kräfte des Geistes durch Sprachen und Literatur, durch Mathematik und Realwissenschaften war das Ziel des Unterrichts.“ (online, Tennenbaum, 2000)¹⁶

¹⁵ Man kann Wilhelm von Humboldt als Initiator des Ediktes von 1812 ‚wegen Prüfung der zu den Universitäten übergehenden Schüler‘ wohl zu Recht als ‚Vater der Abiturprüfung‘ ansehen.

¹⁶ An gleicher Stelle findet sich ein bis ins Detail gehender Anforderungskanon der einzelnen Fächer, der Prüfungsfächer, der zu prüfenden Inhalte und sogar der zu verwendenden Prüfungssprache. So wurden etwa die griechischen Klassiker in Latein abgefragt.

Hier erfolgt eine normative Setzung von curricularen Inhalten nicht aus globalen mathematikspezifischen oder allgemeineren Lernzielen, sondern aus dem Bildungsverständnis heraus. Die didaktische Entscheidung ist der Theorie der materialen Bildung gemäß nach dem ‚Bildungswert‘ der Unterrichtsinhalte getroffen worden.

1.7.3 Zur Standards-Diskussion

TIMSS ist aus ihrer Genese sicher auch als Projekt zu sehen, welches aus der amerikanischen Standards-Diskussion etwa im ‚National Council of Teachers of Mathematics‘ (NCTM) erwachsen ist. Als Ausgangspunkt der Standards-Entwicklung wird oft der Bericht ‚A Nation at Risk‘ der NCEE (National Commission on Excellence in Education) gesehen, wo dargelegt wird:

"The educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a nation and a people. ... We have, in effect been committing an act of unthinking, unilateral educational disarmament." (online, NCEE, 1983).

Hier wird schon 1983 die gleiche Mittelmäßigkeit beklagt, wie sie nach den Ergebnissen von TIMSS zur Zeit in Deutschland Motor für aktuelle Bildungsanstrengungen geworden ist.

Hogan legt dar, dass es gerade Tests waren, die wesentliche Impulse zur Bewältigung der nationalen Erziehungskrise gaben:

In fact, the fires of national consciousness regarding an educational crisis probably have been fueled more by test results than by all other sources of information combined. (online, Hogan, 2000)

Daraus entwickelte sich in den USA ein Ruf nach wesentlich mehr Tests:

Furthermore, the report calls for more testing: more frequent testing, more kinds of tests in more different fields, and more attention to test results. Testing plays a prominent role in the Commission's plans for remedying the currently perceived problem-heady prospects for the measurement field.“ (online, Hogan, 2000)

Dieser Ruf nach häufigeren Tests könnte auch dadurch bedingt sein, dass permanenten Tests mit Lösungsrückfluß an die Schüler (‘feedback’) eine Verbesserung des Unterrichts zugesprochen wurde. So sollen Fehler korrigiert und damit Lernleistungen verbessert werden. Jacobs beschreibt die Lernwirksamkeit von Aufgabenpräsentationen sowie den Einfluß verschiedener Arten von Feedback¹⁷

¹⁷ Feedback-Arten sind z.B. KOR knowledge of results, KVR knowledge of correct result=korrekatives feedback, AUC answer until correct (auch MTF multiple try feedback genannt).

auf die nachfolgenden Lernleistungen (online, Jacobs, 2001). Der wesentliche Wirkungsmechanismus scheint dabei in einer Korrektur von Fehlern gesehen zu werden:

„... in test like events, feedback's primary importance is correcting errors.“ (online, Bangert-Drowns et al. 2001).

Daneben lassen sich jedoch auch andere Wirkmechanismen denken, warum häufige Tests Lernleistungen verbessern könnten.

Hogan sieht allerdings Probleme der Testverfahren, wie sie partiell auch in der vorliegenden Studie untersucht werden:

„... the report does not acknowledge the myriad of measurement problems lurking in the background. What problems? Nothing special, just the traditional problems of validity, reliability, setting standards, norms, and so on. But calling the problems "traditional" makes them no less real, no less difficult.“ (online, Hogan, 2000)

Dennoch hat sich daraus in den USA eine umfassende Diskussion entwickelt mit dem Bestreben, Standards zu setzen und das Erreichen gesetzter Standards testmäßig zu erfassen. Es sprengt hier den Rahmen, diese Diskussion ausführlich darzulegen oder die verwendeten Argumente zu hinterfragen, daher seien im folgenden lediglich einige online verfügbare Literaturstellen angegeben.

Umfangreiches, klassifizierendes Material zu ‚Mathematics Content und Performance Standards‘ von ‚kindergarten‘ bis ‚advanced trigonometry‘ bietet z.B. das ‚Mathematics Standard Committee of San Diego City‘ mit dem Anspruch von „standards as an attempt to define excellence, not minimum competencies...“ (online, MSC, 1999), der Fordham Report gibt ‚State Mathematics Standards‘ in „An Appraisal of Math Standards in 46 States, the District of Columbia, and Japan“ (online, Fordham Report, 2000), das McREL-Institute (Mid-continent Research for Education and Learning) diskutiert Bestrebungen „to survey and consolidate the many national- and state-level efforts to identify what K-12 students should know and be able to do in a variety of subject areas“ (online, McREL, 1999) verbunden mit einer historischen Übersicht zur Standards-Diskussion und Fortentwicklung in den USA ab 1983. Die Standards werden in einer gestuften Hierarchie, gegliedert bis hin zu Detailauflösungen, beschrieben.

Den deutschen Schülern dürfte mehrheitlich eine solche permanente Testpraxis weniger vertraut sein, da hier Klassenarbeiten oder Klausuren als zentrales Mittel zur Erfassung von Schülerwissen und -leistungen eingesetzt werden. Diese Instrumente orientieren sich jedoch eher an der Unterrichtspraxis des Mathematiklehrers, der die Klassenarbeiten oder Klausuren entwirft, als an zentralen Standards. Damit stellen sie ein eher auf die Unterrichtsauswahl des Fachlehrers

ausgerichtetes Kriterium zur Kontrolle von Mathematikleistungen dar, auch wenn sie nach curricularen Vorgaben ausgerichtet sind. Erst in neuerer Zeit wurden in NRW Vergleichsarbeiten als Steuerungselement zum Vergleich von Lernleistungen aufgebaut.

Zur Genese konkreter curricularer Vorgaben oder auch zum Umfang dessen, was einer mathematischen Literalität zuzurechnen ist, unterscheidet Baumert international drei „basale Organisationstypen“:

„das deutsche, kanadische und Schweizerische Modell „zentralisierter Lehrplanentscheidungen auf Länderebene (3 Länder einschließlich Deutschland)“, Entscheidungen auf lokaler Ebene (9 Länder einschließlich USA, dort 16000 regionale Distrikte) Entscheidung auf nationaler Ebene (29 Länder einschließlich Japan, dort zentral bestimmt vom Erziehungsministerium)“. (Baumert et al., 1977, S.181, Grafik mit Erläuterungen S.183 ff)

Er bemerkt dazu, es bestehe

„...wenig Zweifel daran, dass die transnationale curriculare Validität der Leistungstests einen neuralgischen Punkt international verglichenender Schulleistungsforschung darstellt. (Baumert et al., 1997, S.179)

Dennoch wird betont:

Trotz der national unterschiedlichen Entscheidungsverfahren lässt sich ein internationales mathematisches und naturwissenschaftliches Kerncurriculum identifizieren. Dieses Kerncurriculum hat in Mathematik festere Konturen als in den Naturwissenschaften. . (Baumert et al., 1997, S.177)

Nach Baumert ist dieses internationale Kerncurriculum maßgebend als Konstruktionsrichtlinie und als (externes) Validitätskriterium für TIMSS eingesetzt worden.

TIMSS hat auch der Problematik des realisierten Curriculums (Unterrichtsvalidität) ein besonderes Augenmerk gewidmet. Dazu wurden an Begleittests jeweils national valider Tests auch die Schüler anderer Länder gemessen. Baumert kommt zu dem Urteil:

„Insgesamt kann für das Fach Mathematik sehr zuverlässig gefolgert werden, dass Länderunterschiede in den Mathematikleistungen nicht durch nationale Curriculumbesonderheiten erklärt werden können.“ (Baumert et al., 1977, S. 191)

Das deutsche Bildungssystem gestaltet Leistungsdifferenzierung im wesentlichen durch seine verschiedenen Schulformen. Zur Sicherung gleicher curricularer Standards wurden daher etwa in Nordrhein-Westfalen in neuerer Zeit durch das Ministerium für Schule, Wissenschaft und Forschung des Landes Nordrhein-Westfalen (MSWF, jetzt MSJK, Ministerium für Schule, Jugend und Kinder)

Mechanismen zur Qualitätssicherung institutionalisiert. Diese scheinen Diskrepanzen zwischen verschiedenen Schulformen oder Schulen aufzudecken und Aspekte der Beurteilungsdiversität zu bestätigen. So heißt es im ‚Bericht des Ministeriums zur Auswertung des Verfahrens zur Qualitätssicherung im Abitur 2000‘ zwar einerseits, die Auswertung der Zweitkorrekturen durch Lehrkräfte anderer Schulen habe in allen Fächern „eine relativ hohe Übereinstimmung in der Einschätzung der fachlichen Standards in den Schulformen und schulformübergreifend zum Ausdruck gebracht“ (MSWF, 2001, S.12), dass aber andererseits „Schulen bei der Notengebung in höherem Maße die besondere pädagogische Situation der Lerngruppen berücksichtigen als dies die stärker am Standard orientierte Schulaufsicht tut“. (MSWF, 2001, S.12)

Die Genese der von der Schulaufsicht verwendeten fachlichen Standards bleibt offen; es kann aber vermutet werden, dass diese tradierten Standards entsprechen, die nur in mathematischen Teilaспектen variieren. Wohl gemessen an diesen tradierten Standards werden die Bewertungen diskriminiert nach „angemessen, zu gut, zu schlecht“ (MSWF, 2001, S.15). Die berichtete Integration von pädagogischen Gegebenheiten der Lerngruppen in die Bewertung schulischer Leistungen, hier speziell mathematischer ‚Leistungen‘ durch Noten eröffnet dagegen ein Spektrum der Notengebung, welches der intendierten Qualitätssicherung entgegenarbeitet.

Für das Abitur 2001 hat das MSWF einen Vergleich der Leistungskurse Mathematik realisiert. Die Genehmigung der nicht zentral vorgegebenen Abiturvorschläge obliegt in Nordrhein-Westfalen den Schulaufsichtsbehörden, die vorab über das Genehmigungsverfahren eine normierende oder Standards setzende Funktion ausüben; der landesweite Vergleich der eingereichten Aufgaben durch die Schulaufsicht könnte auch für die Mathematikdidaktik von Bedeutung sein.

Die vorliegende Studie will nach ihrer Anlage einen anderen Weg gehen: statt mit TIMSS aus den ‚test responses‘ auf (an einem Kerncurriculum gemessene) Schülerleistungen zu schließen, sollen aus den elaboriert vorliegenden Aufgabenbearbeitungen von den Schülern eingebrachte und eingesetzte mathematische Fertigkeiten oder Fähigkeiten extrahiert werden, wobei keinerlei klassifizierende oder schematisierende, erst recht keine graduell bewertende Vorgaben bezüglich zu erreichender Leistungsstandards gemacht werden. Die Studie will aus Schülerlösungen von TIMSS-Aufgaben rückschließen auf die hinterliegenden mathematischen oder auch außermathematischen Denkprozesse, die die Schüler zur Problemlösung eingesetzt haben. Die Analyse solcher kognitiver oder auch intuitiver Denkwege kann möglicherweise zu einer anderen Beurteilung von Schülerleistungen führen als ein formales TIMSS-Ranking nach vorgegebenen Standards.

In der Literatur werden solche Leistungen oder auch Fähigkeiten gelegentlich unter dem allgemein subsumierendem Begriff ‚abilities‘ gefasst. Im Fordham Report (online, Fordham, 2000) wird das word ability dezidiert für mathematische Fähigkeiten verwendet: ‚reasoning ability‘, ‚ability to deduce consequences‘, ‚ability to estimate‘ etc., aber auch allgemein für intellektuelles Potential wie in: ‚for students of ordinary ability...‘. ‚High ability‘ kann auch für Hochbegabung stehen.

Ein analoger summarischer Gebrauch für allgemeine ‚Fähigkeiten‘ findet sich auch in „Complex problem solving, intelligence, and learning ability“ (Beckmann, Guthke, 1995, S. 177-200) oder in „The structure of cognitive abilities in highly and moderately gifted young people“ (Steffens, Perleth, 1996, S. 113-125).

1.7.4 Mathematische Literacy

In der Nachfolge von TIMSS wurden mit PISA und PISA-E¹⁸ weitere Untersuchungen zu Mathematikleistungen durchgeführt. Dort war der Probenumfang erheblich höher als bei TIMSS.

In der Konzeption von PISA, wo unter anderem die mathematische ‚literacy‘¹⁹ und die mathematische Grundbildung 15-jähriger Schüler erfasst werden soll, wird den oben diskutierten curricularen Aspekten und ihrer Genese nur noch eine Nebenrolle zugewiesen. Vielmehr wird darauf abgestellt, dass Schüler mathematische ‚Kompetenzen‘ gewinnen und diese dann auch nutzen sollen

Die Definition mathematischer Grundbildung im Rahmen von OECD/PISA lautet:

„Mathematische Grundbildung (Mathematical Literacy): Mathematische Grundbildung ist die Fähigkeit einer Person, die Rolle zu erkennen und zu verstehen, die Mathematik in der Welt spielt, fundierte mathematische Urteile abzugeben und sich auf eine Weise mit der Mathematik zu befassen, die den Anforderungen des gegenwärtigen und künftigen Lebens dieser Person als konstruktivem, engagiertem und reflektierendem Bürger entspricht.“ (online, www.mpib-berlin.mpg.de, 2000)

Nach dieser Quelle werden die Begriffe ‚Grundbildung‘ und ‚Literacy‘ als gleichwertig angesehen und können etwa als ‚Modellierungsfähigkeit‘ gedacht

¹⁸ PISA steht für ‚Programme for International Student Assessment‘. Es nehmen teil 32 Länder, davon 28 OECD-Länder, mit typischerweise zwischen 4500 und 10000 getesteten Schülern pro Land. E steht für die in Deutschland durchgeführte Ergänzungsstudie.

¹⁹ Hier wird analog zu den Texten des Max-Planck-Institutes die englische Schreibweise übernommen.

werden. Dabei steht nicht das konventionelle Curriculum im Vordergrund, sondern die Anwendung oder auch Reflexion mathematischen Wissens in unterschiedlichen Kontexten.

Allerdings wird in der Literatur die ‚mathematical literacy‘ auch als Bestandteil einer mathematischen Grundbildung gesehen (vgl. dazu z.B. Neubrand, 2001). Im internationalen PISA-Framework wird ‚mathematical literacy‘ definiert als:

„Mathematics literacy is an individual’s capacity to identify and understand the role that mathematics plays in the world, to make well-founded mathematical judgements and to engage in mathematics, in ways that meet the needs of that individual’s current and future life as a constructive, concerned and reflective citizen.“ (OECD 1999, p 41)

„Mathematical literacy“ wird dabei

„diskutiert als Teilproblem innerhalb der umfassenderen Frage nach dem Beitrag der Mathematik zur allgemeinen Bildung. ... Oft wird der Begriff „mathematische Grundbildung“ verwendet, um diese weitere Sichtweise anzuseigen.“ (online, Neubrand et al., 2000)

Wesentliche Gesichtspunkte („major aspects“) einer mathematical literacy sind die Aspekte der ‚mathematische Kompetenzen (mathematical competencies)‘ und ‚mathematische Ideen (mathematical big ideas)‘. Curriculare Überlegungen, bezeichnet als „mathematical curricular strands²⁰“ sowie Situationen und Kontexte werden als weniger bedeutend („minor aspects“) bewertet, wenn auch offensichtlich zwiespältig:

„Of course one cannot, and should not, ignore traditional strands of the mathematics curriculum. This is why they are explicitly included as a minor organising aspect of the mathematical literacy domain in OECD/PISA. The mathematical curricular strands aspect can help to ensure a balance in the items and a reasonable spread of content in relation to the school curriculum. The content strands for OECD/PISA are: number; measurement; estimation; algebra; functions; geometry; probability; statistics; and discrete mathematics. (online, OECD, 2000a)

Die Zwiespältigkeit wird deutlich in der Formulierung der Ziele („outcomes“) dieser Untersuchung. An erster Stelle wird angeführt

„Outcomes: A basic profile of knowledge and skills among students at the end of compulsory schooling...“ (online, OECD, 2000b)

Augenscheinlich muss die Erfassung von competencies und big ideas doch zu mindest teilweise über die Abfrage von Wissen und Fertigkeiten („knowledges“ und „skills“) erfolgen, die dann in einem erweiterten Kontext interpretiert werden müssen.

²⁰ strand = Strähne, Faden, Litze, Handlungsfaden

Dennoch wird für PISA betont:

„Der Begriff ‚Grundbildung (literacy)‘ wurde gewählt, um zu betonen, dass mathematische Kenntnisse und Fähigkeiten, wie sie im traditionellen Curriculum der Schulmathematik definiert werden, im Rahmen von OECD/PISA nicht im Vordergrund stehen. Statt dessen liegt der Schwerpunkt auf der funktionalen Anwendung von mathematischen Kenntnissen in ganz unterschiedlichen Kontexten und auf ganz unterschiedliche, Reflexion und Einsicht erfordernde Weise.“ (online, Baumert et al., 2000)

Das Konstrukt ‚mathematical literacy‘ beinhaltet also über mathematische Kenntnisse, Fähigkeiten und Fertigkeiten hinaus, diese im Kontextzusammenhang sinnvoll einsetzen zu können, eine Fähigkeit, die erst zusätzlich erworben werden muß:

„Der verständige funktionale Gebrauch von Mathematik muss vielmehr, nach allen gängigen Modellvorstellungen von sinnvollem Wissenserwerb, selbst Gegenstand des Lernens sein, wenn er erfolgreich sein soll.“ (online, Neubrand et al., 2000).

Eine weitere Klassifizierung der ‚mathematical literacy‘ nach Kategorien, Aspekten, Kompetenzen und Kompetenzklassen sowie eine Betrachtung der daraus resultierenden Umorientierung der Aufgaben und Zielvorgaben des Mathematikunterrichtes findet sich in Anhang 10.10 .

Es darf jedoch nicht übersehen werden, dass die Messung der literacy problembehaftet ist. So erzielte Finnland in PISA 2003 den ersten Platz im Bereich Naturwissenschaften und den 2. Platz im Bereich Problemlösen, dazu war das Land im Bereich Lesen sowohl 2000 wie auch 2003 auf dem ersten Platz. Allerdings betont der Artikel von Reinhard Demuth und Oliver Reiser zur 36. internationalen Chemie-Olympiade:

„Die Ergebnisse der viel diskutierten PISA-Studien wurden teilweise auf den Kopf gestellt. So errang Finnland nur zwei Bronze-Medaillen. Zumindest in Chemie scheint das dortige Schulsystem also keine Spitztleistungen zu fördern.“ (Demuth und Reiser, 02/2005, S.70)

Die deutschen Teilnehmer gewannen zwei Gold- und zwei Silbermedaillen bei maximal erlaubten vier Teilnehmern pro Land.

Dies weist wiederum darauf hin, dass international vergleichende Großuntersuchungen zu Bildungsstandards noch immer der wissenschaftlichen Hinterfragung bedürfen.

Die Tagung der Fachabteilung der GDM ‘internationale Großuntersuchungen‘ vom 26. und 27.11.04 bestätigt dies:

Der Anteil von MC-Items konnte durch den Einsatz von „fill-in“-Items niedriger gehalten werden. Rückmeldungen an die Schulen erfolgten in Form eines „informativen Feedbacks“ unter Vermeidung formaler Sanktionen und Leistungsbewertungen. (Büchter und Leuchter, 2004)

1.7.5 Standards und Bildungsstandards

Angesichts dieser Aspekte bleibt zur Beurteilung der Aussagequalität des TIMSS-Rankings zu fragen, an welchen Bildungsvorstellungen oder auch nur Unterrichtsinhalten sich die TIMSS-Testitems orientieren und ob dies zumindest partiell deckungsgleich war mit den jeweiligen Vorstellungen der beteiligten Länder über Bildungsziele und Unterrichtsinhalte. Diese Frage greift deutlich über eine curriculare Abdeckung der Unterrichtsinhalte hinaus. Sollten etwa schulformorientiert Unterrichtsschwerpunkte auf Handlungs- oder Projektorientierung gelegt worden sein, so lässt sich ein Unterrichtserfolg kaum durch MC-TIMSS-Items messen.

Für TIMSS 2 stellt sich somit die Frage, ob dort konkrete Kenntnisse, Fähigkeiten und Fertigkeiten des implementierten Curriculums im Zentrum standen oder eher der verständige Gebrauch dieser, wobei auch ein Fokus auf mathematical literacy indirekt durch curriculare Vorgaben gesteuert sein kann. Als generelle Frage bleibt, ob der Konstruktion der TIMSS-Items ein lernzielorientiertes Unterrichtsverständnis z.B. im Piaget'schen Sinne zu Grunde lag.

Zur Einschätzung der Aussagequalität des TIMSS-Rankings bleibt daher weiterhin zu hinterfragen, was genau durch TIMSS gemessen wurde. Dieser Frage vorgelagert ist die, was durch TIMSS gemessen werden sollte.

Aus dem technical report zu TIMSS entnimmt man dazu als Zielvorstellung:

„.... IEA studies have as a central aim the measurement of student achievement in school subjects...“ (online, TIMSS, 1998d)

Im Kontrast zum Fokus auf abstrakt gesehene Kompetenzklassen wird hier eine explizite Fokussierung auf das Curriculum als „broad explanatory factor underlying student achievement“ (online, TIMSS, 1998d) herausgestellt. Dies deutet auf eine im Vergleich zu Deutschland historisch anders gewachsene Einstellung zu Unterricht und Messung von Unterrichtserfolg, der hier als ‚achievement‘, also generalisierend als „Leistung“ der Schüler angesprochen wird.

Es lässt sich zusammenfassend formulieren, dass das Konstrukt der mathematischen Literalität, wie es für PISA verwendet wurde, abheben will von mathematischen Elementaroperationen und eher auf verbindendes Wissen setzt, aber dennoch elementare Fähigkeiten oder abilities als Kriterien vorausgesetzt und verwendet werden. Letztere haben zumindest den Vorteil der leichteren Erfassbarkeit: zu messen, ob ein Schüler das Addieren von Brüchen beherrscht, scheint einfacher standardisierbar als durch Testverfahren zu ermitteln, ob er versteht, kritisch zu analysieren, oder ob er mathematisches Sachwissen in sozialen, politischen und wirtschaftlichen Beziehungen sinnvoll anwenden kann.

1.7.6 Bildungspolitischer Diskurs nach TIMSS und PISA

In neuerer Zeit werden durch das Ministerium für Schule, Jugend und Kinder (MSJK)²¹ Mechanismen zur Qualitätssicherung institutionalisiert. So werden etwa Zweitkorrekturen der Abiturarbeiten von Lehrern anderer Schulen und Schulformen durchgeführt. Dies scheint Diskrepanzen im Procedere der Leistungsmessung zwischen verschiedenen Schulformen oder Schulen aufzudecken. So heißt es im Bericht des Ministeriums zur Auswertung des Verfahrens zur Qualitätssicherung im Abitur 2000 zwar einerseits, es gebe „eine relativ hohe Übereinstimmung in der Einschätzung der fachlichen Standards in den Schulformen und schulformübergreifend“ (MSWF, 2001, S.12), dass aber andererseits „Schulen bei der Notengebung in höherem Maße die besondere pädagogische Situation der Lerngruppen berücksichtigen als dies die stärker am Standard orientierte Fachaufsicht tut“. (MSWF, 2001, S.12)

Die Genese der dabei von der Schulaufsicht verwendeten fachlichen Standards bleibt offen; es kann aber vermutet werden, dass diese den oben angeführten tradierten Standards entsprechen und als historisch gegeben angesehen werden. Vermutlich gemessen an diesen tradierten Standards werden die Bewertungen diskriminiert nach „angemessen, zu gut, zu schlecht“ (MSWF, 2001, S.15). Die berichtete Integration von pädagogischen Gegebenheiten der Lerngruppen in die Bewertung schulischer Leistungen, hier speziell mathematischer ‚Leistungen‘ durch Noten eröffnet dagegen ein Spektrum der Notengebung, welches der intendierten Qualitätssicherung entgegenarbeitet (und so einen Vergleich erteilter Schulnoten mit Ergebnissen internationaler Schulleistungstests problematisch macht).

Für das Abitur 2001 hatte das MSWF einen Vergleich der Leistungskurse Mathematik vorgesehen. Die Genehmigung der nicht zentral vorgegebenen Abiturvorschläge obliegt in Nordrhein-Westfalen den Schulaufsichtsbehörden, die vorab über das Genehmigungsverfahren eine normierende oder Standards setzende Funktion ausüben, dennoch dürfte der nachfolgende Vergleich im landesweiten Maßstab durch die Schulaufsicht auch für die Mathematikdidaktik von Bedeutung sein

Zu Genese und konkreten Setzungen für mathematische Standards oder auch einer ‚mathematical literacy‘ heißt es in der von der Konferenz der Kultusminister (KMK) getroffenen „Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10), (Beschluss der Kultusministerkonferenz vom 04.12.2003)“ und die „Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (Beschluss der Kultusministerkonferenz vom 04.12.2003)“ neutral:

²¹ vormals Ministerium für Schule, Wissenschaft und Forschung des Landes Nordrhein-Westfalen (MSWF)

„Die Standards basieren auf fachspezifisch definierten Kompetenzmodellen, die aus der Erfahrung der Schulpraxis heraus entwickelt wurden. Sie beziehen international anerkannte Standardmodelle - u.a. theoretische Grundlagen der PISA-Studie und den gemeinsamen europäischen Referenzrahmen für Sprachen - ein. (online, KMK 2003a, S.4)

In der Pressemitteilung der KMK heißt es zur Frage der Genese von Standards weiter:

„Mit der konkreten Erarbeitung der Bildungsstandards waren Arbeitsgruppen betraut, die aus Fachdidaktikern und Schulpraktikern aus den Ländern in der Bundesrepublik Deutschland bestehen. Die Arbeiten der Fachkommissionen wurden durch eine Steuerungsgruppe unter wissenschaftlicher Beteiligung aus den Bereichen der Unterrichtsforschung und Fachdidaktik koordiniert. Die Ergebnisse der von der Bundesregierung in Auftrag gegebenen Expertise „Zur Entwicklung nationaler Bildungsstandards“ vom Februar 2003 wurden bei der Erarbeitung der Standards einbezogen.“ (online, KMK, 2003b, S.10)

Die Vielschichtigkeit des ‚standards‘-Begriffes in der internationalen Diskussion wird z.B. in (online, Klieme, E. et al., 2003, S.17) dargelegt. International werden ‚standards‘ meist ergebnisbezogen interpretiert und verwendet, wie etwa in Schweden, wo das nationale Curriculum ein

„auf Bildungsziele von mittlerem Abstraktionsgrad abzielendes Dokument“ (online, Klieme, E. et al., 2003, S.25)

ist, wogegen

„in den USA ... traditionell der Begriff Standard wesentlich synonym zu *Performance Standard* benutzt wurde, (vgl. McLeod, Stake, Schapelle, Mellissinos & Gierl 1996), und die Messung von Leistung stand im Vordergrund.“ (online, Klieme, E. et al., 2003, S.25).

Dies dürfte dem Blickwinkel entsprechen, unter dem die TIMSS-Aufgaben konzipiert wurden.

Klieme et al. legen in dieser Expertise dar, dass Bildungsstandards auf allgemeinen Bildungszielen beruhen. Die Bildungsstandards beschreiben dann diejenigen Kompetenzen (als grundlegende Handlungsanforderungen, nicht als Listen von Lehrstoffen und Lerninhalten), die Schüler bis zu einem gewissen Zeitpunkt erworben haben sollen.

„Gute‘ Bildungsstandards werden danach daran gemessen, ob sie folgende Bedingungen erfüllen:

1. Fachlichkeit, 2. Fokussierung 3. Kumulativität 4. Verbindlichkeit für alle 5. Differenzierung: 6. Verständlichkeit, 7. Realisierbarkeit (online, Klieme, E. et al., 2003, S.17)

Im Entwurf der KMK „Bildungsstandards im Fach Mathematik für den Hauptschulabschluss nach Klasse 9, (Stand: 23. 04. 04)“ werden in den Bildungsstandards angesprochenen Kompetenzen konkretisiert als

„Auseinandersetzung mit mathematischen Inhalten: mathematisch argumentieren, mit symbolischen, formalen und technischen Elementen der Mathematik umgehen, kommunizieren, mathematisch modellieren, Probleme mathematisch lösen, mathematische Darstellungen verwenden“ (online, KMK, 2004 S.7)

Die Autoren der amerikanischen nationalen Mathematiklehrervereinigung NCTM (National Council of Teachers of Mathematics, online erreichbar über <http://www.nctm.org>) stellten der traditionellen, auf Basisfähigkeiten beruhenden Sichtweise des Mathematikunterrichts schon bald stärker eine prozessorientierte Sichtweise gegenüber:

„....(bis heute hat NCTM keine testbasierten Standards im klassischen Sinne vorgelegt), sondern sie beinhalten eher eine Vision des guten Mathematikunterrichts, sind also Standards für professionelles Handeln von Mathematiklehrern (opportunity to learn-standards). Zugleich legen sie Inhaltsdimensionen des Mathematikunterrichts fest, sind also auch content standards. (online, Klieme, E. et al., 2003, S.26.)

Allerdings wird dort explizit formuliert:

„Es fehlt noch an Wissen darüber, was genau Standards bedeuten, wie sie zu formulieren sind, welche Rolle sie bei der Qualitätsentwicklung im Bildungssystem spielen können und wie ihre Einhaltung überprüft werden kann. Das vorliegende Gutachten hat daher zum Ziel, das Konzept der Bildungsstandards fachlich zu klären ...“ (online, Klieme, E. et al., 2003)

Umfangreiche Materialien zum „System der Standardsetzung und –überprüfung in NRW“, zum Projekt VerA (Vergleichsarbeiten), zu Parallelarbeiten und Lernstandserhebungen in Nordrhein-Westfalen nach den Vorstellungen des MJSK NRW sind über <http://www.bildungsportal.nrw.de/BP/Schule> zugänglich.

Für den Thementag auf der Bildungsmesse Didacta am 10.2.2004 hat das MJSK eine Schrift vorgelegt (aktualisiert am 19.4.2004), die den wissenschaftlich-didaktischen Forschungsdiskurs über Begrifflichkeiten wie Standards und Kompetenzen auf die Schulen übertragen und für die unterrichtliche Praxis aufbereiten will. Die Begrifflichkeiten werden dabei erweitert um die Termini „Bildungsstandards“ und „Kernlehrpläne“. Nach der vom Ministerium verwendeten Terminologie gilt für Bildungsstandards:

„Mit der Entwicklung von Bildungsstandards wird deswegen ein Perspektivenwechsel vorgenommen. Standards greifen allgemeine Bildungsziele auf und legen fest, welche Kompetenzen Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt ihres Bildungsgangs erworben haben sollen. Im Unterschied zu herkömmlichen Lehrplänen, die die Wege zur Zielerreichung beschreiben und strukturieren, geben Bildungsstandards die erwarteten Lernergebnisse vor („Outputorientierung“)“ (MJSK, 2004, S. 1)

Allerdings ergibt sich eine gewisse Widersprüchlichkeit dadurch, dass das MJSK NRW unter den übergreifenden Zielen von Standardsetzung und Standardüberprüfung an erster Stelle die

„Verbesserung der Schülerleistungen insgesamt und *insbesondere im unteren Leistungsbereich*“ (MJSK, 2004, S. 3)

aufführt, wogegen die zugrunde liegenden nationalen Bildungsstandards, veröffentlicht von der KMK, zitiert auf der gleichen Seite, ein anderes Maß anstreben:

„Die nationalen Bildungsstandards verstehen sich als sogenannte Regelstandards, die ein *mittleres Anforderungsniveau* festlegen.“ (MJSK, 2004, S. 3)

Den neu in die Begrifflichkeit eingeführten ‚Kernlehrplänen‘ wird eine Normierungsfunktion auf dem Wege zu standardisierten Abschlußprofilen zugewiesen:

„Kernlehrpläne in Nordrhein-Westfalen greifen die KMK-Bildungsstandards für den Mittleren Schulabschluss auf und zeigen durch Zwischenstufen, wie man diese Standards erreichen kann. Sie beschreiben deshalb neben einem Abschlussprofil für das Ende der Sekundarstufe I die Anforderungen, die am Ende der Klassen 6, 8 und 10 erreicht sein müssen, in Form von Kompetenzen.“ (MJSK, 2004, S. 4)

Ihre Funktion wird wie folgt beschrieben:

„Die Kernlehrpläne sind somit die Umsetzung der Bildungsstandards der KMK in Nordrhein-Westfalen auf der curricularen Ebene.“ (MJSK, 2004, S. 4)

Eine Vergleichbarkeit schulischer Standards soll nach den Vorstellungen des Landes NRW durch die gleichartige Struktur der Kernlehrpläne in den verschiedenen Fächern hergestellt werden:

„In Kapitel 1 der Kernlehrpläne wird aufgezeigt, welche *allgemeinen Bildungsziele* mit dem jeweiligen Fach verfolgt werden und welchen Stellenwert das jeweilige Fach im Gesamtzusammenhang schulischer Grundbildung hat. Diese allgemeinen Bildungsziele stecken den Rahmen ab für die Konkretisierung von fachbezogenen Kompetenzen.

In Kapitel 2 der Kernlehrpläne werden - abgestimmt mit den Standards der KMK - *verbindliche Anforderungen* am Ende der Sekundarstufe I definiert. Diese sind nicht nur fachlich relevant, sondern zugleich auch in besonderem Maße bedeutsam für den weiteren Bildungsweg der Schülerinnen und Schüler sowie für die Bewältigung ihres persönlichen und späteren beruflichen Alltags.“ (MJSK, 2004, S. 4)

Man mag darüber streiten, ob dies der Intention der KMK entspricht. Der nachfolgende Satz:

„Eine direkte Arbeit der Lehrkräfte mit den KMK-Standards ist nicht vorgesehen.“ (MJSK, 2004, S. 4)

weist jedoch darauf hin, dass es bei der nationalen Adaption der didaktischen Entwicklung auf Länderebene noch einer Vereinheitlichung im Gebrauch der Fachtermini sowie eines weiteren politischen Diskurses bedarf.

Ansätze einer gemeinsamen bildungspolitischen Normierung sind etwa erkennbar in der Ausgestaltung des Projektes VerA, bei dem das Land Nordrhein-Westfalen mit Berlin, Brandenburg, Bremen, Rheinland-Pfalz und Schleswig-Holstein kooperiert, ausgehend von einer Initiative in Rheinland-Pfalz.

Die neueren politischen Bestrebungen, dem Bund übergreifende Kompetenzen in der bisherigen Bildungshoheit der Länder zuzuordnen, sind bisher am Widerstand der Länder gescheitert (Stand Januar 2005).

1.8 Forschungsfragen

Ausgangspunkt für die Forschungsfragen dieser Arbeit war die Diskrepanz zwischen den TIMSS-Leistungen Schweizer und deutscher Schüler. Monokausale Erklärungen sind nach den in 1.6 dargelegten Gründen nicht zu erwarten.

Die Original-Schülerarbeiten waren zum Zeitpunkt des Entwurfs dieser Studie nicht unbeschränkt für wissenschaftliche Analysen freigegeben. Selbst bei frei zugänglichen Schülerarbeiten wäre jedoch der hohe Grad an MC-Aufgaben ein Forschungshindernis.

Sowohl zur Erforschung von Ursachen der graduellen Leistungsunterschiede zwischen Schülern verschiedener Nationen wie auch zum Verständnis von Denkprozessen bei Schülern und zur wissenschaftlichen Absicherung geplanter Änderungen von Unterrichtsstrukturen, Inhalten und Formen sind didaktische Detailanalysen elaborierter Schülerarbeiten unverzichtbar. Lediglich aus deskriptiven Befunden heraus abgeleitete Aussagen über zu initiierende Änderungen von Mathematikunterricht stützen sich auf statistische summarische Daten, bei denen nicht gesichert ist, ob eine wenig zufriedenstellende Leistung möglicherweise durch aufgabenimmanente, andere unterrichtliche oder sonstige außerunterrichtliche Faktoren bedingt ist.

Daher bedarf es gerade wegen des hohen Maßes politischer und unterrichtlicher Handlungsaktivitäten, die als Folge von TIMSS bereits eingeleitet wurden, noch umfangreicher Ursachenforschung, um didaktische und methodische Konsequenzen und Strukturänderungen von Unterricht wissenschaftlich abzusichern.

Die vorliegende Studie will einen Schritt in diese Richtung tun. Der allgemeine Fokus auf Bedingungsfaktoren für unterschiedliche Leistungen Schweizer und deutscher Schüler wurde daher verengt auf die Leitfrage, ob, wie und in wel-

chem Umfang aus elaborierten Aufgabenbearbeitungen Zusatzinformationen für zielgerichtete vergleichende didaktische Analysen extrahierbar sind.

Der Vergleich der Schülerleistungen Schweiz / Deutschland wird damit auch als ein Beispiel für zu entwickelnde Methoden zum didaktischen Detailvergleich gesehen – ohne Anspruch auf vollständige Erklärung.

Dazu bedarf es der Entwicklung eines Systems, welches den Vergleich der elaborierten Lösungswege ermöglicht und bedingende Faktoren sonstiger Genese extrahieren kann.

Voraussetzung für den Erfolg der Studie ist, dass die Schüler bereit und in der Lage sind, ihre Lösungswege explizit darzulegen und zu kommentieren. Da die TIMSS-Aufgaben eher auf eine MC-Abfrage angelegt sind, ist dies keineswegs sicher. Gelingt es also z.B. den Schülern, sogar bei einer Aufgabe wie K03 (nach TIMSS-Notation), wo Drehoperationen „im Kopf“ zu vollziehen sind, einen beschreibenden Text ihrer Lösungswege zu formulieren? Wie ergiebig sind die Beschreibungen? Bei nicht hinreichender Beschreibung wird eine vergleichende Detailanalyse deutlich erschwert.

Es sollen konkret folgende Forschungsfragen ggf. auch länderspezifisch differenziert beantwortet werden:

1.8.1 Zur Datengewinnung und Aufbereitung

1. Läßt sich ein möglichst signierobjektives Schlüsselsystem so entwickeln, dass vergleichende Aussagen zu Bearbeitungsqualitäten entnehmbar sind?
2. Welche Kodierungstiefen/-stufen sind den vorliegenden Aufgaben angemessen? Sind dynamische Schlüsselerweiterungen erforderlich?
3. Wie kann eine Kodierung der Lehrer- und Schülerbegleitquestionnaires so erfolgen, dass auch Dateninterpretationen zu Fragen wie einer Passung von Lehrer- und Schüleraussagen zur Curriculumvalidität, Übungsintensität etc. von Aufgaben möglich werden?

1.8.2 Zur Interpretation der Ergebnisse

1. Wird die Diskrepanz in den TIMSS-Schülerleistungen Schweiz zu Deutschland reproduziert?
2. Bleibt der relative Leistungsabstand bei einer fehlertoleranten Aufgabenbewertung erhalten oder erklären sich Leistungsunterschiede schon allein aus einer unterschiedlichen Anzahl leichter oder schwerer Rechenfehler?

3. Gibt es länderspezifische Unterschiede in Intensität und Qualität der Aufgabenbearbeitungen? Inwieweit wird insbesondere die These verifiziert oder falsifiziert, dass die Diskrepanz auf unvollständigen Aufgabenbearbeitungen oder auf einer Häufung von Nichtbearbeitungen beruht?
4. Gibt es länderspezifisch qualitative Unterschiede in den freien Kommentierungen der Questionnaires? Welche Aspekte sind isolierbar?
5. Welcher Anteil der Aufgaben wurde geraten? Gibt es hier länderspezifische Unterschiede? Gibt es möglicherweise statistisch signifikante Unterschiede zwischen Ratehäufigkeit und Rateerfolg, die auch auf einen Einsatz von ‚testwiseness‘-Strategien hindeuten könnten?
6. Gibt es länderspezifisch markante Häufungen von Fehlertypen und wodurch sind sie bedingt?
7. Gibt es aufgabenspezifische Probleme und sind sie möglicherweise curricular bedingt?
8. Sind die Aufgaben durchgängig curriculum valide oder gibt es bereichsspezifische Abweichungen?
9. Wie weit decken sich die Angaben zur Curriculumvalidität in Lehrer- und Schülerquestionnaires? Sind Korrelationen zwischen Validität und Bearbeitungsgüte erkennbar?
10. Sind Korrelationen zwischen Curriculumvalidität und Lösungshäufigkeit erkennbar?
11. Gibt es Hinweise auf die Maskierung von Schülerleistungen durch Sekundäreffekte, und sind didaktische Gründe dafür extrahierbar oder erkennbar?
12. Führt die exemplarische Analyse konkreter Aufgabenbearbeitungen in dieser Studie zu Modifizierungen gegenüber den in der Fachliteratur beschriebenen deskriptiven Befunden?
13. Welche generalisierenden Aussagen sind zum Vergleich der Leistungen der Schweizer und deutschen Schüler möglich?

1.9 Planung der Untersuchung

Den Forschungszielen entsprechend umfaßt die geplante Untersuchung folgende Grobaspekte:

- Auswahl der Aufgaben sowie Entwurf und Bereitstellung der Testhefte und der Begleitfragebögen („questionnaires“) unter Beachtung didaktischer Vorerlegungen,
- Probelauf,
- Durchführung der Tests in der Schweiz und in Deutschland,
- Bearbeitung und Kodierung aller Arbeiten anhand eines geeignet entwickelten aussagefähigen Schlüsselsystems,
- Bewertung der Schülerlösungen nach verschiedenen didaktischen Aspekten, mit Extraktion länderspezifischer Unterschiede.

Dieser Ablaufplan lässt sich unter Beachtung allgemein anerkannter Verfahrensschritte für den Ablauf einer Studie²² in Detailschritte auflösen, die dann schrittweise erarbeitet werden können. Dabei ist Prozesskohärenz erforderlich, um bei der Planung keine Detailschritte zu übersehen, die später möglicherweise zu Unterbrechungen der Arbeit führen. Dazu wurde ein zeitlich geordneter Ablaufplan der Detailschritte entwickelt. Zu jedem Schritt wurde überprüft, ob er auf der Basis der zeitlich voraus notierten Aspekte überhaupt realisierbar sei. So erzwingt beispielsweise der Prozessschritt „Durchführung“ die Frage nach dem Testmaterial etc. Als Kontrollinstrument diente eine vorab gedanklich zeitinverse Simulation des Verfahrensablaufs. Dies hat sich rückschauend bewährt.

Zur Realisierung und Auswertung der Studie lassen sich damit chronologisch geordnete Einzelschritte erarbeiten:

1. Kontaktaufnahme zu Schweizer Schulbehörden,
2. Auswahl geeigneter Klassen auch nach Vorgaben der Schweizer Schulbehörden,
3. Bewertung und Auswahl der freigegebenen TIMSS-Aufgaben nach Testbereich, Inhalt, Schwierigkeitsgrad und Anspruchshöhe sowie nach dem in der TIMSS-Studie ermittelten Schwierigkeitsgrad,
4. Neuübersetzung und teilweise sprachliche Adaption der freigegebenen TIMSS-Aufgaben,
5. Didaktische Überlegungen zu Umfang, Aufgabenauswahl und –verteilung auf mehrere Versionen von Testheften unter der Prämisse einer homogen Verteilung von Anforderungs- und mathematischen Teilbereichen,
6. Entwicklung von zugehörigen Fragebögen: Schülerbewertungsbogen, Lehrerbegleitbögen zur Erfassung der curricularen Validität der Aufgaben und Testverfahrensanweisungen sowie Schülerhinweistexte mit motivierenden Elementen für die Schüler,
7. Probelauf („pilot survey“) in Deutschland, daraus Entwurfsfehlerkorrekturen und Überarbeitung,
8. Produktion des Testmaterials,
9. weitgehend zeitsynchrone Durchführung der Tests in der Schweiz und in Deutschland in vergleichbaren Schulformen und Klassen,
10. Entwicklung eines aufgabenspezifischen Bewertungsschlüsselsystems, welches zugleich methodenqualifizierende Aussagen zuläßt und die Qualität der Einzelschritte der Schülerlösungen einer statistischen Analyse zugänglich macht,
11. schlüsselspezifische Bearbeitung und Kodierung aller Testhefte, dabei Übertragung der ermittelten Schlüsselwerte in elektronische Form,

²² Eine Basis dazu kann das von Cohen und Manion angegebene Flußdiagramm „Planning a survey“ (Cohen und Manion, 1998, S.84) sein, obgleich es hier nicht als absolute Planungsgrundlage diente.

12. Rohanalyse: Ermittlung von Übereinstimmungen mit oder Abweichungen von den TIMSS-Ergebnissen,
13. Feinanalyse (a): Untersuchung auf extramathematische Besonderheiten wie etwa Muster spezifischer Fehlerhäufungsbereiche, Bearbeitungsqualität oder auch –intensität etc.,
14. Feinanalyse (b): Detailvergleiche von Schülerlösungs wegen unter Einbeziehung der freien Schülerkommentare,
15. Feinanalyse (c): Vergleich der curricularen Bewertungen von Schülern und Lehrern,
16. Feinanalyse (d): didaktische Analyse eines Teilbereiches auf konkrete Schülerleistungen jenseits der TIMSS-Bewertung.

2 Entwurf der Studie

Das Forschungsziel ‚Wiederholung eines Teils des TIMSS-Tests Population 2 im Vergleich Schweiz zu Deutschland so, daß *Schülerlösungswege erkennbar und analysierbar werden*‘ erfordert für das geänderten Studiendesign Überlegungen zur Vergleichbarkeit, zu Probenumfang, Methoden der Datenerhebung, Testpopulation, Zeitaspekten, Kriterien und Auswahl einer TIMSS-analogen Aufgabenverteilung etc.

Es wurde großer Wert auf gleichartige Aufgabenverteilung und zeitgleiche Testbearbeitung in vergleichbaren Populationen unter vergleichbaren Bedingungen gelegt. Die folgenden Inhalte dieses zweiten Kapitels erläutern die entsprechenden ‚technischen‘ Überlegungen und Entscheidungen.

2.1 Probenumfang

Der erforderliche Probenumfang ist abhängig von der geplanten Verwendung des erhobenen Datenmaterials. Cohen/Manion geben eine Zahl von etwa 30 Exemplaren pro Aufgabe als Minimum an, falls statistisch signifikante Analysen des Datenmaterials geplant sind (Cohen und Manion, 1998, S.89).

Von den 151 TIMSS-Aufgaben der Population 2 unterliegen etwa 1/3 der Geheimhaltung für mögliche weitere Verwendung. 102 Aufgaben wurden veröffentlicht (cluster I bis Z nach TIMSS-Nummerierung).

Unter Zugrundelegung der o.a. Mindeststichprobengröße und Verwendung aller freigegebenen TIMSS-Aufgaben ergäbe sich eine Gesamtzahl von 102x30 oder 3060 *pro Land*, also insgesamt 6120 Aufgaben. Damit liegt eine Datenmenge vor, deren elaborierte Auswertung kaum noch von einer einzelnen Person zu bearbeiten ist, wenn diese Bearbeitung nicht nur Kontrolle auf richtig oder falsch sein soll, sondern auch didaktisch auswertend, klassifizierend und bewertend.

Die geplante Aufgabenzahl von 12 Aufgaben pro Schüler (zur Diskussion dieser Setzung s.u.) fordert unter den vorstehenden Bedingungen eine Schülerzahl von $3060:12 = 255$ entsprechend etwa 10 Klassen in Deutschland (bei einer angenommenen Klassenfrequenz von 25 Schülern pro Klasse) und 17 Klassen in der Schweiz (bei 15 Schülern pro Klasse). Vermittelt wurden aber von der Schweizer Schulverwaltung vier Klassen mit durchschnittlich 15 Schülern. Die Aufgaben- und damit die Klassenzahl musste dieser Prämisse angepasst werden.

Als realistische und für einen einzelnen Auswerter in angemessener Zeit noch zu bewältigende *Gesamtaufgabenzahl* der Studie wurde maximal 3200 angesehen.

Somit kann die Arbeit keine statistisch signifikanten Ergebnisse liefern, bietet aber folgende Vorteile:

„.... they [Kleinstudien, „small scale surveys“, Anm. des Verf.] are far less complicated to set up, are considerably less expensive, and can prove perfectly adequate when researchers do not intend to generalize their findings beyond the sample in question...“ (Cohen und Manion, 1998)

Letztlich wurden fünf Testheft-Typen (bezeichnet mit A bis E) mit je 12 verschiedenen TIMSS-Aufgaben entworfen. Somit wurden 60% der freigegebenen Aufgaben eingebunden in insgesamt 234 Testhefte, entsprechend 2808 zu analysierenden TIMSS-Aufgaben.

2.2 Datenerhebung

Da die Studie Schweizer und deutsche Schüler betrifft, bedurfte die Art der Datenerhebung („sample collecting“) besonderer Vorüberlegungen. Es war von Anfang an das Anliegen des Verfassers, das Testverfahren in beiden Ländern vor Ort zu begleiten, um möglichst Gleichartigkeit der Testumstände zu sichern. Eine solche persönliche Teilnahme versprach weitere Zusatz- und Sekundärinformationen durch:

- Besuch von Schweizer Schulhäusern, dabei
- Kennenlernen von Schulgebäuden, ihrer Architektur, Ausstattung und Atmosphäre,
- Gespräche mit den Schulleitern,
- Gespräche mit den Kollegen, in deren Klassen die Tests durchgeführt wurden, über Unterricht im allgemeinen, über die Bedeutung der Fragebögen und über die Curriculumvalidität der Testaufgaben bzw. über klassen- und hausinterne Abweichungen vom Standardcurriculum,
- Motivierung zur Ausfüllung der umfangreichen Lehrerbegleitbögen,
- Kennenlernen der Klassenzimmer und ihrer Einrichtung
- Kennenlernen des Schülerverhaltens,
- gleichartige Testeinführung und damit möglichst gleiche Testmotivation in allen Klassen²³,
- Beobachtung des Schülerverhaltens während der Tests,
- Zurückhaltung der Fachlehrer (sie neigten – wie vorab vermutet – in unterschiedlich starkem Maße dazu, Hilfen zu geben),
- Verhinderung gegenseitiger Schülerhilfen.

²³ Dabei wirkt natürlich die Anwesenheit eines für die Klasse fremden deutschen Lehrers in der Schweiz anders als in Deutschland – allein schon sprachlich bedingt.

Zu wesentlichen Beobachtungsaspekten wurden vom Verfasser stichwortartig Protokolle geführt. Eine Diskussion von Facetten dieser Protokolle folgt unter 3.3 und 3.4.

Da eine persönliche Teilnahme weder technisch noch zeitlich gesichert war, mußte alternativ ein Versand der Fragebögen per Post („postal questionnaires“) eingeplant werden. Dies erforderte eine auch auf diesen Fall abgestellte Strukturierung der Lehrerinformationsblätter, um bei Abwesenheit des Verfassers möglichst gleichartige Testabläufe zu sichern.

2.3 Testpopulation

Die Alterscharakteristik der Schüler in TIMSS Population 2 wird folgendermaßen beschrieben:

„All countries that participated in TIMSS were to test students in the two grades with the largest proportion of 13-year-olds in both mathematics and science.“ (online, TIMSS 1998a, S.8)

Dementsprechend bot es sich an, Schüler der Klasse 7 am Ende des Schuljahres zu testen, weil so eine Altersstruktur mit einem Schwergewicht bei etwa 13 Jahren zu erwarten ist. Zum Vergleich mit TIMSS wurde trotz zugesicherter Anonymität der Tests entweder das Geburtsdatum oder das Alter in Jahren und Monaten erhoben²⁴.

Da der Verfasser langjährig Mathematik am Gymnasium unterrichtet hat und seine eigene Berufserfahrung primär aus dem Gymnasialbereich z.B. bei der Beurteilung von Curriculumvaliditäten oder Anforderungsniveaus von Aufgaben in die Probenauswahl einbringen wollte, wurden für Deutschland als Zielgruppe Schüler aus Klassen 7 von zwei *Gymnasien* ausgewählt.

Nach dieser Prämissen wählte Herr Prof. Dr. Armin Hollenstein²⁵, Vizedirektor des Sekundarlehramts der Universität Bern, als Vermittler zu Berner Schulen unter Beachtung weiterer Randbedingungen wie Bereitschaft der Schulleiter und Kollegen, Belastung der Schulen etc. zwei ihm als vergleichbar und geeignet erscheinende Schulen, die er wie folgt charakterisierte:

²⁴ Somit wäre den Fachlehrern theoretisch bei Angabe des Geburtsdatums eine Re-Identifikation der Schüler möglich, daher werden keine korrelierten Angaben zu Geburtsdatum und Klasse/Testheft-Nummer publiziert.

²⁵ email über <http://www.sla.unibe.ch/personen/hollenstein.html>

Tabelle 1: Am Test partizipierende Schweizer Schulen

7KA CH	Schule 1 (öffentliche Schule) Klasse Nr. 1	Sekundarklasse, Gruppe, in der Schüler mit guten Mathematik-Leistungen zusammengezogen werden (sog. "Spezialklasse")
7KE CH	Schule 1 Klasse Nr. 2	Sekundarklasse, Mittelfeld-Gruppe (14 Schüler)
7G1 CH	Schule 2 (privates Gymnasium, volle staatliche Anerkennung, Hausmatur eidgenössisch anerkannt), Klasse Nr. 1	7. Klasse, unterrichtet nach dem allgemein gültigen Sekundarlehrplan. Ein eigentliches Unter- oder Progymnasium existiert in Bern nicht.
7G2 CH	Schule 2 Klasse Nr. 2	7. Klasse (eine Parallelklasse zu 7G1 CH)

Die beiden Klassen der Schule 2 wurden vom gleichen Lehrer unterrichtet. Er schätzte beide Klassen als recht leistungsstark ein, Klasse zwei noch etwas stärker als Klasse eins.

In beiden Klassen wurde die Geometrie bisher wenig behandelt, da der hauseigene Stoffverteilungsplan nach persönlicher Mitteilung des Fachlehrers für die Geometrie einen Schwerpunkt in der Klasse 8 setzt.

In Deutschland wurde der Test mit vier siebten Klassen von zwei Gymnasien sowie zusätzlich mit einer achten Klasse durchgeführt. Durch die Einbeziehung der achten Klasse können Lernfortschritte von der siebten zur achten Klasse untersucht werden.

Die Testdurchführungen in NRW geschahen unter Beachtung der datenschutzrechtlichen Vorschriften des Erlasses zur „Durchführung empirischer Untersuchungen und Befragungen in Schulen“ (BASS 1999/2000, 10-45 Nr. 2). Die Schüler gaben gelegentlich trotz Hinweis auf die Anonymität der Tests ihre Namen auf den Testheften an. Diese Namen werden aber weder verwendet noch sonst in irgendeiner Form veröffentlicht. Die entsprechende Spalte in der erhobenen Gesamtdatentabelle wurde gelöscht. Die Identifikation einzelner Testhefte erfolgt lediglich durch fortlaufende Kennziffern.

2.3.1 Vereinbarungen zur Nomenklatur

Um eine problemlose Vergleichbarkeit mit den TIMSS-Daten und anderen Publikationen zu sichern, wurden die aus dem angelsächsischen Sprachraum stammenden Abkürzungen übernommen:

Wie bereits erwähnt sind die TIMSS-Aufgaben nach *sechs inhaltlichen Dimensionen* strukturiert. Es bedeuten:

- n Zahlen und Zahlenverständnis (‘fraction and number sense’)
- a Algebra (‘algebra’)
- m Messen und Maßeinheiten (‘measurement’)
- g Geometrie (‘geometry’)
- d Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung (‘data representation, analysis, and probability’)
- p Proportionalität (‘proportionality’).

Die „performance expectation“ lässt sich als „Schwierigkeitsgrad“ oder „Anspruchsniveau“ interpretieren und ist hier nach Ziffern aufsteigend geordnet:

- 1 Wissen (‘knowing’)
- 2 Beherrschung von Routineverfahren (‘performing routine procedures’)
- 3 Beherrschung von komplexen Verfahren (‘using complex procedures’)
- 4 Anwendungsbezogene mathematische Probleme (‘solving problems’²⁶).

Weiter werden die Abkürzungen zum *Aufgabentyp* übernommen. Es bedeuten:

- m reine MC-Aufgaben (‘multiple choice items’)
- s Kurzlösungen (‘short answer items’)
- e elaborierte Aufgabenlösungen (‘extended answer items’).

Zu Vergleichszwecken werden für alle Aufgaben ihre TIMSS-Namen wie I01, K08, R14 etc. beibehalten.

Die hier vereinbarten Abkürzungen werden in den folgenden Tabellen aus Platzgründen ohne erneute Erläuterung verwendet. Aus den TIMSS-Publikationen übernommene Tabellen werden unter Beibehaltung dieser Abkürzungen ins Deutsche übertragen.

2.3.2 TIMSS-Aufgabenverteilung

Die TIMSS-Aufgaben wurden nach Inhaltsbereichen und Schwierigkeitsstufen diskriminiert und deklariert. Ihre Gesamtverteilung nach Inhaltsbereichen und Schwierigkeitsstufen war:

²⁶ Man beachte dazu die Ausführungen in Kapitel 1.6.1 zur bekannten problematischen Übersetzung des Terminus ‘solving problems’). Die Übersetzung wurden entnommen aus (Bauernert et al., 1997, S.48)

Tabelle 2: TIMSS-Aufgabenverteilung nach inhaltlichen Bereichen

Bereich	Gesamtzahl Aufgaben	davon multiple-choice-Aufgaben	davon Aufgaben mit Kurzlösung	davon elaboreierte Aufgaben
n	51 (37)	41 (27)	9 (9)	2 (2)
a	27 (18)	22 (13)	3 (3)	2 (2)
m	18 (12)	13 (7)	3 (3)	2 (2)
g	23 (17)	22 (16)	1 (1)	0 (0)
d	21 (12)	19 (10)	1 (1)	1 (1)
p	11 (6)	8 (3)	2 (2)	1 (1)
Summe	151 (102)	125 (76)	19 (19)	7 (7)
Die Zahlen in Klammern beziehen sich auf freigegebenen Aufgaben. (nach TIMSS 1994-95, S.iv)				

Tabelle 3: TIMSS-Aufgabenverteilung nach Anspruchsbereichen

Bereich	Gesamtzahl Aufgaben	davon multiple-choice-Aufgaben	davon Aufgaben mit Kurzlösung	davon elaboreierte Aufgaben
1	33 (16)	31 (13)	2 (2)	0 (0)
2	38 (30)	32 (24)	6 (6)	0 (0)
3	32 (19)	28 (15)	4 (4)	0 (0)
4	48 (38)	34 (24)	7 (7)	7 (0)
Die Zahlen in Klammern beziehen sich auf die freigegebenen Aufgaben. (nach TIMSS 1994-95, S.iv)				

Eine Liste der freigegebenen Aufgaben mit zugehörigen Aufgabentiteln findet sich in (online,TIMSS 1998c). Sie werden dort exakt so wiedergegeben, wie sie den teilnehmenden Ländern übergeben wurden, also noch vor Übersetzung und evtl. nationaler Adaption durch die nationalen Zentren.

2.3.3 Kriterien zur Aufgabenauswahl

Alle freigegebenen Aufgaben wurden daraufhin analysiert und es wurde tabellarisch erfasst, welche mathematischen Mikrobereiche jeweils angesprochen wurden und welche Fähigkeiten oder Fertigkeiten zur Bewältigung einzusetzen waren. Diese Analyse diente bei der späteren Auswertung als ein Hilfsmittel zur Analyse der Schülerlösungswege. Es folgt ein beispielhafter Auszug aus dieser didaktischen Aufgabenbewertung. Die Analyse war weiter eines der Hilfsmittel zur Planung einer homogenen Aufgabenverteilung auf die Testhefte. Die Spalte „diffic(ulty) index“ gibt den aus den TIMSS-Ergebnissen international ermittelten Gesamtschwierigkeitsgrad (international difficulty index‘) der jeweiligen Aufgaben an (mit einer Spanne zwischen 326 (leicht) und 699 (sehr schwer) sowie zwei Spitzenwerten von 737 und 817).

Tabelle 4: Beispiele mathematischer Mikrobereiche in TIMSS-Aufgaben und zur Lösung erforderliche(s) Wissen/Fähigkeiten/Fertigkeiten

Aufgabe	Art	diffic. index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
I 01	a	628	Bedeutung einer Variablen kennen	Sachverhalte ‚rückwärts‘ mathematisieren
I 02	n	530	Bruchanteile im Kopf und schwierige Formulierung	Prozentanteilgefühl; Aufgaben präzise lesen
I 04	a	591	zwei Folgen weiterführen: ausprobieren oder kgV	Zahlenfolgen weiterschreiben oder kgV anwenden
I 06	n	427	Brüche ordnen / vergleichen	z.B. Regel kennen, dass Vergrößerung nur des Zählers den Wert eines Bruches erhöht
...				
O 01	d	535	Ablesen im gekrümmten Graphen	Aufnehmen eines komplexen sprachlichen Zusammenhangs, Kenntnis von zweiachsigen Zustandsdiagrammen, umgekehrte Ablesung $y \rightarrow x$
O 02	n	680	prozentuale Steigerung errechnen aus absoluten Zahlen	Differenz als Aufschlag auf Grundwert erkennen und vom Zahlenanteil auf Prozentanteil umrechnen, Vermeidung der Falle Aufschlag = Prozentzahl
O 05	d	587	Anzahlen aus Wahrscheinlichkeiten	Verknüpfung Häufigkeit / Wahrscheinlichkeit ‚rückwärts‘ durch geschicktes Erweitern der vorgegebenen Häufigkeit
O 06	m	465	Rechnen mit Zeitangaben	Rechnen mit Bruchanteilen einer Stunde und Addieren im 60er-System
...				

Im TIMSS-Test waren – wie bereits in 1.6.4 erwähnt – 70 Aufgaben in 90 Minuten zu bearbeiten. Somit standen pro Aufgabe nur ca. 1,3 Minuten zur Verfügung. Dies impliziert zwei Bearbeitungsaspekte. Zum einen ist denkbar, dass testerfahrene Schüler neben einer mathematischen Bearbeitung der Aufgaben auch Testlösestrategien („testwiseness“-Strategien) einsetzen oder sich teilweise auf das Erraten von Lösungen beschränken. Zum anderen erforderte die Testkonstellation einen sehr langen Zeitraum intensiver Konzentration auch von Schülern, die eher an einen 45-Minuten-Unterrichtstakt gewöhnt sind.

Nach den Forschungsfragen dieser Studie sollten beide Aspekte möglichst in den Hintergrund gedrängt werden, um eingesetzte mathematische Fähigkeiten oder auch typische Fehlerquellen aufdecken zu können. Daher wurden die zeitlichen Beschränkungen aufgehoben, um zu Aufgabenbearbeitungen zu gelangen, die auch bei ausführlicher Darstellung der Lösungswege nicht durch Zeitdruck erzwungen wurden.

Die Anzahl der Aufgaben wurde daher in Abstimmung mit der oben kalkulierten Maximalzahl von Aufgabenbearbeitungen auf 12 pro Testheft, also in der Regel auch pro Schüler²⁷, festgelegt. Damit standen bei einer angesetzten Testzeit von 90 Minuten selbst bei einer eingerechneten Anlaufzeit incl. Erläuterungs-, Motivations- und Austeilphasen von 6 Minuten und einer individuellen Pausen- oder Entspannungszeit von weiteren 6 Minuten immer noch 6,5 Minuten reine Bearbeitungszeit pro Aufgabe zur Verfügung.

Die Aufgaben sollten auf fünf Testhefte verteilt werden, um ein Abschreiben zu minimieren. Keine Aufgabe kommt in mehr als einem Testheft vor. Daher erreicht die Studie wie bereits erwähnt eine fast 60%ige Abdeckung der freigegebenen TIMSS-Aufgaben. Dies erschien sinnvoller als die Einarbeitung von Verankerungs-Elementen („anchor items“). Diese hätten ohnehin für jede Inhaltsdimension konstruiert werden müssen. Dazu ist aber der Probenumfang zu gering.

Bei der Auswahl aus den freigegebenen Aufgaben wurden folgende Aspekte bedacht:

- Zum Zwecke der Vergleichbarkeit mit den TIMSS-Ergebnissen war es geboten, das prozentuale Verhältnis der Sachbereiche des TIMSS-Tests in den Testheften zu reproduzieren.
- Ebenfalls aus Gründen der Vergleichbarkeit sollte die Verteilung „leichter“ und „schwierigerer“ Aufgaben gleichmäßig sein. Als Maßstab für die Qualität „Schwierigkeitsgrad“ bietet TIMSS den bereits zu Tabelle 4 erläuterten „international difficulty index“ an. Die Verteilung der Aufgaben auf die Testhefte wurde auch nach diesem Index weitgehend homogen gehalten, so dass nach diesem Maßstab leichte und schwere Aufgaben etwa gleichmäßig berücksichtigt wurden.
- Die Diskriminierung der Aufgaben in leichte und schwere nach dem international difficulty index entsprach nicht immer dem, wie der Verfasser sie nach seiner Unterrichtserfahrung eingeordnet hätte. Alle Aufgaben wurden nach diesem rein subjektiven Empfinden vorab in eine der Schwierigkeitsklassen eins (leicht) bis vier (schwer) eingeordnet. Nachdem die Aufgaben anhand der ersten Kriterien ausgewählt waren, bot es sich an, so weit wie möglich auch noch eine gleichmäßige Verteilung dieses subjektiven Bewertungsschlüssels zu erreichen. In jedem Fall hatten aber die erstgenannten Kriterien Vorrang bei der Aufteilung, so dass diese nachgeordnete subjektive Einordnung in Schwierigkeitsstufen keinerlei implizierenden Charakter hatte.

²⁷ Allerdings waren viele Schweizer Schüler sehr früh mit der Bearbeitung ihres Testheftes fertig und verlangten ein zweites Heft.

Es bleibt anzumerken, dass ein solches aus subjektiver Unterrichtserfahrung gespeistes Einordnen von Aufgaben und Aufgabenteilen ohne wissenschaftlich strenge Absicherung zu den dienstlichen Aufgaben jedes Mathematiklehrers in Nordrhein-Westfalen gehört, wenn er seine Abiturvorschläge kategorial klassifiziert: bei der Vorlage der Abituraufgaben bei der Dienstaufsichtsbehörde ist jeder Aufgabenteil nach einem dreistufigen Schwierigkeitsraster zu gewichten. Die Richtlinien des Landes NRW sehen drei Stufen von Kategorie 1 (rein reproduktive Aufgaben) bis zu Kategorie 3 (Transferleistungen) vor. Diese Bereiche werden in den Richtlinien Mathematik wie folgt klassifiziert (und dort auch detaillierter beschrieben):

Bereich 1: Wiedergabe von Sachverhalten, Beschreibung und Verwendung gelernter und geübter Verfahrenstechniken,

Bereich 2: selbstständiges Auswählen, Anordnen, Bearbeiten und Verknüpfen bekannter Sachverhalte,

Bereich 3: planmäßiges Verarbeiten komplexer Gegebenheiten mit dem Ziel, zu selbstständigen Lösungen, Gestaltungen oder Deutungen, Folgerungen, Begründungen, Wertungen zu gelangen. ... (Richtlinien Mathematik S II, 1999, S.71 ff).

Die Richtlinien bieten allerdings keinerlei Kriterien für die adäquate Klassifizierung der Aufgaben und auch der Teilaufgaben. Da die Zuordnung speziell zu den Stufen 2 und 3 wesentlich vom unterrichtlichen Kontext bedingt wird, ist dies auch nicht weiter verwunderlich. Unter 10.9 finden sich weitere Ausführungen zum Vergleich dieses dreistufigen Kategoriensystems mit dem dreistufigen Kompetenzklassenmodell.

2.4 Nationale Adaption

Dieser Abschnitt befasst sich mit Überlegungen zur Übersetzung der TIMSS-Aufgaben und Problemen, die bei der nationalen Adaption durch spezielle Aufgabenformulierungen entstehen können. Unter TIMSS waren dafür die nationalen Testzentren zuständig. In Schmidt, W.H. et al, 1996b, S. 209 ff, findet sich die zugehörige Liste der nationalen Koordinatoren.

Für die vorliegende Studie wurden die ausgewählten Aufgaben neu übersetzt. Zur Berücksichtigung nationaler Gegebenheiten wie länderspezifischer Notationen oder auch Einheiten boten sich gelegentlich leichte Anpassungen an. Bei jedem einzelnen Aufgabentext wurde versucht, mögliche vermutete Verschiedenartigkeiten zwischen deutschem und Schweizer Verständnis zu eliminieren, wie etwa bei der Ersetzung des internationalen Neutralnamens Alana durch Peter oder durch das ausdrückliche Erwähnen von Franken und DM statt der Verwendung von ‚zeds‘ als Währungseinheit.

Wie sich später im Probelauf in Deutschland (vgl. Abschnitt 2.8) erwies, waren diese Anpassungen nicht immer trivial oder von untergeordneter Bedeutung, sondern hatten punktuell erheblichen Einfluss auf das Verständnis der Aufgabe und damit auch auf die Lösungshäufigkeiten.

Als Beispiel für die Bedeutung der Aufgabenpräsentation sei im Vorgriff die Aufgabe Q07 vorgestellt. Sie lautet im Original:

Q07. *$P = LW$. If $P = 12$ and $L = 3$, then W is equal to*

- A. $\frac{3}{4}$ B. 3 C. 4 D. 12 E 36

Sie wurde bewusst wörtlich übernommen als

Q07. *$P = LW$. Falls $P = 12$ und $L = 3$ ist, dann hat W den Wert*

- A. $\frac{3}{4}$ B. 3 C. 4 D. 12 E 36

Wie den Schüler- und Lehrerkommentaren zu entnehmen ist, hatte diese Aufgabenpräsentation deutlichen Einfluß auf die Lösungshäufigkeit, da sie in dieser Formulierung offensichtlich nicht verstanden wurde. Dies dürfte auch der Grund für den internationalen Schwierigkeitsgrad von immerhin 519 sein. Eine Präsentation der Aufgabe als „Womit muss ich 3 multiplizieren, um 12 zu erhalten?“ dürfte dagegen zu einer deutlich höheren Lösungshäufigkeit führen. Die Aufgabenproblematik wird hier also im wesentlichen durch die Aufgabendarbietung begründet.

In der Aufgabe R09 wurde bei der Übersetzung aus der englischen Aufgabenstellung

R.09 „Which one of the following is FALSE when a , b , and c are different real numbers“

- | | | |
|----------------------------|--------------------|--------------------|
| A. $(a+b) + c = a + (b+c)$ | B. $ab = ba$ | C. $a + b = b + a$ |
| D. $(ab)c = a(bc)$ | E. $a - b = b - a$ | |

bewußt das Wort „reelle“ Zahl verwendet, obwohl die Aufgabe auch ohne den Fachbegriff erfassbar gewesen wäre. Da der Begriff „reelle“ Zahl in der Klasse 7 noch nicht begrifflich eingeführt sein dürfte, lässt sich hier fragen, ob dies im Lehrerquestionnaire gerügt wird und ob möglicherweise dadurch ein merkbarer Einfluß auf die Bearbeitungs- bzw. Lösungshäufigkeit ausgeübt wird. Denkbar ist auch, daß das unbekannte Wort einfach überlesen und die Aufgabe dennoch bearbeitet wird. Einige Schüler wiesen explizit auf Probleme mit dem Terminus „reell“ hin.

Ähnliche Effekte lassen sich bei Aufgabe K04 erwarten, wo das englische „ $x/2 < 7$ is equivalent to“ übernommen wurde als „äquivalent zu“. So gewinnt das gän-

gige englische Wort der gehobenen Umgangssprache für Gleichwertigkeit den Charakter eines mathematischen Fachwortes, und es ist denkbar, dass dadurch die Aufgabenbearbeitung gehemmt wird. Eine Übersetzung mit „...bedeutet das Gleiche wie...“ hätte vielleicht diese denkbare Zusatzhemmung vermieden.

Eine anders gelagerte Problematik entsteht, wenn Fachbegriffe tatsächlich unbekannt sein sollten wie z.B. „Schwerlinien“, „Kongruenz“ etc. Dabei handelt es sich dann allerdings nicht mehr um die Problematik der sprachlichen Präsentation, sondern um Probleme der Curriculumvalidität.

Bei einer Reihe von Aufgaben mussten graphische Elemente eingefügt werden. Hier wurden in der Regel die TIMSS-Zeichnungen in geeigneter Skalierung übernommen. An wenigen Stellen wurden aus didaktischen Gründen Skalenverhältnisse bewußt verändert, um Lösungen durch Ausmessen unmöglich zu machen.

2.5 Curriculumkonforme Zusatzaufgaben

Neben den TIMSS-Aufgaben wurden für ein weiteres Forschungsvorhaben algorithmisch orientierte Aufgaben eingestreut. Da die Bearbeitung aller Aufgaben dieses Testheftes F deutlich mehr Zeit erforderte als bei den anderen Testheften, wurde dafür Sorge getragen, dass durch die unterschiedlichen Bearbeitungszeiten möglichst keine Frustrationseffekte entstanden. Als Weg bot sich die mündliche Information vor Testbeginn an. Die Schüler wurden darüber informiert, dass ein Teil der Testhefte deutlich mehr Bearbeitungszeit erfordere als die übrigen; sie erfuhren allerdings nicht, um welches Testheft es sich dabei handele. Somit konnten sich auch andere Schüler mit Zeitproblemen – falls sie denn auftaten – darauf berufen, dass dies nicht ungewöhnlich sei.

Diese zusätzlichen Testhefte wurden für die vorliegende Studie nicht weiter verwendet, da ihnen kein TIMSS-Analogon entspricht. Sie gehen aber in die Festlegung des Probenumfangs mit ein.

2.6 Testhefte

2.6.1 Didaktische Funktion

Form und Präsentation von Testblättern oder Testheften können didaktische Funktionen übernehmen. So kann allein durch ein ausreichendes oder ein knappes Platzangebot ein Impuls zu einer ausführlichen oder weniger elaborierten Aufgabenbearbeitung entstehen. Die Integration eines Schülerfragebogens in jede einzelne Aufgabenbearbeitung kann über Gewöhnungseffekte zu einer höheren

Ausfüllquote führen. Daher wurden folgende augenscheinlich zweckmäßigen Anforderungen an das Design der Testblätter gestellt:

- ausreichend Platz zum Schreiben pro Aufgabe (je ein Doppelblatt DIN A 4), ohne bei Bearbeitung das Doppelblatt wenden zu müssen,
- Ergebnis- und weitere Abfrageraster in immer gleicher Form an gleicher Stelle, um so möglicherweise die Ausfüllhäufigkeit zu erhöhen (Gewöhnungseffekt),
- Schülerdaten fest mit dem Fragebogen verbunden, aber nur einmal auszufüllen,
- leichte Katalogisierbarkeit und Möglichkeit der Zuordnung im Auswertungsprozeß,
- vorgeschaltetes Schülermotivationsblatt.

Diesen Forderungen entsprachen die entworfenen Aufgabenhefte durch folgenden Aufbau:

- Deckblatt mit Informations- und Motivationstexten,
- dort auch Raum für Schülerdaten und Auswertungsvermerke,
- jede Aufgabe auf einem Doppelblatt A4, links oben die Aufgabe, rechts unten die Fragen zur Aufgabe,
- Rückseite zwangsläufig leer,
- ausreichender Heftrand zur Durchlochung und späteren Archivierung.

2.6.2 Homogene Aufgabenverteilung

Die nach den in 2.4.3 dargelegten Kriterien ausgewählten TIMSS-Aufgaben wurden unter den bereits diskutierten Aspekten möglichst homogener Verteilung auf die fünf AufgabenTesthefte A, B, C, D und E aufgeteilt. Die Aufgaben sind in den Testheften gekennzeichnet durch ihre TIMSS-Nummer. Zusätzlich sind sie von 1 bis 12 durchnummieriert. Die nachfolgende Tabelle zeigt die Verteilung der Aufgaben auf die Testhefte.

Tabelle 5: Verteilung der Inhaltsbereiche und TIMSS-Aufgaben auf die Testhefte A bis E

	Nr.	A	B	C	D	E
n	1	K01	Q09	K02	I06	I05
a	2	P10	N13	I04	L13	Q02
g	3	K03	R10	L15	J16	M05
n	4	K09	N16	Q06	O09	J12
p	5	V03	T2ab	R14	M06	L14
a	6	J18	R09	Q01	Q07	R11
d	7	V02	I09	R08	N18	O05

	Nr.	A	B	C	D	E
n	8	J14*	R07	M04	M08	N17
m	9	Q03	I03	U2ab	N15	V04
g	10	M07	P09	O03	K08	I08
n	11	O02	Q08	R13	K06	P16
a	12	L16	L11	K04	T1ab	I01

*: die Aufgabe J14 wurde gegenüber der TIMSS-Aufgabe leicht modifiziert

Der bewußte Wechsel der Inhaltsbereiche in der Aufgabenabfolge dient nicht nur als Element der Abwechslung. Bei Gruppierung nach Sachbereichen bestünde die Gefahr, dass sich bei (nicht geforderter, aber denkbarer und auch wahrscheinlicher) linearer Bearbeitung der Testhefte Ermüdungserscheinungen auf einen Inhaltsbereich konzentrieren und somit möglicherweise fälschlich dem Aufgabentyp zugeschrieben würden.

Die folgende Tabelle zeigt die Verteilung nach dem internationalem Schwierigkeitsgrad aus TIMSS.

Tabelle 6: Schwierigkeitsgrad der Aufgaben

n	1	K01	464	Q09	558	K02	487	I06	427	I05	448
a	2	P10	540	N13	576	I04	591	L13	326	Q02	568
g	3	K03	478	R10	551	L15	623	J16	548	M05	565
n	4	K09	563	N16	580	Q06	610	O09	571	J12	593
p	5	V03	603	T2ab	699 815	R14	617	M06	634	L14	693
a	6	J18	594	R09	603	Q01	595	Q07	519	R11	584
d	7	V02	675	I09	490	R08	565	N18	541	O05	587
n	8	J14*	585	R07	583	M04	615	M08	575	N17	611
m	9	Q03	636	I03	603	U2ab	621 737	N15	492	V04	610
g	10	M07	457	P09	617	O03	581	K08	639	I08	597
n	11	O02	680	Q08	587	R13	642	K06	604	P16	637
a	12	L16	615	L11	640	K04	606	T1ab	627 631	I01	628

2.6.3 Schülerinformationsblatt

In der TIMSS-Folgediskussion wurde der Verdacht geäußert, dass die Schüler möglicherweise nicht genügend motiviert waren, den Test sorgfältig zu bearbeiten (vgl. hierzu auch die Verfahrenskritik unter 1.6.4). Um einer solchen abweisenden Arbeitshaltung entgegenzuwirken, sollte das Deckblatt der Testhefte als Schülerinformationsblatt mit einem eingearbeitetem Motivationalelement gestaltet werden. Es boten sich folgende Design- und Gestaltungsfaktoren an:

- auflockerndes Grafikelement (Bild „von Formeln rauchender Kopf“)
- Hinweis auf den bereits durchgeführten Test
- Vorstellung des aktuellen Tests als *Wettbewerb* Schweiz/Deutschland
- Hinweis auf mehr als ausreichende Bearbeitungszeit
- Hinweis auf geplante Rückmeldung, wer „*Sieger*“ wurde
- Möglichkeit, sich einen Phantasienamen geben zu können

Das Grafikelement sollte an die heute den Schülern in allen Fächern geläufige Form der Auflockerung von Textmaterial anknüpfen. Der Hinweis auf den bereits durchgeführten Test und die Vorstellung des aktuellen Tests als Wettbewerb zusammen mit dem Versprechen einer Rückmeldung appellierte an den sportlichen Ehrgeiz.

Dieses Versprechen der Rückmeldung wurde eingehalten. Zusätzlich erhielten die betreuenden Lehrer unmittelbar nach Abschluss der Tests auf Wunsch alle verwendeten Testhefte als Muster zur Besprechung der Aufgaben mit den Schülern. Dieser behutsame Umgang mit der „Ressource“ Schüler und Lehrer geschah auch mit dem Ziel, keine negative Haltung gegenüber mathematikdidaktischen Tests aufkommen zu lassen. Fehlende Informationen und Rückkopplungen sprechen sich herum und fördern sicher nicht die Bereitschaft zur weiteren Testteilnahme an zukünftigen Untersuchungen²⁸.

Der Hinweis auf die ausreichende Bearbeitungszeit sollte Ängste nehmen, die möglicherweise durch Erinnerung an Klassenarbeiten mit knapp bemessener Zeit entstehen könnten. Schließlich diente der sicher anspornende Hinweis, dass aus der ganz eigenen Schülerleistung Informationen zur Verbesserung von Mathematikunterricht kommen könnten, als denkbares Motivationalelement zur sorgfältigen Arbeit.

²⁸ Daher war diese Rückmeldung auch ausdrücklicher Wunsch von Herrn Prof. Dr. Hollenstein, um die Bereitschaft der Lehrer zu weiteren Testteilnahmen nicht zu gefährden. Das Fehlen von Rückmeldungen wurde auch von einem der beiden Schulleiter als ärgerliches Faktum angesprochen und als Grund angegeben, warum er selbst mit *seiner Klasse diesmal* nicht teilnehme.

Eine entsprechende Wirkung kann jedoch nur dann erwartet werden, wenn den Schülern auch Zeit zur Lektüre des Informationsblattes gegeben wird. Diese nur vordergründig triviale Bedingung muß bei der Testdurchführung sorgfältig beachtet werden. Andernfalls besteht die Gefahr, dass der Testleiter aus seiner Vertrautheit mit dem Test heraus dazu übergeht, die Zeit zur Lektüre des Informationsblattes zu kürzen. Dies wäre kontraproduktiv zum Motivationsziel.

Zusätzlich wurde Platz zur Erfassung der benötigten Personaldaten gegeben. Für die Auswertung waren Alter und Geschlecht relevant, zusätzlich hatten die Schüler die Möglichkeit, ihren Namen zu nennen oder sich einen Phantasienamen²⁹ zu geben.

Da ohnehin Anonymität zugesichert wurde, ist es berechtigt zu fragen, warum dann überhaupt eine Möglichkeit zur Antwortehrlichkeit vorgesehen war. Für die Aufnahme dieses Feldes gab es im wesentlichen zwei Gründe psychologischer Natur: zum einen sollte die Namensnennung eine gewisse Verbundenheit mit dem Testheft herstellen, zum anderen sollte die Möglichkeit des Phantasienamens eben die Phantasie anregen oder – anders ausgedrückt – einen Kreativitäts-impuls liefern. Beide Aspekte lassen die Namensabfrage als mögliches Motivationselement indirekter Art erscheinen.

Als Nebeneffekt wurde die spätere Zuordnung von Zweit-Testheften zum Bearbeiter erleichtert und sicherer gemacht: so konnte abgefangen werden, wenn ein Schüler – wie vorgekommen – sein Geburtsdatum zweimal verschieden angab („Zahlendreher“).

Die Altersangabe war entweder durch das Geburtsdatum oder – falls Angst bestand, anhand dieses Datums nachträglich identifiziert zu werden – durch Altersangabe in Jahren und Monaten möglich.

Das Schülerinformationsblatt ist im Anhang 10.2 vollständig abgedruckt.

2.6.4 Schülerlösungsblatt

Jede der Aufgaben wurde auf einer Doppelseite , also 2 DIN A4-Seiten, präsentiert. Damit wurde ein Umschlagen bei der Bearbeitung und Kommentierung einer Aufgabe unnötig. Zu jeder Aufgabe wurden analog zu TIMSS Lösungen mit Lösungskennbuchstaben vorgegeben. Zusätzlich wurde die gefundene Lö-

²⁹ Von dieser Möglichkeit wurde rege und mit viel Phantasie Gebrauch gemacht und traf damit wohl die Interessenlage der Schüler. Beispiele Schweizer Phantasienamen sind Bäri, Pamela Andersrum, Sorreti Sara, Zumbrunnen Simon, Tinkiwinki,, Fagabundä-Otto, Nuggi, Dipsy, Sorreti Sara, XTXer, MfG, Arqua? und Pikage, deutsche Schüler wählten etwa Frau Holle, Sadam Hussein, Cypress Hill, Aschenputtel, Wu-Tang, Diamond Dallas Page, Scyces oder Blümli.

sung auf der rechten Seite noch einmal ausdrücklich abgefragt: sie sollte dort explizit niedergeschrieben oder angekreuzt werden. Dieses Verfahren gestattet die Registrierung von Fällen mit evtl. korrekter Lösung in der Bearbeitung, aber irrtümlich falsch angekreuztem Wert. Zusätzlich sollten die Schüler damit nach Bearbeitung der Aufgabe auf die rechte Seite geleitet werden, die zu jeder Aufgabe einen kleinen Fragebogen enthielt.

2.6.5 Schülerfragebögen

Die Schülerfragebögen („Questionnaires“) zu *jeder* Aufgabe waren nach Bekanntheitsgrad der Aufgabe, Curriculumvalidität, Zeitraum seit Behandlung im Unterricht und Eigeneinschätzung zur Leistung differenziert. Zusätzlich wurden Informationen zum Lösen durch Raten abgefragt. Weiter waren freie Kommentare möglich.

Mein Ergebnis ist also:

Eine Aufgabe wie diese ist für mich neu.

Ähnliche Aufgaben kenne ich aus dem Unterricht.

Das ist schon länger her (mehr als sechs Monate).

Das ist schon etwas her (drei bis sechs Monate).

Das haben wir vor kurzem erst gemacht (weniger als drei Monate).

Ich konnte sie damals gut.

Ich kann sie noch gut.

Ich habe sie nie gut gekonnt.

Das Ergebnis kann man leicht erraten.

Ich rate das Ergebnis, weil mir dazu kein Rechenweg einfällt.

Zu der Aufgabe möchte ich noch sagen, daß ...

Die Informationen zur Curriculumvalidität aus Schülersicht sind nicht notwendig korrekt und auch nicht notwendig übereinstimmend mit der Sicht der Lehrer. Durch Vergleich mit den Lehrerbegleitheften lässt sich somit untersuchen, ob Lehrer- und Schülersicht zur Curriculumvalidität identisch sind oder sich nur teilweise überdecken.

Über MC-Aufgaben hinaus kann hier diskriminiert werden, ob eine Lösung geraten wurde oder nicht. Dies setzt eine ehrliche Antwort voraus. Da aber Raten durch die Fragen nicht als Lösungsweg verboten, sondern als Lösungsweg ausdrücklich zugelassen wurde, kann wohl im Wesentlichen Antwortehrlichkeit unterstellt werden. Somit lässt sich der Frage nachgehen, wie hoch der Rateanteil selbst bei ausreichender Bearbeitungszeit ist. Die offene Form der Formulierung mit vorgefertigter Begründung sollte ebenfalls Hemmungen abbauen und zu ehrlichen Antworten führen.

Im Raum für frei formulierte Zusatzbemerkungen konnten weitere aus Schüler-sicht relevante Angaben gemacht werden. Bei der Auswertung fanden sich hier Informationen über die Schwierigkeit der Aufgabe aus Schülerperspektive, ihr Gefühl für die Richtigkeit des Rechenweges oder zur weiteren Erklärung des Lösungsweges. Gelegentlich fanden sich auch Äußerungen, die Rückschlüsse auf die Haltung der Schüler zur Schule allgemein oder zur Mathematik zuließen.

2.7 Lehrerbegleitbuch

2.7.1 Lehrerinformationsblatt

Die Lehrer der teilnehmenden Schweizer Schulen wurden zusammen mit der Bitte um Testteilnahme vorab per email über die Testintentionen informiert, so dass das Informationsblatt nur noch einmal schlagwortartig das Ziel der Unter-suchung erläuterte.

Das Lehrerinformationsblatt für die den Test betreuenden Lehrer wurde zu ei-nem Zeitpunkt entwickelt, an dem noch nicht abzusehen war, ob der Verfasser selbst in den Schweizer Schulen die Tests betreuen konnte. Dies erforderte eine Gestaltung des Designs derart, dass möglichst gleichartige Testabläufe gesichert waren.

Als Hilfe dazu wurde der Testablauf in seinen wesentlichen Schritten schriftlich skizziert, und es wurden entsprechende Handlungsanweisungen vorgegeben.

Analog zu den Schülertexten wurde versucht, durch motivierende Begründung zu erreichen, dass die Lehrerfragebögen zu den Aufgaben bearbeitet wurden. Eine Motivation war hier besonders erforderlich, da es sich, wie unten beschrie-ben, um eine umfangreiche Arbeit handelte.

Damit die Lehrer nicht zur Lektüre des Schülerhinweistextes auf die Testhefte zurückgreifen mussten, wurde der Schülerinformationstext auch auf die Leh-reinformationsblätter gedruckt.

Der vollständige Text des Lehrerinformationsblattes (im Original beidseitig be-druckt) findet sich im Anhang 10.3.

2.7.2 Lehrerfragebogen

Analog zum Schülerquestionnaire wurde ein Lehrerfragebogen entworfen. Er sollte *zu jeder Aufgabe* folgende Informationen erbringen:

- Curriculumrelevanz aus Lehrersicht
- Einschätzung des Schwierigkeitsgrades bei unbekannten Aufgaben
- Übungsintensität bei behandelten Aufgaben
- Zeitraum zwischen unterrichtlicher Behandlung und Test
- sonstige Bemerkungen

und damit eine Kontrolle bzw. Wertung der Schülerangaben ermöglichen oder auch Differenzen in der Einschätzung der Curriculumrelevanz Schüler/Lehrer aufzeigen. Die Abfrage umfaßte die Fragekategorien oder ‚Dimensionen‘ Bekanntheitsgrad der Aufgaben, Übungsintensität, Zeitdimension und freie Zusatzinformationen: Sie war wie folgt gestaltet:

Für die SchülerInnen ist die Aufgabe meiner Ansicht nach:

<input type="checkbox"/> unbekannt und auch nicht durch Nachdenken lösbar			
<input type="checkbox"/> unbekannt, aber mit Überlegen lösbar			
<input type="checkbox"/> gut bekannt	<input type="checkbox"/> bekannt	<input type="checkbox"/> weniger bekannt	
<input type="checkbox"/> gut geübt	<input type="checkbox"/> etwas geübt	<input type="checkbox"/> kaum geübt	<input type="checkbox"/> nicht geübt
<input type="checkbox"/> vor kurzem < 1 Monat	<input type="checkbox"/> vor einiger Zeit 1..3 Monate	<input type="checkbox"/> schon länger her 3..6 Monate	<input type="checkbox"/> schon lange h > 6 Monate

evtl. Bemerkung:

Die Dimensionen sowie die Stufungen innerhalb der Kategorien ergeben in der gewählten Form ein handhabbares Antwortraster, von dem eine einheitliche personenunabhängige Bearbeitung zu erhoffen war.

Die Abfrage der Zeitdimension (verstrichene Zeit seit Behandlung von Aufgaben vergleichbaren Typs) kann trotz geringen Stichprobenumfangs zur Untersuchung eventueller Korrelationen zwischen verstrichenem Zeitraum und Bearbeitungsqualität und somit evtl. auch als Indikator für Langzeitlernfähigkeiten oder Vergessensfaktoren dienen. Zusätzlich wird ihr jedoch eine Funktion als qualitätshebendes Instrument zugeschrieben: durch Nachdenken darüber, vor welcher Zeit die Aufgabe im Unterricht behandelt wurde, wird auch erneutes Nachdenken darüber erzwungen, ob sie überhaupt behandelt wurde. So soll eine Markierung als ‚im Unterricht behandelt‘ vermieden werden, die nur deshalb gesetzt wurde, weil üblicherweise solche Aufgaben bereits behandelt sind, in der speziellen Klasse aber möglicherweise nicht. Damit wird also das erreichte Lernziel („attained curriculum“) bezogen auf die spezielle Klasse hinterfragt.

Wie dargelegt diente das Lehrerbegleitbuch zur Erfassung bzw. zum Abgleich der curricularen Validität der Aufgaben. Dazu wurde *zu jeder Aufgabe* die oben

abgedruckte Bewertungstabelle vorgegeben, und zwar jeweils mitsamt der Aufgabe auf einer Seite.

Somit waren vom betreuenden Lehrer immerhin 72³⁰ solcher Abfragen zu bearbeiten, und es konnte nicht vorausgesetzt werden, dass sich alle Lehrer dieser umfangreichen Aufgabe unterzogen. Dem entsprechend enthielt das Lehrerinformationsblatt explizit einen Hinweis auf die Bedeutung dieser Zusatzinformationen.

2.8 Probelauf in Deutschland

Auch bei sorgfältiger Kontrolle sind beim Testentwurf und beim Druck der Fragebögen, hier also der Testhefte, Fehler nicht auszuschließen, seien es vorher nicht erkennbare Probleme mit den ausgewählten Aufgaben, weniger günstige sprachliche Adaptionen, Druckfehler, Fehler bei der Vorgabe der Antworten, der vorgefertigten Auswahlfelder oder auch bei der Zusammenstellung der Einzelaufgaben zu den Gesamttestheften. Weiter ist denkbar, dass es bei der Übersetzung und Aufgabenformulierung zu aus Schülersicht unklaren Formulierungen kommt.

Die Testhefte wurden daher in kleiner Menge gedruckt und Schülern eines Grundkurses Mathematik der Jahrgangsstufe 13 zur Bearbeitung und Kontrolle vorgelegt.

Die Bearbeitung der Testhefte erfolgte altersentsprechend nicht in 90 Minuten, sondern die Schüler waren nach durchschnittlich 25 Minuten mit der Bearbeitung ihrer Testhefte fertig.

Es wurden wenige Druckfehler und falsch vorgegebene Antwortalternativen gefunden (z.B. „richtig ist...“ statt „mein Ergebnis ist ...“, Umformulierungen, die durch die andere textliche Präsentation im Vergleich zu TIMSS erforderlich wurden), teilweise mussten die Antwortvorgaben aufgabengemäß gesplittet werden.

Bei der Durchsicht wurde festgestellt, dass die Aufgabe J14 in ihrer ursprünglichen Form: „Dividiere 24,56 durch 0,16“ wenig Aufschluss über Kenntnisse zur Kommapositionierung bei der Dezimaldivision gab, da die Schüler lediglich die Dezimalkommata weglassen mussten. Die Aufgabe wurde daher modifiziert zu „Dividiere 24,56 durch 1,6“.

³⁰ Dabei sind die Zuatzaufgaben mit eingerechnet.

Die Aufgabe B.I03 lautet in der englischen Version

B.I03 The number of 750 mL bottles that can be filled from 600 L water is ...

Die Übersetzung im Probelauf („pilot survey“) war

B.I03 Die Anzahl von 750 mL-Flaschen, die man mit 600 L Wasser füllen kann, ist

Diese direkte Übersetzung empfanden alle (!) Schüler als sehr unklar, so unklar, dass vor Lösung nach dem Sinn gefragt wurde.

Nach Umformulierung zu

B.I03 Wie viele 750 mL-Flaschen kann man mit 600 L Wasser füllen?

erschien den Schülern die Aufgabe leicht. Es kann hier nur vermutet werden, dass die aufgrund der Konstruktion mit einem Relativsatz komplexere erste Frageversion das eigentliche Hindernis zum Verständnis ist, wohingegen die zweite Frageversion direkt ist und damit den Weg zum mathematischen Problem erst freimacht. Die spitzfindige Antwort ‚keine‘ – da in keine 750mL-Flasche 600 L Wasser passen – wäre in beiden Frageversionen denkbar und kann somit nicht das Verständnisproblem gewesen sein.

Die Aufgabe C.Q06 lautete in der Erstfassung:

C.Q06 Die Familie Huber verbraucht etwa 6000 L Wasser pro Woche. Wieviel Liter Wasser verbrauchen sie pro Jahr?

A. 30 B. 240 000 C. 300 000 D. 2 400 000 E. 3 000 000

Der Wechsel zum Plural-„sie“ blieb unbeantwortet, allerdings betrachteten die Schüler die Aufgabe als unlösbar. Auf Nachfrage wurde moniert, dass die Frage nicht mehr das Wort „etwa“ enthalte und daher keine der angebotenen Lösungen korrekt sei.

Die Frage wurde daher umformuliert zu

C.Q06 Die Familie Huber verbraucht etwa 6000 L Wasser pro Woche. Wieviel Liter Wasser verbraucht Familie Huber ungefähr pro Jahr?

A. 30 B. 240 000 C. 300 000 D. 2 400 000 E. 3 000 000

Bei einigen Aufgaben wurde angeführt, dass der Lösungsweg nicht beschreibbar sei, wie etwa bei der Drehung des Würfelsatzes im Kopf. Als Wege wurden angegeben „durch Vorstellen“ oder „räumliches Vorstellungsvermögen“. Da die TIMSS-Aufgaben allerdings wohl kaum auf Analysen des Lösungsweges hin optimiert wurden, war schon vorab zu erwarten, dass sich nicht alle Aufgaben gleich gut für Lösungsbeschreibungen eigneten³¹. Daher erfolgten keine entsprechenden Änderungen.

³¹ Dennoch wiesen die Schülerbearbeitungen später auch dort oft aufschlußreiche Kommentierungen und Aufgabenbeschreibungen auf, wo dies eher nicht zu erwarten war.

2.9 Präparation des Testmaterials

Nach Einarbeiten der Änderungen aus dem Probelauf wurden die Testhefte und Lehrerbegleitbücher in ausreichender Zahl (350 Testhefte, 12 Lehrerbegleitbücher) gedruckt, sortiert, geheftet, und gelocht. Um sicherzustellen, dass die einzelnen Testhefte in etwa gleich häufig zur Verwendung kamen, wurden sie für die spätere Austeilung gruppiert: jeweils ein Satz Testhefte A bis F bildeten eine zusammengebundene Verteileinheit. Damit musste der betreuende Lehrer eine der Schülerzahl entsprechende Anzahl von Sechserbündeln öffnen. So ergab sich automatisch eine gleichmäßige Verteilung der Testhefte³².

³² Allerdings unter Benachteiligung der letzten Testhefte dieser Bündel; um auch dort eine ausreichende Anzahl von Testheften zur Verteilung zu bringen, hat der Verfasser bei den Tests, an denen er teilnahm, gezielt ein oder zwei dieser Testhefte eingestreut.

3 Realisierung und Datenpräparation

Die folgenden Ausführungen 3.1 bis 3.4 geben einen Einblick in den Ablauf der Tests in den beteiligten Schulen und die Schulsituation. Zusammen mit den Protokollen des jeweiligen Testablaufs entsteht ein (weder bewertender noch bewerteter) Eindruck zu den unterschiedlichen Arbeits- und Unterrichtsumgebungen.

3.1 Zeitplan

Die Tests wurden sowohl in Deutschland als auch in der Schweiz in den jeweils beiden letzten Wochen vor Beginn der Sommerferien durchgeführt. Sie fanden statt in Deutschland in der Zeit vom 31.5. bis zum 14.6., in der Schweiz vom 22.6. bis zum 28.6.1999.

Der gewählte Zeitraum der Tests am Ende des Schuljahres ließ erwarten, dass der Unterrichtsstoff der 7. Klassen nach curricularen Vorgaben so weit wie möglich behandelt war.

Als Nachteil war zu vermuten, dass einige Schüler aufgrund schon feststehender Nichtversetzung keinerlei Interesse an einer sorgfältigen Testbearbeitung hatten. Dies wurde nicht nur von den Fachlehrern betätigt, sondern bei Durchsicht der Schülerarbeiten sind deutlich Bearbeitungen zu erkennen, die grobe Kritzeleien, Verwendung von Fäkalsprache in den Bemerkungen, kaum Aufgabenbearbeitung etc. aufweisen.

Somit war zu entscheiden, ob solche „Nichtbearbeitungen“ in das Bewertungsvolumen aufgenommen werden sollten. Eine Aussonderung von Daten, auch nachvollziehbar und begründet, wäre allein deshalb schon bedenklich, weil die o.a. Gründe zwar hier, nicht aber in der TIMSS-Studie extrahierbar sind. Allein schon aus Gründen der Vergleichbarkeit wurden daher solche Arbeiten mit in die Auswertung einbezogen.

3.2 Schulbeschreibungen

Die beteiligten Schulen werden in den später folgenden Tabellen kodiert als GER1, GER2, CH3 und CH4. Diese Kennungen stehen für:

- GER1: ein Stadtrandgymnasium in einem Schulzentrum (gemeinsam mit einer Hauptschule) in Mönchengladbach. Am Test beteiligt waren drei siebte

Klassen (7A, 7B, 7C), eine Kombination von Schülern dieser Klassen als gesonderte Lateingruppe (7AB L mit Beginn der zweiten Fremdsprache Latein in Klasse 7) und eine Klasse 8A als Vergleichsgruppe,

- GER2: ein Innenstadtgymnasium mit einer Klasse 7C aus Mönchengladbach,
- CH3: eine „Oberstufenzentrum“ genannte öffentliche Schule in einem Vorort von Bern, mit einer Sekundarklasse 7KA, in der Schüler mit guten Mathe-matik-Leistungen zusammengezogen werden (sog. „Spezialklasse“) und ei-ner Sekundarklasse 7KE als „Mittelfeldgruppe“,
- CH4: ein Innenstadtgymnasium in Bern (privat, aber mit voller staatlicher Anerkennung, Hausmatur eidgenössisch anerkannt, eine Klasse 7G1, unterrichtet nach dem allgemein gültigen Sekundarlehrplan und zusätzlichem Stoff, aber mit Fachschwerpunkt Geometrie erst in Klasse 8 sowie einer zweiten Klasse 7G2 als Parallelklasse dazu (nach Auskunft des Fachlehrers „etwas stärker“ als die erste Klasse“).

3.3 Klassenräume

Zur evtl. Beurteilung des Lernumfeldes ist eine Kenntnis der Klassenraum-ausstattung hilfreich. Daher wurden stichwortartig Klassenraumbeschreibungen notiert. Im einzelnen erschienen folgende Beobachtungen relevant:

Schulen GER1 und GER2: Klassenräume bestückt mit Bänken, Tafel, ggf. Pro-jektionswand für OHP, Lehrerpult, ggf. etwas Wandschmuck (Bilder, Seitenta-fel), Waschbecken, große Fenster,

Schule CH3: eine Klasse 7, alle Tische höhenverstellbar, Frontalanordnung, Jungen und Mädchen jeweils für sich in den Reihen, eine Wand des Schulrau-mes bis auf den Fußboden als Fensterwand mit Ausblick ins Grüne, Waschbek-en, reichhaltiger Wandschmuck mit offenbar selbstgemalten farbigen Bildern von Würfelschnitten etc. (Schnittebenen, einbeschriebene Tetraeder, ...), mit Kandinsky-Bildern, drei Groß-Comicbildern über der Tafel, Overheadprojektor und Projektionsleinwand, 3 Computer, Radio, Dartscheibe, Duden Bände 1 bis 10 und 10-bändiges Dudenlexikon.

Schule CH4: eine Klasse 7, einzeln verstellbare Bänke mit Klappe, Glaswand mit Öffnung zum grünen Innendach mit aufstehendem Gewächshaus, Projekti-onstafel, um die Tafel als Wandschmuck viele gemalte Tigerköpfe, Schweizer-flagge über der Tafel.

3.4 Testverlaufsprotokolle

Zum Ablauf der Tests, bei denen der Verfasser anwesend war, wurden Protokolle gefertigt. Man erhält auch ohne weitere Kommentare einen Eindruck der Testsituationen und des Testablaufs im Ländervergleich.

Testverlauf Kl. 7C GER1, 7.6.99 , 3.u.4. Std.

10.07 Beginn mit 3. Std.

Ruhe herstellen, Testerläuterungen, Genese, Austeilen, starkes Kichern

10.18 Beginn, schwierig, Ruhe herzustellen

10.21 absolute Ruhe

10.24 leises Kichern,

Aufforderung zur Konzentration und Ruhe, ruhig bis

10.40 eigentlich Pause, Arbeit aber durchgehend über Pause hinweg

10.43 leichte Unruhe

10.55 sechs Testhefte werden abgegeben (5 Jungen, 1 Mädchen, Testhefte AAABCE)

11.10 starke Unruhe

ein Mädchen möchte wissen, „ob man auch erfahren kann, ob man es richtig gemacht hat“

11.21 alle abgegeben bis auf 4 Testhefte

11.24 alle abgegeben bis auf 2 Schüler

Gesamteindruck: Die Aufforderung zur Alleinarbeit war sehr schwierig durchzuhalten, erst recht nach Abgabe einiger Testhefte; es gab einen starken Wunsch nach Besprechung mit den Nachbarn.

Testverlauf Klasse 7KA CH3, 22.6.99 , 7.25 – 9.00

7.45 nach Vorstellung und Verteilen sehr ruhige Bearbeitung

unter Aufsicht vom Verfasser, da der Fachlehrer den Raum verlassen musste (anderweitige dienstliche Aufgaben)

8.50 leichte Unruhe, weil Fachlehrer zurückkehrt
danach ruhige Arbeit

Gesamteindruck: Es gab eine große Konzentration bei der Arbeit; der Lehrer wollte bei Problemen helfen durch Hinweis darauf, ob das schon „drangewesen“ sei. Nach meiner Bitte, dies zu unterlassen, blieb Ruhe während der gesamten Arbeitszeit.

Testverlauf Klasse 7KE CH3, 22.6.99 , 7.25 – 9.00

Testverlauf nach Mitteilung der Fachlehrerin, da der Verfasser beim Test in der Parallelklasse war.

- 8.10 Beginn, Klasse begann sehr spät, da Lehrerin noch anderweitig zeitlich gebunden
bis auf einen Schüler allerdings „sehr früh“ fertig
danach wurden teilweise die Hefte getauscht, ohne dass sie weiter bearbeitet wurden. Die Schüler meinten: „zum Angucken“, um zu wissen, was „die anderen“ bearbeiten mussten.

Testverlauf Klasse 7G1 CH4, 25.6.99 , 13.20 – 14.50

- 13.20 Vorstellung, Testerläuterung, Austeilen
13.27 Beginn mit absoluter Ruhe
13.50 Tendenz zur Diskussion mit dem Nachbarn, Hinweis und Bitte, dies zu unterlassen
13.55 wieder Ruhe bis gegen 14.20
14.25 viele Schüler sind fertig und verlangen nach einem weiteren Testheft
mit entsprechendem Vermerk auf dem Ersttestheft wurde auf Wunsch ein Zweittestheft ausgeteilt
nach Austeilen wieder Ruhe
14.40 erste Abgaben
14.48 letztes Testheft wird abgegeben.

Gesamteindruck: intensive Einzelarbeit

Testverlauf Klasse 8A GER1 , 7.6.99, 800 – 9.44

- 8.00 Beginn mit Erläuterungen
Hinweis auf „lieber durchstreichen als radieren“ (da dann Lösungswege besser transparent werden)
8.10 häufiges Nachfragen
9.00 abfallende Konzentration
9.15 Konzentration weitgehend verlorengegangen
sieben Schüler fertig mit ihrem Testheft
9.17 16 Schüler haben abgegeben

3.5 Bewertungsschlüssel

Zur Gewinnung von Hinweisen auf signifikant unterschiedliche Schülerleistungen anhand der elaborierten Itembearbeitungen ist ein Quervergleich über alle Schülerbearbeitungen zu jeder Aufgabe erforderlich.

Dies impliziert unmittelbar, dass die Bivalenz richtig/falsch nicht mehr letztlich entscheidendes Produkt der Schülerbearbeitungen ist. Vielmehr ist es eine deutlich darüber hinaus gehende Frage, auf welchem Weg der Schüler seine Aufgabenbearbeitung angeht, ob man seinen Aufgabenbearbeitung Qualitäten zuschreiben kann, selbst wenn sie fehlerbehaftet sind, wie weit er vordringt und woran er möglicherweise scheitert.

Ab einer gewissen Anzahl zu vergleichender Aufgabenbearbeitungen ist aber der Einzelkorrektor mit der Memorierung aller aufgabentypischen Besonderheiten der Schülerlösungswege überfordert. Daher bedarf es zur vergleichenden Analyse einer Kodierung, die Schülerlösungen kategorial und klassifizierend erfasst und die „Qualität“ der Einzelschritte der Schülerlösungen oder die Probleme, die zum Scheitern bei der Aufgabenbearbeitung führten, einer (auch statistischen) Analyse zugänglich macht.

Dazu erscheint ein kategoriales System, welches alle Aufgabenbearbeitungen gleichartig kodiert – ggf. mit aufgabenspezifischen Unterschlüsselungen – eher geeignet als eine aufgabenbezogene Fehlerkodierung. Es erfasst damit mehr als die meist nur dichotomen Antwortkategorien bei TIMSS. Die hier eingeführten „polytomous Antwortvariablen“ (Rost, 1996, S.89) sind allerdings nicht als äquidistant zu interpretieren. Sie umfassen eher Klassen der „Bearbeitungsqualität“ in einem ordinalen Rahmen.

Im einzelnen wurden folgende naheliegenden Anforderungen an das Schlüsselsystem gestellt:

- Schlüsselermittlung über ein Standardlösungsblatt
- mit Musterschlüsselvorgaben,
- kodiert als metrische Schlüssel,
- Schlüsselmetrik mit zumindest teilweise ordinalem Charakter,
- möglichst Unabhängigkeit von der Person des Korrektors, also Nachvollziehbarkeit auch durch Fremdkorrektoren, idealerweise Annäherung an Signierobjektivität,
- Möglichkeit zur TIMSS-analogen Leistungsbewertung,
- Differenzierung nach Fehlertypen,
- Erfassung von Ratehäufigkeit und Ratequalität,
- Differenzierung in der Lösungsqualität nicht nur bei falsch, sondern auch bei richtig gelösten Aufgaben,
- dynamisch mögliche Schlüsselerweiterung.

Besondere Aufmerksamkeit verdient die Anforderung, Ratehäufigkeit und Ratequalität zu erfassen. Dies sind Bearbeitungsmerkmale, die sich aus MC-Tests keinesfalls direkt erfassen lassen. Sie können aber bei geeigneter Anlage diskriminieren, ob testwiseness-Strategien eingesetzt wurden, ob ‚blindes‘ Raten bei schwierigen Problemen vorlag oder ob Raten wegen der Einfachheit der Frage eingesetzt wurde. Zur Erfassung der Ratestrategien wurde der aufgabenspezifische Schülerfragebogen entsprechend ausgelegt. Da ‚Raten‘ als Problemlösestrategie ausdrücklich zugelassen wurde, kann von einem hohen Maß an Antwortreabilität ausgegangen werden.

Zur Gesamtkodierung wurde ein Grobraster angelegt mit Kennziffern von

- Null bis neun für geratene oder deutlich fehlerbehaftete Lösungswege, wobei die Fehlerart mit kodiert wurde;
- 10 bis 12 für im Wesentlichen korrekte Lösungswege, die aber nach TIMSS bereits als falsch bewertet worden wären, und
- 13 bis 20 für Lösungswege, die nach TIMSS als korrekte Lösung bewertet worden wären, gleich ob eher zufällig richtig, richtig geraten oder korrekt bearbeitet.

Nachfolgend werden beispielhaft das vollständige Lösungs- sowie das Bewertungsblatt für das Testheft A aufgeführt:

Tabelle 7: Lösungsblatt zu Testheft A

1.	K01	D	
2.	P10	B	4m
3.	K03	A	
4.	K09		$\frac{115}{24} = 4\frac{19}{24}$
5.	V03	D	5 9
6.	J18	B	3
7.	V02		Gebäude A: 9600/Jahr oder 800/Monat ist günstiger B: 9900/Jahr oder 825/Monat
8.	J14	C	15,35
9.	Q03	B	1500 Minuten ist die längste Zeit (25 Stunden)
10.	M07	D	80 Grad
11	O02	C	25 %
12.	L16		x = 7

Tabelle 8: Ergebniskodierung für Testheft A

Nr.	Fehlerart
20	Weg/Rechnung korrekt und korrekt angekreuzt/angegeben
19	Weg/Rechnung teilweise falsch, aber Ergebnis korrekt und korrekt angekreuzt
18	kaum/kein Weg/Rechengang vorhanden, aber korrekt angekreuzt/angegeben
17	Weg/Rechnung nicht nachvollziehbar, aber Ergebnis korrekt und korrekt angekreuzt
16	Rechenfehler heben sich weg, daher korrektes Ergebnis und korrekt angekreuzt
15	Weg/Rechnung falsch, aber korrekt angekreuzt/angegeben
14	explizit als geraten gekennzeichnet und korrekt angekreuzt auch: als unklar bezeichnet und korrekt geraten
13	keiner der Fälle 14 - 20, aber Ergebnis korrekt angekreuzt/angegeben, z.B. auch / abgemessen statt gerechnet / abgeschätzt / ausprobiert / ... /
12	Weg/Rechnung und Ergebnis korrekt, aber nicht angekreuzt
11	Weg/Rechnung und Ergebnis korrekt, aber falsch angekreuzt/angegeben
10	Weg/Rechnung korrekt, aber ein leichter Rechenfehler, daher falsch angekreuzt/angegeben
9	Weg/Rechnung korrekt, aber mehrere leichte Rechenfehler, daher falsch angekreuzt
8	Weg/Rechnung vom Ansatz her korrekt, aber ein/mehrere grobe(r) Rechenfehler
7	Weg/Rechnung vom Ansatz her korrekt, aber ein/mehrere grobe(r) Denkfehler, auch: falsche räumliche Drehung / als „Wissenslücke“ gekennzeichnet /
6	Weg/Rechnung vom Ansatz her falsch und daher falsch angegeben/angekreuzt
5	Rechengang sinnleer, Ergebnis falsch angegeben/angekreuzt
4	kein/e Weg/Rechnung vorhanden und Ergebnis falsch angegeben/angekreuzt
3	explizit als geraten gekennzeichnet und falsch angekreuzt
2	als „nicht verstanden“ oder „kann ich nicht“ gekennzeichnet und nicht bearbeitet
1	keiner der Fälle 2 - 12 , aber Ergebnis falsch angekreuzt/angegeben, z.B. auch / nicht angekreuzt / mehrfach angekreuzt / nach Ansatz Bearbeitung abgebrochen / ... /
0	Aufgabe nicht bearbeitet

Die Kennziffern sind von einem externen Korrektor nachvollziehbar. Nicht präzise definierbar sind die Schlüsselkriterien „leichter“ und „grober“ Rechenfehler. Bei der Kodierung wurden Kopfrechenfehler wie etwa $7 * 8 = 54$ oder Abschreibefehler bei Zahlen oder vertauschte Ziffern als leichte Rechenfehler eingestuft. Grobe Fehler waren dagegen Verstöße gegen Rechenregeln, falscher Umgang mit Brüchen etc.

Ebenfalls nicht immer aus den Kodierhinweisen ableitbar ist der Unterschied zwischen groben Rechenfehlern (Schlüsselziffer 8) und groben Denkfehlern (Schlüsselziffer 7). Der Schlüssel 7 wurde daher nur dann gesetzt, wenn die Bearbeitung Fehlansätze im Rechen- oder Gedankengang erkennen ließ wie beispielsweise falsch basierter Dreisatz, nicht logische Folgerung etc. Die nicht immer eindeutige Abgrenzung zwischen diesen Fehlerarten wird allerdings nur relevant, wenn sich eine statistische Datenanalyse speziell darauf stützt. Dies ist hier nicht der Fall.

Das Kategorienschema erscheint insgesamt als einfach genug, um eine hinreichende Signierobjektivität zu erzielen. Eine darüber hinaus gehende Absicherung auf dem Wege der Parallelsignierung durch Zweit- oder Drittkorrektoren, Berechnung von Übereinstimmungsmatrizen und -koeffizienten etc. sprengt bei weitem den Rahmen des in dieser Einzelarbeit realisierbaren Aufwands.

Die oben vorgegebene grobe Dreigliederung des Bewertungsrasters wurde vor Durchsicht der Testhefte entworfen. Eine teilweise weitere Untergliederung geschah als interaktiver Prozess bei Durchsicht der Testhefte, wenn aufgabentypischen Fehler oder Lösungswege erkennbar wurden.

3.6 Datenpool

Insgesamt wurden (Probelauf und Zweittesthefte eingeschlossen) 274 Testhefte bearbeitet. Eine Analyse der Verteilung der Testhefttypen auf die Klassen ergab, dass die angestrebte homogene Verteilung der Testheftvarianten innerhalb der einzelnen Klassen und sogar bei den Zweittestheften erfüllt werden konnte.

Die erste Auswertungsphase bestand darin, nach Durchnummerierung der Testhefte und Erfassung des Testheft-Typs sowie des Geburtsdatums bzw. Alters in Jahren alle 2772 TIMSS-konformen Aufgaben der Testhefte A bis E durchzuarbeiten, unter parallel laufender Schlüsselerweiterung zu kodieren und tabellarisch festzuhalten.

Bei der Durchsicht erwies es sich als sinnvoll, neben der Kennzeichnung der Lösung durch Schlüsselziffern weitere Bemerkungen bzw. Hinweise mit aufzunehmen, wenn Schülerlösungswege besonders charakteristische oder sonst auffällige Elemente enthielten. Gleichzeitig ergaben sich so bei Mehrfachkennzeichnungen zusätzlich Hinweise auf Aufgaben, bei denen eine vergleichende Analyse lohnend schien. Wie weiter unten dargelegt wird, erwies sich dieses so implementierte Zweitschlüsselsystem als ausgesprochen wertvoll zur Lokalisierung didaktischer Besonderheiten. Insbesondere ergaben sich durch Häufung derart markierter Aufgaben Hinweise auf Curriculuminvaliditäten.

Die anschließende Übertragung in die elektronische Tabelle(n) folgte unter doppelter Kontrolle: die Werte wurden sowohl horizontal als auch vertikal lesend mit der Handschrift verglichen.³³ Die statistischen Analysen der folgenden Kapitel basieren auf dieser Gesamt-Datentabelle. Sie wird hier auf Grund ihres Umfangs nicht abgedruckt.

Bei den didaktischen Bewertungen wurden zusätzlich die Schülerskripte herangezogen.

³³ Es wurden 8 Übertragungsfehler gefunden und eliminiert. Dies ist bei 2772 übertragenen Zahlen eine akzeptable Fehlerrate.

Die Originaltabelle enthielt auch die Schülernamen oder frei gewählte Pseudonymnamen, soweit sie von den Schülern angegeben wurden. Da Anonymität der Auswertung zugesichert war, wurden die Schülernamen gelöscht, nicht aber die Geburtsdaten/Altersangaben, da daraus Informationen über die Altersstreuung in den Populationen gewonnen werden können.

Bei nicht anonymer Aufgabenbearbeitung ergäbe sich ein breites weiterführendes Forschungsfeld. Denkbar wäre, auf Korrelationen zwischen TIMSS-Testleistungen und schulischen Leistungsbewertungen zu untersuchen. Dabei wären jedoch zahlreiche rechtliche Probleme zu bedenken. So könnte ein solcher Datenabgleich auch in die Persönlichkeitsrechte der betroffenen Lehrer eingreifen.

4 Detailanalysen zur Bearbeitungsqualität

Das Kapitel 4 befaßt sich mit extrinsischen Strukturen wie etwa sozialen Skripten oder ähnlichen Faktoren, die neben mathematischen Aspekten den TIMSS-Erfolg determinieren können. Neben curricularen Aspekten wird insbesondere darauf eingegangen, ob länderspezifischen Unterschiede in der Leistungsbereitschaft der Schüler vorlagen. Dieser Aspekt scheint zur Zeit in der öffentlichen wie auch in der fachdidaktischen Diskussion an Gewicht zu erhalten. Bedingende Faktoren dafür sind ebenfalls Gegenstand öffentlicher Diskussion, bedürfen aber wissenschaftlicher Hinterfragung.

Diesem Aspekt der außercurricularen Testbedingungen wird in den Augen des Verfassers bisher noch zu wenig Aufmerksamkeit gewidmet. Eine wie auch immer geartete Leistungsabfrage ist dann wenig aussagekräftig, wenn die Testpopulation nicht bereit ist, auf die Testanforderungen einzugehen. Im wesentlichen muss daher die Leistungsbereitschaft als determinierender Faktor mit erfasst werden.

Die Leistungsbereitschaft kann aus den soziokulturellen Determinaten des jeweiligen Schulsystems folgen oder auch aus den gleichen Gründen fehlen; eine Interpretation vergleichender Testergebnisse bedarf aber der Berücksichtigung eben dieses wesentlich diskriminierenden Faktors.

Die Bemerkung von Haenisch zur Verfahrenskritik gewinnt damit erneut an Gewicht. In der vorliegenden Studie wurde versucht, diesen Faktor durch entsprechend motivative Einführungsblätter gering zu halten.

Leistungsbereitschaft ist ein schwer zu messender Faktor. Nur gelegentlich ergaben sich aus den Schülerkommentaren dazu direkte Hinweise. Die Studie geht daher einen indirekten Weg über die Bearbeitungsinsentität. Insbesondere werden die Konsequenz und Ausdauer bei der Bearbeitung der Aufgaben, die Auslassung von Aufgaben zum Ende der Testhefte oder die Auslassung von Aufgabenketten untersucht. Es können Indikatoren für Leistungsbereitschaft sein, wenn ein Proband gleich ganze Reihen von Aufgaben überspringt oder die Aufgaben am Ende des Testheftes einfach nicht mehr bearbeitet. Zusätzlich können freie Aufgabenkommentare Anzeichen für Leistungs- oder Bearbeitungsbereitschaft geben. Solche Bemerkungen wurden gelegentlich in sehr drastischer Form geäußert. Dabei handelte es sich zumindest in Deutschland in einer Klasse um Schüler, die bereits wussten, dass sie am Ende des Schuljahres nicht in die nächst höhere Klasse versetzt wurden. (Diese bei einem anonymen Test sonst nicht zugängliche Information konnte gewonnen werden, weil der Verfasser den Test neben dem Fachlehrer mit betreute und die Schüler sowie als Vorsitzender der Versetzungskonferenz auch ihre Leistungen kannte.)

4.1 Vergleich der Altersstrukturen

Zum Vergleich der Länderdaten dieser Studie miteinander und mit TIMSS ist es von Bedeutung festzustellen, ob sich die Altersstrukturen der Probanden gleichen, da die Lern- und Wissensstände sicherlich auch altersabhängig sind. Das Auswahlkriterium für die TIMSS-Testpopulationen wurde bereits in 2.3 zitiert:

„All countries that participated in TIMSS were to test students in the two grades with the largest proportion of 13-year-olds in both mathematics and science.“(online, TIMSS 1998a, S.8)

Baumert/Lehmann bemerken dazu:

„Um den internationalen Vergleich der 7. und 8. Jahrgangsstufen zu erhalten, wurde die vorgegebene Populationsdefinition in Deutschland nicht exakt eingehalten. Die beiden angrenzenden Jahrgänge mit dem größten Anteil 13jähriger – die internationale Zielgruppe – sind in der Bundesrepublik im Unterschied zu den meisten anderen teilnehmenden Ländern die Jahrgangsstufen 6 und 7.“ (Baumert et al., 1997, S.51-52)

In der vorliegenden Studie wurde für alle teilnehmenden Schüler das Alter für das Datum 20.6.1999 berechnet, da dieser Tag zwischen den Testtagen in Deutschland und der Schweiz lag, jeweils nur mit wenigen Tagen Abstand.

Für Deutschland ergab sich ein Durchschnittsalter von 13 Jahren und 7 Monaten, für die Schweiz von 13 Jahren und 10 Monaten. Die Anzahl der ‚älteren‘ Schüler, definiert als ein Alter mehr als 10 Monate über dem Durchschnittsalter, war in beiden Ländern mit weniger als drei Schülern gering.

Das von Baumert angeführte Item „Klassenwiederholung“ wurde hier nicht explizit erhoben. Man entnimmt aber aus den Datentabellen, dass kaum „Ausreißer“ in der Altersstruktur und somit vermutlich auch wenig Repetenten in den Klassen sind.

Die Tabelle listet die Altersverteilung der Schüler im Vergleich zu TIMSS. Die TIMSS-Vergleichsdaten (Zeilen 4 und 5) wurden entnommen (Baumert et al., 1997, S.52).

Tabelle 9: Altersverteilung der Schüler nach TIMSS und der vorliegenden Studie

Alter	≤ 12	13	14	15	16	17
GER Klassen 7	5	84	11	0	0	0
CH Klassen 7	1	70	29	0	0	0
TIMSS 2 GER JgSt 7	2	71	23	4	0	0
GER Klasse 8	0	4	68	28	0	0
TIMSS 2 GER JgSt 8	0,1	2	70	22	5	0,6

(Angaben gerundet in Prozent)

Die Aussage, dass gut ein Viertel der deutschen Siebklässler zum TIMSS-Erhebungszeitpunkt schon 14 oder 15 Jahre alt war (vgl. Baumert et al., 1997, S. 51-52), wurde hier nicht bestätigt. Zwar sind 29 % der Schweizer Schüler zum Erhebungszeitpunkt bereits 14 Jahre, aber nur 11 % der deutschen Schüler 15jährige finden sich weder in der Schweiz noch in Deutschland in den untersuchten Klassen 7.

Die Altersstruktur der Achtklässler entspricht dagegen etwa den TIMSS-Daten mit dem Unterschied, dass hier keine 16jährigen Schüler (oder noch älter) in der Klasse sind, dafür aber ein höherer Anteil von 15jährigen.

Die Altersstruktur der Schweizer Schüler entspricht daher insgesamt etwa der Verteilung der TIMSS-Studie, die deutschen Schüler sind ähnlich verteilt, aber etwas jünger.

4.2 Schülerleistungsvergleich nach TIMSS-Maßstab

Für Vergleiche von Befunden aus Anschlußstudien mit den vorliegenden TIMSS-Daten ergibt sich ein grundlegendes Problem: die TIMSS-2-Daten liegen nicht nach Schulformen separiert vor. Die öffentlich verfügbaren Daten bilden einen Querschnitt über alle beteiligten Schulformen der jeweiligen Länder. In der vorliegenden Studie werden Gymnasialklassen aus Deutschland und vergleichbare Schweizer Klassen miteinander verglichen. Für relative Leistungsdimensionen ist dies unerheblich. Bei absoluten Werten kann man aus den Vergleichen der TIMSS-Hauptkohorten nach (Baumert et al., 1997, S. 131) ersatzweise davon ausgehen, dass die jeweiligen Prozentangaben für Gymnasiasten über denen des schulformübergreifenden Durchschnitts liegen.

Um einen Hinweis zur Vergleichbarkeit des Länder-Rankings in dieser Studie mit dem TIMSS-Ranking zu gewinnen, wird im Folgenden der Frage nachgegangen, ob sich der aus dem TIMSS-Ranking bekannte Leistungsrückstand der deutschen gegenüber den Schweizer Schülern qualitativ reproduziert.

Dazu wurden die Schülerlösungen anhand der Bewertungsschlüssel so bewertet, wie sie auch in TIMSS bewertet worden wären.

- Es wurden *die* Aufgaben als korrekt gelöst angesehen, die tatsächlich richtig angekreuzt oder bei denen die richtige Lösung angegeben war, gleich aus welchen Gründen dies erfolgte (gelöst oder geraten).
- Aufgaben, die nur deshalb falsch waren, weil sie trotz richtiger Lösung falsch angekreuzt waren oder die lediglich einen leichten Rechenfehler enthielten, sonst aber korrekt gelöst waren, wurden als falsch gewertet.

- Aufgaben, die explizit als geraten gekennzeichnet und richtig geraten waren, wurden als richtig gewertet.

Die nachfolgende Tabelle zeigt das Resultat dieser ersten Analyse.

Tabelle 10: Leistungsvergleich nach TIMSS-Maßstab

Klasse	Schule	Schülerzahl	Anzahl bearbeiteter Aufgaben	korrekte Lösungen in %
7A	GER Gymnasium 1	21	260	54
7B	GER Gymnasium 1	23	286	54
7C	GER Gymnasium 1	21	260	52
7 AB L*	GER Gymnasium 1	19	236	58
7C	GER Gymnasium 2	23	286	55
<hr/>				
7KA	CH Oberstufenzentrum	18	223	73
7KE	CH Oberstufentzentrum	12	149	73
7G1	CH freies Gymnasium	16	198	66
7G1**	CH freies Gymnasium	12	94	68
7G2	CH freies Gymnasium	16	198	74
7G2**	CH freies Gymnasium	10	82	74
<hr/>				
8A	GER Gymnasium 1	24	298	61
*: Lateingruppe: Schüler aus 7A und 7B, die Latein als 2. Fremdsprache gewählt haben.				
**: Schüler, die ein zweites Testheft zur Bearbeitung erbatzen.				

Man entnimmt direkt:

- Der aus den TIMSS-Daten bekannte Leistungsabstand Schweiz / Deutschland wird auch bei kleinen Schülergruppen stabil reproduziert.
- Etwa die Hälfte der Schweizer Schüler haben während der 90-minütigen Bearbeitungszeit ein zweites Aufgabenheft ganz oder teilweise bearbeitet. Dabei bleibt die Lösungsrate stabil.
- Die Schweizer Klassen 7G1 und 7G2 wurden vom gleichen Mathematiklehrer unterrichtet. Er schätzte vorab (mündliche Mitteilung) die Klasse 7G2 als „leistungsstärker“ als die Klasse 7G1 ein. Dieses Ranking durch den unterrichtenden Lehrer wird hier bestätigt.
- In Deutschland wird ein Zuwachs korrekter Lösungen von Klasse 7 (Mittel etwa 54%) zu Klasse 8 (61%) erkennbar, allerdings bleiben die Leistungen der Klasse 8 GER hinter allen Klassen 7 CH (Mittel etwa 71%) erkennbar zurück.

Somit wird der TIMSS-Rankingunterschied Schweiz / Deutschland qualitativ stabil reproduziert.

4.3 Mängeltoleranter Schülerleistungsvergleich

Es wurde vermutet (vgl. dazu die o.a. Kritik von Haenisch), dass die deutschen Schüler die Aufgaben möglicherweise nicht mit der gleichen Sorgfalt bearbeitet haben wie die Schweizer Schüler, so dass z.B. Rechen- und sonstige Flüchtigkeitsfehler oder auch „nachlässige“ Aufgabenbearbeitungen³⁴ mit für den Rankingunterschied ursächlich zeichnen könnten.

Zur Untersuchung dieser Vermutung wurden in einem weiteren Schritt die Bedingungen für das Urteil „korrekte Lösung“ bzw. das Werturteil „richtige Bearbeitung der Aufgabe“ schrittweise reduziert. Daraus ergeben sich möglicherweise Hinweise auf die Qualität der Aufgabenbearbeitung über eine richtig/falsch-Bivalenz hinaus.

Die Reduzierung erfolgte zweistufig.

- Im ersten Schritt wurden auch diejenigen Aufgaben als korrekt gelöst angesehen, die zwar korrekt gelöst, aber nicht oder falsch angekreuzt waren oder bei denen ein falsches Ergebnis lediglich auf einem leichten Rechenfehler beruhte - bei ansonsten richtigem Lösungsweg.
- Im zweiten Schritt wurden auch diejenigen Aufgaben noch als „richtig“ bearbeitet angesehen, die weitergehende Mängel enthielten wie beispielsweise Beantwortung der inversen Frage (z.B. korrekte Angabe des Benzinverbrauchs statt der erfragten Restmenge Benzin), falsche Lösungen aufgrund ungenauer Zeichnungen, mehrere leichte Rechenfehler, vertauschte Koordinaten etc.

Den jeweiligen Dekrementen in der Bewertungsstrenge entsprechen somit Inkremente in der Lösungshäufigkeit.

Die folgende Tabelle zeigt die resultierenden Lösungshäufigkeiten aufgeschlüsselt für die einzelnen Klassen.³⁵

³⁴ Die Qualität der Aufgabenbearbeitungen wird in 4.5 weiter untersucht.

³⁵ Auf eine Angabe der Standardabweichungen wird auf Grund der geringen Probandenzahl verzichtet.

Tabelle 11: mängeltoleranter Schülerleistungsvergleich (1)

	Schü-ler-zahl	mängel-frei	leichte Mängel toleriert	weitergehende Mängel toleriert
7a GER GY1	260	54	57	61
7b GER GY1	286	54	57	63
7c GER GY1	260	52	58	63
7L GER GY1	236	58	60	64
7c GER GY2	286	55	58	62
7KA CH	223	73	74	79
7KE CH	149	73	75	80
7G1 CH	198	66	69	74
7G1CH +	94	68	70	79
7G2 CH	198	74	79	84
7G2 CH +	82	74	78	85
8a GER GY1	298	61	65	70
Angaben in % richtiger Lösungen, Werte gerundet				

Daraus folgt gemittelt über die Klassen der jeweiligen Länder:

Tabelle 12: mängeltoleranter Schülerleistungsvergleich (2)

	fehlerfrei	leichte Mängel toleriert	weitergehende Mängel toleriert
7 GER	55	58	63
7 CH	71	74	81
8 GER	61	65	70
Angaben in % richtiger Lösungen, Werte gerundet			

Man entnimmt:

- Die Anzahl fehlerhaft markierter Aufgaben oder solcher mit einem leichten Rechenfehler, die zur TIMSS-Wertung „falsch“ führen, ist gering. In der Gesamtwertung ergeben sich für beide Länder jeweils etwa 3 % Lösungszuwachs.
- Bei weitergehender Mängeltoleranz erhöht sich der prozentuale Anteil korrekter Lösungen für beide Länder vergleichbar um noch einmal 5 bis 7 Prozent.

Der relative Leistungsabstand Schweiz zu Deutschland bleibt also bei einer gestuft mängeltoleranten Bewertung der Schülerlösungen erhalten. Somit kann der Leistungsunterschied Schweiz/Deutschland nicht durch gehäufte leichte oder gröbere Rechen- bzw. Flüchtigkeitsfehler bei deutschen Schülern erklärt werden.

4.4 Ratehäufigkeit und Rateerfolg

In der vorliegenden Arbeit kann auch der Frage nachgegangen werden, wie weit Lösungen durch Ratetechniken gewonnen wurden.

Zur Bedeutung von Ratetechniken für die TIMSS-Ergebnisse führt Lind in einem Exkurs über die dort eingesetzten Testmodellklassen aus:

„Die Planer der TIMSS griffen auf eine Modellklasse zurück, [...] die auch das sogenannte Raschmodell umfaßt. [...] Bei der Testkonstruktion zu TIMSS wurde das Raschmodell auch für Aufgaben verwendet, bei denen nur eine Auswahlantwort anzukreuzen war. Da auch ein total unwissender Proband solche „MC-Items“ durch Raten richtig beantworten kann, ist das Raschmodell für Tests mit derartigen Aufgaben eigentlich unangemessen. Die Planer der TIMSS haben es trotzdem verwendet, da sie davon ausgingen, daß die konstruierten MC-Items für die Zielpopulation leicht genug waren und daher nicht zum Raten animierten.“ (Lind, 1999, S. 349 ff)

In der vorliegenden Untersuchung wurde den Schülern nicht nur *nicht* verboten, Ergebnisse zu raten, sondern im Schülerfragebogen war explizit die Position vorgegeben „ich rate das Ergebnis, weil mir dazu kein Rechenweg einfällt“, wohl unterschieden von der Position „das Ergebnis kann man leicht erraten“ für Aufgabenstellungen, die dem Schüler so trivial oder offensichtlich erschienen, dass er dafür keine Rechnung benötigt oder einsetzen will. Somit lässt sich ein Überblick über Ratehäufigkeit und Rateerfolg gewinnen. Da Raten als Lösungstechnik zugelassen war, liegt die Vermutung nahe, dass i.w. ‚ehrliche‘ Antworten gegeben wurden.

Die nachfolgende Tabelle enthält neben den aufgeschlüsselten Ratehäufigkeiten zusätzlich eine Aufschlüsselung für die nicht bearbeiteten Aufgaben. Auch hier wurde diskriminierend nach Art der Nichtbearbeitung aufgeschlüsselt.

Man erhält für die Klassen in absoluten Zahlen und prozentual:

Tabelle 13: geratene und abgebrochene Aufgaben

Klasse	geraten		nicht bearbeitet oder abgebrochen		
	r	f	nicht verstanden	abgebrochen	nicht bearbeitet
7A GER 1	13	11	16	15	13
7C GER 1	11	16	9	12	12
8A GER 1	8	13	4	3	7
7B GER 1	6	7	24	11	26
7AB L GER 1	7	5	23	11	17
7C GER 2	15	25	6	7	24
7 KA CH 3	5	5	6	5	4
7 KE CH 3	5	1	2	2	5

Klasse	geraten		nicht bearbeitet oder abgebrochen		
	r	f	nicht verstanden	abgebrochen	nicht bearbeitet
7 G1 CH 4	5	4	6	9	5
7 G2 CH 4	2	4	4	2	2
Zahlen absolut	77	92	100	81	121
Summen in %	3,3	4,2	4,5	3,6	5,4

Die in der Schweiz zusätzlich bearbeiteten Testhefte werden aus Gründen der Gesamtdatenintegrität nicht mehr berücksichtigt. Denn es ist nicht zu verfolgen, ob dabei Aufgaben allein aus Zeitgründen nicht zu Ende geführt oder evtl. gar nicht bearbeitet wurden. Außerdem ist denkbar, dass die Zweitbearbeitungen in anderer Qualität, mit anderem Konzentrationsgrad oder anderer Motivationslage etc. erfolgte. Entsprechend werden im Folgenden auch die Daten der Klasse 8a sowie der Klasse 13 nicht weiter verfolgt. Lediglich die Lateinklasse 7AB L wird weiter als separate Klasse geführt und in die Auswertungen mit einbezogen, da hier die Schüler den Test in einer vollständig neuen Testumgebung in anderer Klassenzusammensetzung durchführten (und, wie aus Tabelle 10 zu entnehmen ist, in dieser Zusammensetzung auch etwa 4% über der Lösungsrate der Einzelklassen lagen). Daraus zu schließen, dass Schüler unter den strengen Anforderungen des Lateinunterrichtes auch in anderen Fächern bessere Leistungen erbringen, war nicht Gegenstand der Studie und kann daher hier nicht belegt werden.

Aus der Tabelle der absoluten Bearbeitungszahlen aller Klassen folgt: Von 2261 bearbeiteten Aufgaben wurden insgesamt 77 richtig und 95 falsch geraten. Dies entspricht etwa 7,5 % geratener Aufgaben mit leichtem Überhang zum falschen Raten. Dies spricht dafür, dass über alle Aufgabenbearbeitungen hinweg gesehen höchstens punktuell, nicht aber durchgängig testwiseness-Strategien eingesetzt wurden. Die relative Gleichverteilung erweckt damit Vertrauen in die Korrektheit der Angaben in den Schülerquestionnaires, die das Raten erfragten.

Differenziert man jedoch nach Schweizer und deutschen Schülern³⁶ getrennt, so ergibt sich ein deutlich anderes Bild:

³⁶ Die Zahlenunterschiede zur vorherigen Tabelle folgen aus dem Fehlen der Klasse 8 GER.

Tabelle 14: Absolute Zahl geratener und abgebrochener Aufgaben nach Ländern

Klasse	Schü- ler- zahl	geraten		nicht verstan- den	abge- bro- chen	nicht bear- beitet
		r	f			
7A GER 1	21	13	11	16	15	13
7C GER 1	21	11	16	9	12	12
7B GER 1	23	6	7	24	11	26
7AB L GER 1	21	7	5	23	11	17
7C GER 2	23	15	25	6	7	24
Summe GER	109	52	64	78	56	92
<hr/>						
7 KA CH 3	18	5	5	6	5	4
7 KE CH 3	12	5	1	2	2	5
7 G1 CH 4	16	5	4	6	9	5
7 G2 CH 4	16	2	4	4	2	2
Summe CH	62	17	14	18	18	16

Hier fällt zum einen die erhebliche Anzahl abgebrochener Aufgabenbearbeitungen bei deutschen Schülern auf: immerhin werden 226 Aufgaben (78+56+92) nicht oder nicht angemessen bearbeitet. Dies bedeutet für jeden Schüler zwei in der Bearbeitung abgelehnte Aufgaben. Bei den Schweizer Schülern beträgt die Anzahl 52 (18+18+16) abgelehnter Aufgaben bei 62 Schülern, also nicht einmal die Hälfte der deutschen Ablehnungen.

Eine aufgeschlüsselte Untersuchung der *geratenen* Aufgaben ergibt qualitativ das gleiche Bild:

Tabelle 15: Relativer Anteil geratener Aufgaben nach Ländern

	Anzahl bearbei- teter Aufgaben	% geratene Aufgaben	davon richtig geraten in %	davon falsch geraten in %
GER	1138	10,2	4,6	5,6
CH	710	4,4	2,4	2,0

Die Schweizer Schüler haben eine prozentual nur halb so große Ratehäufigkeit. Die Gleichverteilung des Rateerfolges bleibt etwa erhalten.

Man kann dies einer unterschiedlichen Qualität der Testbearbeitungen zuschreiben – umso mehr, als die Geometrieaufgaben (wie in 5.1 dargelegt) für die Schweizer Schüler im wesentlichen curriculuminvalide waren.

4.5 Bearbeitungsintensität und Bearbeitungsqualität

Neben innermathematischen Aspekten können auch Unterschiede in der Bearbeitungsintensität und -qualität für die Ranking-Unterschiede zwischen der Schweiz und Deutschland verantwortlich sein. Baumert erwähnt

„leistungsrelevante Persönlichkeitsmerkmale wie Ausdauer, Interessen oder andere motivationale Merkmale etc. (Baumert, Lehmann et al., 1997, S. 130)

Als außermathematische, aber dennoch ergebnisrelevante Aspekte könnten beobachtet werden:

- Akzeptanz der Bearbeitungsmotivation,
- Gesamtbereitschaft zur Testbearbeitung,
- Bereitschaft zur Lösung der Einzelaufgaben,
- eingebrachter Zeitaufwand vor Bearbeitungskapitulation,
- Fähigkeit zur Aktivierung mathematischen Vorwissens,
- Bereitschaft zur vollständigen Bearbeitung des Testhefts.

Solche Persönlichkeitsmerkmale oder auch andere soziale Skripte könnten zu einer länder- oder sogar klassenverbandspezifischen differenzierenden Bearbeitungsgüte führen. Damit wäre auch einsichtig, warum die von Haenisch formulierte Testkritik im Ländervergleich selektiv angewendet werden muss. Es wurde auch in der Diskussion nach Pisa (2004) vermutet, dass solche Aspekte wesentlich bedingende Faktoren für Schülererfolge bei Schulleistungsvergleichen sind.

Zur Beurteilung der Bearbeitungsintensität wurden hier indirekte Faktoren untersucht.

Ob Bereitschaft zur vollständigen Bearbeitung der Testhefte vorlag, kann etwa aus der Summe der Schlüsselwerte „geraten oder nicht bearbeitet“ gewonnen werden. Auch hier sind natürlich implizite Korrelationen zwischen Bearbeitungsintensität und Aufgabentyp sowie Schwierigkeitsgrad gemäß dem ‚international difficulty index‘ denkbar. So tritt bei den deutschen Schülern der Aufgabentyp a (algebra) markant häufig als geraten oder nicht bearbeitet auf. Bei den Schweizer Schülern gilt dies eher für den Typ g (geometry). Andererseits könnte auch die Positionierung im Testheft, obwohl hier bewußt gestreut, Einfluß auf die Bearbeitung gehabt haben. Die folgende detaillierte Tabelle dient zur Analyse solcher Beziehungen. Sie weist nach Aufgabentyp und Position im Testheft auf, oft Aufgaben *nicht* bearbeitet wurden.

Tabelle 16: Anzahl nicht bearbeiteter Aufgaben nach Aufgabentyp und Position im Testheft

Aufgabe Nr.	1	2	3	4	5a	5b	6	7	8	9a	9b	10	11	12
Typ	n	a	g	n	p	p	a	d	n	m	m	g	n	a
7A GER 1	1	2	0	1	6	2	1	4	1	1	1	1	3	5
7C GER 1	0	1	1	0	1	0	3	1	2	1	1	2	1	7
7B GER 1	1	3	0	0	11	5	4	4	0	3	2	6	3	8
7AB L GER 1	0	4	2	0	5	1	3	0	0	4	3	3	3	12
7C GER 2	0	3	1	0	2	2	3	2	1	1	4	2	1	8
Summen	2	13	4	1	25	10	14	11	4	10	11	14	11	40
7 KA CH 3	0	0	0	0	3	1	1	0	0	0	0	2	1	2
7 KE CH 3	0	0	0	0	2	0	1	2	0	0	0	0	1	1
7 G1 CH 4	0	0	1	0	3	1	0	0	0	0	1	4	0	1
7 G2 CH 4	0	0	2	0	2	2	1	1	0	4	0	0	0	0
Summen	0	0	3	0	10	4	3	3	0	4	1	6	2	4

Man erkennt, dass die letzten Aufgaben der Testhefte bei den deutschen Schülern bemerkenswert oft nicht bearbeitet wurden. Bei den Schweizer Schülern gab es keine äquivalenten Häufungen am Testende. Unter der Annahme, dass die Testhefte zumindest im ersten Bearbeitungsdurchgang linear in der durch das Testheft vorgegebenen Reihenfolge bearbeitet wurden, kann daher vermutet werden, dass die deutschen Schüler gegen Ende des Tests nicht immer daran interessiert waren, ihre Aufgabenbearbeitungen zu vervollständigen. Ein Zeitproblem dürfte hier im Vergleich zu TIMSS nicht ursächlich sein, da es ja deutsche und Schweizer Schüler gleich betroffen hätte.

Bei den Schweizer Schülern ist ein solcher Effekt der frühzeitigen Bearbeitungsaufgabe nicht erkennbar. Im Gegenteil verlangten viele Schweizer Schüler nach einem zweiten Testheft und bearbeiteten es zumeist vollständig.

Die auffällige Häufung von fehlenden Bearbeitungen zu Aufgabe 5a sowohl bei deutschen wie auch bei Schweizer Schülern ist möglicherweise korreliert mit dem Aufgabentypus. In allen Testheften war lediglich Aufgabe 5 (Typ Proportionalität) von (zufällig) hohem Schwierigkeitsgrad.

Die Vermutung, die Bearbeitung sei bei den deutschen Schülern gegen Testheftende nachlässiger als bei den Schweizer Schülern, lässt sich indirekt anhand eines weiteres Indizes testen: durch eine Bestimmung der Anzahl zusammenhängender Ketten nicht bearbeiteter Aufgaben *vom Ende der Testhefte her*. Hier gibt es bei den deutschen Klassen drei Zweierketten, fünf Dreierketten und zwei Viererketten zusammenhängend geratener, abgebrochener oder nicht bearbeiteter Aufgaben, bei den Schweizer Schülern dagegen keine einzige.

Daneben gibt es solche Ketten auch *innerhalb* der Testhefte, und zwar in Deutschland achtzehn Zweier-, elf Dreier-, zwei Vierer- und zwei Fünferketten, in der Schweiz dagegen lediglich sechs Zweierketten.

Diese Kettenlängen innerhalb der Testhefte ist ein zusätzlicher markanter Diskriminator: es entsteht der Eindruck, als hätten deutsche Schüler des öfteren Teilstabschnitte entweder einfach überschlagen oder nicht bearbeiten wollen. Möglicherweise wurden auch Aufgaben bei der Erstbearbeitung übersprungen in der Absicht, diese später nachzuholen, ohne dass dies noch zeitlich realisiert wurde.

Die nachfolgenden beiden Diagramme geben dazu einen graphischen Vergleich. Dargestellt ist die *Anzahl nicht bearbeiteter* Aufgaben. Angegeben ist für die einzelnen Klassen, wieviel Prozent der Schüler keine, eine, zwei etc. Aufgaben *nicht bearbeitet* haben.

Es ist optisch direkt zu entnehmen, dass die deutschen Schüler erheblich mehr Aufgaben *nicht bearbeitet* haben als die Schweizer Schüler.

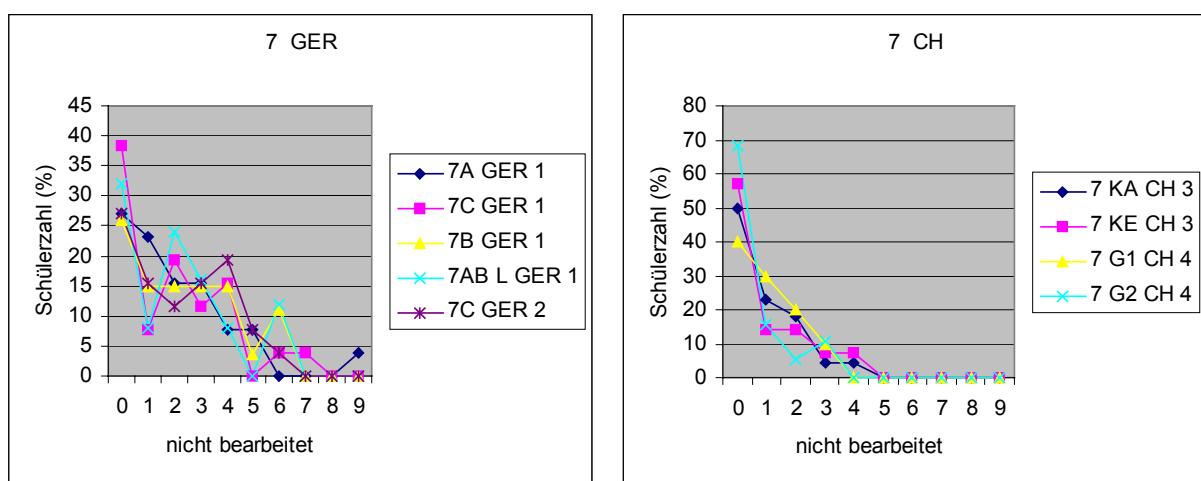


Abbildung 1: Anzahl nicht bearbeiteter Aufgaben im Vergleich GER/CH

Im Wesentlichen gilt: die Schweizer Schüler haben im Mittel eine bis zwei Aufgaben geraten bzw. nicht bearbeitet; bei den deutschen Schülern werden deutlich mehr als zwei Aufgaben geraten oder nicht bearbeitet

Dagegen geht die Anzahl insgesamt nicht bearbeiteter Aufgaben im Einzelfall bei den deutschen Schülern weit über den Schweizer Grenzwert von 5 nicht bearbeiteten Aufgaben entsprechend 42% hinaus bis zu einer Nichtbearbeitung von 75 %. Auch dies gibt einen Hinweis auf eine höhere Bearbeitungsintensität seitens der Schweizer Schüler.

Vergleicht man die Schweizer Schüler jedoch mit der Klasse 8A in Deutschland (oben nicht aufgeführt), so sind die Bearbeitungsintensitäten etwa gleich. Dies

könnte wiederum ein Hinweis auf curriculare Diskrepanzen sein; allerdings sind auch entwicklungspsychologische oder andere Gegebenheiten dieser Klassen als Ursache denkbar.

Weitere Informationen über die Bearbeitungsintensität lassen sich auch durch Vergleiche der jeweiligen Schülerquestionnaires mit den Lehrerangaben zur curricularen Validität und zurückliegendem Zeitraum seit unterrichtlicher Bearbeitung erhalten. Bei Diskrepanzen zwischen Lehrer- und Schülerangaben kann wohl unterstellt werden, dass die Lehrerinformation eher korrekt ist als entgegenstehende Schüleraussagen (solange es sich nur um Einzelaussagen handelt). In Zweifelsfällen können mehrere Schülerquestionnaires zur gleichen Aufgabe verglichen werden. Möglicherweise hat der Schüler dann die Analogie zu geübten Aufgaben nicht erkannt oder den Aufgabentyp und vielleicht auch oder nur die entsprechenden Lösungswägen vergessen. Denkbar ist aber auch, dass der Schüler vorschützt, der Aufgabentyp sei noch nicht behandelt worden, um so einen Grund zu finden, die Aufgabe nicht zu bearbeiten. Andere Aspekte sind zu bedenken, wenn der betreuende Lehrer die Aufgabe als unterrichtlich zwar nicht bearbeitet bezeichnet, aber seiner Ansicht nach vom Schüler gelöst werden kann.

Die sich hier öffnenden Zusammenhänge zwischen Aufgabenbearbeitungen und Lehrer- und Schülerquestionnaires bilden einen weiterführenden Problemkreis, zu dem im Rahmen dieser Arbeit lediglich Teilespekte beleuchtet werden können.

4.6 Lehrer- und Schülerquestionnaires

4.6.1 Informationsgehalt und Forschungsfragen

Mit den Lehrer- und Schülerquestionnaires stehen Zusatzinformationen zur Verfügung. Im Lehrerquestionnaire wurden folgende Kategorien erhoben:

Tabelle 17: Zusatzinformationen Lehrerquestionnaire

Bekanntheitsgrad:	<i>gut bekannt / bekannt / weniger bekannt / unbekannt</i>
Übungsintensität:	<i>gut geübt / etwas geübt / kaum geübt / nicht geübt</i>
Zeitdimension:	<i>vor kurzem (< 1 Monat) / vor einiger Zeit (1..3 Monate) / schon länger her (3..6 Monate) / schon lange her (> 6 Monate)</i>
Lösungserwartung, falls unbekannt:	<i>auch nicht mit Nachdenken lösbar / mit Überlegen lösbar</i>
ggf. freie Zusatz-informationen	<i>... je nach Kommentierung ...</i>

Das Schülerquestionnaire erfasst die Aspekte

Tabelle 18: Zusatzinformationen Schülerquestionnaire

Bekanntheitsgrad:	<i>Eine Aufgabe wie diese ist für mich neu. / Ähnliche Aufgaben kenne ich aus dem Unterricht.</i>
Eigeneinschätzung:	<i>Ich konnte sie damals gut. / Ich kann sie noch gut. / Ich habe sie nie gut gekonnt.</i>
Zeitdimension:	<i>Das ist schon länger her (mehr als sechs Monate). / Das ist schon etwas her (drei bis sechs Monate). / Das haben wir vor kurzem erst gemacht (weniger als drei Monate).</i>
Ratemöglichkeit:	<i>Das Ergebnis kann man leicht erraten. / Ich rate das Ergebnis, weil mir dazu kein Rechenweg einfällt.</i>
ggf. freie Zusatz-informationen	<i>Zu der Aufgabe möchte ich noch sagen, daß ...</i>

Die Fragen waren jeweils so formuliert, dass sie aus der Perspektive des Beantwortenden wie Feststellungen erschienen. Dieser Aspekt wurde bei den Schülerantworten dadurch verstärkt, dass die anzukreuzenden Antworten jeweils als ganze Sätze formuliert waren (wogegen bei den Lehrern Ankreuzen von Auswahlstichworten vorgegeben wurde). Durch diese Formulierung als Aussagen sollte ein gesteigertes Maß an Antwortkorrektheit bewirkt werden, da vermutlich eher aus einer Stichwort- oder Werteauswahl heraus falsch angekreuzt wird als dass ein aus Sicht des Ankreuzenden erkennbar falscher Satz markiert wird.

Im Lehrerquestionnaire wurde die Zeitdimension detaillierter abgefragt als bei den Schülern, da kaum zu erwarten ist, dass Schüler bei zeitlich zurückliegenden Aufgaben präzise zwischen ein, drei oder vier Monaten seit stofflicher Behandlung im Unterricht differenzieren, wogegen der Lehrer dies vermutlich anhand der ihm vertrauten Unterrichtssequenzen präzisieren kann.

Die Einzelauswertung der Questionnaires ergab Hinweise auf

- die Curriculumvalidität,
- den subjektiv empfundenen Schwierigkeitsgrad,
- das Zeitintervall seit Behandlung im Unterricht,
- den Bekanntheitsgrad,
- den Übungsgrad,
- das Maß, in dem sich der Schüler sicher im Aufgabentyp fühlte,
- die Einschätzung der aufgabenspezifischen Schülerfähigkeiten bei bekannten Aufgaben,
- die Bearbeitungsbereitschaft bei unbekannten Aufgaben,
- die Bewertung der Lösungsmöglichkeit bei unbekannten Aufgaben.

Aus der Verknüpfung der Daten aus beiden Questionnaires erhält man weit reichende Zusatzinformationen. Durch eine Analyse der jeweils aufgabenbezogene-

nen Schüler- und Lehrerangaben lassen sich unterschiedliche Einschätzungen aufdecken und damit Informationen dazu gewinnen, wie weit Planung, Durchführung und Ergebnisse von Unterricht miteinander konsistent sind. So mag es durchaus sein, dass Schüler und Lehrer den Schwierigkeitsgrad einer Aufgabe deutlich unterschiedlich wahrnehmen. Denkbar ist auch, dass aus Sicht des unterrichtenden Lehrers bekannter und behandelter Stoff für die Schüler aus anderem Blickwinkel als neue Problematik erscheint.

Aus der Gegenüberstellung der Lehrerinformationen mit dem Durchschnitt der Schülerinformationen entsteht ein Kontrollinstrument zur Einschätzung des Wahrheitsgehalts der Schülerquestionnaires. Intuitiv erwartet man, dass Lehrer- und Schülerangaben zu einer Aufgabe ähnlich sind. Ist dagegen die Lehrerangabe konträr zu dem Durchschnitt der Schülerangaben, so sollten Gründe für diese unterschiedlichen Einschätzungen existieren, je nachdem, ob die Diskrepanz zum Zeitintervall, zum Schwierigkeitsgrad oder zur Unterrichtsvalidität besteht.

Durch Verknüpfung der Questionnaires mit den tatsächlichen Schülerleistungen eröffnet sich ein weites Feld von länderspezifisch analysierbaren Fragestellungen.

Solche Analysen können weitere Hinweise auf soziale Skripte ergeben. Zusätzlich lassen sich auch durch länderinternen Vergleich der Lehrerquestionnaires Informationen über die Curriculumähnlichkeit bei verschiedenen Schulen und Lehrern gewinnen. Werden curriculare Normen nur im großen Zeitraster gesetzt, so ist die Bedeutung der Curriculumvalidität in Vergleichstests zu überdenken. Dies ist hier speziell für die Schweiz wichtig, wo der Schwerpunkt der Geometrie in die Klasse 8 verlegt wurde.

In solchen Fällen wäre es aufschlussreich, den Lösungserfolg der Schüler genau in den Aufgaben, die *nicht* im Unterricht besprochen wurden, eingehend zu analysieren und ggf. länderspezifisch zu vergleichen.

Nicht zuletzt ist es prinzipiell denkbar, aus den Angaben zu den ‚content categories‘ teilweise das realisierte Curriculum der letzten Monate zu rekonstruieren.

Bei unbekannten Aufgaben sind aus der relativen Häufigkeit der beiden Positionen ‚unbekannt und auch nicht durch Überlegen lösbar‘ und ‚unbekannt, aber mit Überlegen lösbar‘ Rückschlüsse darauf möglich, wie der Lehrer die Leistungsfähigkeit seiner Schüler einschätzt. Hier könnte auch der Frage nachgegangen werden, ob es Korrelationen zwischen solchen Einschätzungen und den tatsächlich erbrachten Schülerleistungen gibt. Dies könnte content-spezifisch nach dem ‚Pygmalion-in-the-Classroom-Effekt‘ sein:

„Can a label, such as ‚gifted‘, have a measureable effect on a person's behavior even when the label is applied randomly? Can it even change what is often as-

sumed to be an inborn characteristic such as intelligence?“ (Rosenthal, Jacobson, 1968)

Alle angedeuteten Fragestellungen erfordern eine Rasteruntersuchung der Daten, um aus den Testheften die jeweils in Bezug auf die spezielle Fragestellung interessanten Aufgabenbearbeitungen heraus zu präparieren. Dies impliziert eine dazu geeignete Datenstruktur, also die Entwicklung eines Schlüssel- oder Codesystems auch für die Begleitquestionnaires. Zur bequemeren Auswertung sollten die Schlüssel dezimal, ordinal, möglichst analog (Gleichheit ist aufgrund der leicht unterschiedlichen Begleitfragen nicht erzielbar) und leicht handhabbar sein, sowohl bei der Kodierung als auch bei der anschließenden Auswertung.

4.6.2 Synchrone Kodierung

Wie bei jeder Kodierung umfangreicher „sets“ von Itemantworten empfiehlt es sich, die enthaltenen Informationen durch differenzierende Codes zu erfassen, um den (auch in dieser Studie) aufwendigen Kodievorgang möglichst nur einmal zu durchlaufen.

Im Lehrerquestionnaire sind die Kategorien polytom³⁷ und geordnet. Die geordneten Alternativantworten der jeweiligen Kategorie schließen einander aus, so dass eine eindeutige Zahlkodierung möglich ist. Eine übliche Kodierung mit ganzzahligen aufsteigenden Werten könnte eine Äquidistanz suggerieren, die jedoch nicht vorliegt. Zusätzlich sollte Null für „keine Angabe“ als Indiz für die Bearbeitungsgüte möglich sein. Der geplante Aufgabeninformationsvektor sollte zudem bequem handhabbar sein. Daher bot es sich an, den Ziffern einer Zahl entsprechende Informationsgehalte zuzuordnen, wobei eine Ziffer redundant für mögliche Zusatzinformationen aus den freien Antworten vorgehalten wurde. Unter Beachtung des Prinzips der gleichsinnigen Polung auch zum Zwecke der leichteren Kodierarbeit und der anschließenden Lesbarkeit wurde gewählt:

Tabelle 19: Komponenten des Kodervektors – Lehrerquestionnaires

Bekanntheitsgrad und Lösungserwartung	Übungsintensität
1000 unbekannt und auch nicht durch Nachdenken lösbar	100 nicht geübt
2000 unbekannt, aber mit Überlegen lösbar	200 kaum geübt
3000 weniger bekannt	300 etwas geübt
4000 bekannt	
5000 gut bekannt	400 gut geübt

³⁷ Zur Unterscheidung von dichotom (zweigeteilt), polytom (mehrgeteilt), ungeordnete oder geordnete Kategorien sowie zur Kodierung der Antwortvariablen vgl. auch Rost, J., 1996, S. 89 ff.

Zeitdimension		Zusatzinformationen
10 schon lange her	> 6 Monate	0 Es gibt keinen freien Zusatztext
20 schon länger her	3..6 Monate	
30 vor einiger Zeit	1..3 Monate	
40 vor kurzem	< 1 Monat	9 Es gibt freien Zusatztext
Der Kodiervektor ergibt sich als Summe der einzelnen Zahlen in den Bereichen.		

Die Positionsauswahl folgte der äußereren Form des Abfragekastens so, dass bei der Kodierung sequentiell vorgegangen werden konnte.

Die so gewählte Kodierung lässt eine Aufgabenbeurteilung schon anhand des Zahlenwertes zu. Es entsprechen „große“ Ziffern je Stelle (ausgenommen Einerziffer) wie etwa 5440 einer Erwartung von „Item-Leichtigkeit“, hier z.B. ‘*gut bekannt und vor kurzem gut geübt*’.

Einige Kodierungsmöglichkeiten sind in Kombination redundant bzw. ggf. nicht verträglich: wenn eine Aufgabe unbekannt ist, so kann sie zumindest nicht gut bekannt oder geübt sein³⁸. Die Kodierung unterscheidet aber dennoch zwischen einer 1 als Hunderterziffer für „nicht geübt“ und einer 0 als Hunderterziffer für „keine Angabe“. Bei anderen Kombinationen wird analog verfahren.

Auch scheinbar widersprüchliche Angaben wie der Code 2220 („*unbekannt, aber mit Überlegen lösbar, kaum geübt, vor 3..6 Monaten*“) können im Einzelfall Sinn ergeben, wenn z.B. nicht ganz dem Typus entsprechende, aber dennoch etwa analoge Aufgaben kurz behandelt wurden. Ähnlich mag auch eine Angabe wie „*bekannt, aber nicht geübt*“ durchaus berechtigt sein. Solche Kodierungen können daher nicht verworfen werden.

Für den möglichen Zusatztext wurde keine Kodierung eingebracht, da er als freie Formulierung nicht vorhersehbar ist und somit vorab keine klassifizierende Kategorien bildbar sind. Da er aber voraussichtlich wesentliche Informationen zur didaktischen Validität bezogen auf die konkrete Klasse enthält, wurden solche Zusatzinformationen separat textlich erfasst und teilweise transskribiert.

Die Kodierung der Schülerquestionnaires orientiert sich an der Lehrerquestionnaires. Eine absolute Äquivalenz ist aufgrund der differierenden Fragekategorien nicht möglich. Zusätzlich sind Mehrfachantworten möglich, die nicht leicht in ein dezimal kodierendes Schlüsselsystem einzuordnen sind. Eine alternative Kodierung über Zweierpotenzen, wobei Mehrfachnennungen durch Addition der Schlüsselziffern erfasst werden könnten, schien hier zu aufwendig und überdies wenig erklärend.

³⁸ Es wurden allerdings in sich widersprüchliche Kodierungsequenzen gefunden.

Tabelle 20: Komponenten des Kodiergektors – Schülerquestionnaires

Bekanntheitsgrad und Lösungserwartung	Übungsintensität
1000 Eine Aufgabe wie diese ist für mich neu.	100 Ich habe sie nie gut gekonnt
2000 unbekannt, aber mit Überlegen lösbar	200 Ich konnte sie damals gut.
3000 Ähnliche Aufgaben kenne ich aus dem Unterricht	300 Ich kann sie noch gut.
Zeitdimension	Zusatzinformationen
10 schon länger her > 6 Monate	0: Es gibt keinen freien Zusatztext
20 schon etwas her 3..6 Monate	1 Ich rate das Ergebnis, weil mir dazu kein Rechenweg einfällt.
30 vor kurzem < 3 Monate	2 Das Ergebnis kann man leicht erraten
	3 = 1 und 2
	9 Es gibt freien Zusatztext
Der Kodiergektor entsteht als Summe der einzelnen Kodeziffern.	

Für interaktiv während der Kodierung erfasste Details, die eine Erweiterung des Schlüsselsystems erfordern, stehen somit noch weitere Einerziffern zur Verfügung. Sie wurden bei Bedarf tabellarisch ergänzt.

Die Übertragung in die elektronische Datentabelle erfolgte mit Fehlerkontrolle analog zum Verfahren bei der Kodierung der Testheft-Bearbeitungen.

Transskripte der Lehrerquestionnaires, beispielhafter Schülerlösungswege und Schülerkommentierungen finden sich in Anhang 10.6 bis 10.8.

4.7 Zu den Lehrerquestionnaires

Die Lehrerfragebögen verlangten von den Lehrern Beurteilungen und Kommentare zu *allen* gestellten Aufgaben. Daher wurde ein erstes Augenmerk auf die Präferenzen bei der Bearbeitung gerichtet. Sie lassen sich allerdings nicht quantifizieren. Dies verbietet schon die geringe Fallzahl. Dennoch sind sehr wohl graduelle Unterschiede in der Bearbeitung bemerkbar.

Dagegen lassen sich die Angaben zur Curriculumvalidität im Ländervergleich quantitativ auswerten; es sind so auch Aussagen zur Lösungshäufigkeit gelöster curriculumvalider Aufgaben möglich.

Zusätzlich lassen sich gesetzte curriculare Schwerpunkte – ebenfalls im Ländervergleich – aufweisen.

4.7.1 Präferenzen bei der Kommentierung

Bei der Kodierung der Questionnaires ließ sich eine deutlich differenzierte Art der Beantwortung der Fragen erkennen. Der Lehrer der Klasse 7A G1 GER gab im Wesentlichen nur den Bekanntheitsgrad an, nicht aber Übungsintensität und Zeitdimension (erkennbar an den meist fehlenden Hunderter- und Zehnerziffern). Dafür kommentierte er immerhin 13 Aufgaben, wobei sich die Hälfte der Kommentare darauf bezog, warum die Aufgaben von den Schülern nicht zu lösen seien.

Es ist beachtenswert, dass er kein einziges Mal den Bekanntheitsgrad ‚gut bekannt‘ wählte.

Aus den fehlenden Kodierungen zur Übungsintensität könnte man vermuten, dass hier Bekanntheitsgrad und Übungsintensität als gleichartige Informationen angesehen wurden etwa in der Art, dass eine gut bekannte Aufgabe nur dadurch gut bekannt sein könne, dass sie auch gut geübt wurde.³⁹

Der Lehrer der Klasse 7C G2 GER beantwortete detailliert alle Fragen, gab aber keinen einzigen Zusatzkommentar. Es lässt sich daher nicht entscheiden, ob er diese nicht für nötig erachtete, weil ihm die Aufgaben als angemessen erschienen, oder ob dies aus Zeit- oder anderen Gründen geschah.

Der Lehrer der Klasse 7 KA CH markierte alle Abfragedimensionen und nahm zusätzlich ausführlich (insgesamt 14 mal) zu den Aufgaben Stellung. Dabei bezogen sich 10 Kommentare auf seines Erachtens nicht lösbarer Aufgaben sowie die Gründe dafür.

Die Lehrer der Klassen 7 KE CH und 7 G1,2 CH kodierten alle Abfragen, nahmen aber kaum zusätzlich Stellung. Sie begründeten im Wesentlichen je zweimal knapp, dass Begriffe noch nicht eingeführt seien.

Zusammenfassend ist festzuhalten, dass in die Bearbeitung identischer Questionnaires offensichtlich Teile der einzelnen Lehrerprofile eingespiegelt sind. Dies wurde beim Design der Lehrerfragebögen in so starkem Maße nicht erwartet: beim Entwurf der Fragekategorien und beim Design der Abstufungen innerhalb der Kategorien war ein eher einheitliches Antwortverhalten vermutet worden. Eine länderspezifisch charakteristische Haltung ist allerdings schon auf Grund der viel zu kleinen Fallzahl nicht extrahierbar.

³⁹ Dass diese Ansicht zu überdenken ist, zeigt folgende bemerkenswerte Kommentierung einer Schweizer Schülerin zu Aufgabe V04: „Hierzu möchte ich noch sagen, dass wir es vor kurzer Zeit gelernt haben. Wir haben aber noch nicht geübt.“

Bei der Kodierung der Lehrerkommentare lassen sich zum Teil Gehalte erfahren, die Hinweise auf sonst unbeachtete Differenzierungen liefern. Als Beispiel sei die Kodierung 1000 des Lehrers der beiden Klassen 7 Ch4 G1 und 7 CH 4 G2 bei Aufgabe E.O05 genannt, verglichen mit den gleichwertigen Kodierungen 1000 und 1009 bei den Aufgaben F05 und F08. In eine statistische Auswertung geht lediglich der Gehalt „unbekannt und auch durch Nachdenken nicht lösbar“ ein, obwohl hier unterschiedliche Qualitäten vorliegen: Aufgabe F05 kann auch durch Nachdenken nicht gelöst werden, weil der Begriff „Schwerlinie“ unbekannt ist. Bei Aufgabe F08 wurde das doppelte Minuszeichen noch nicht behandelt.

Hier liegen also Sachargumente für Unlösbarkeit vor, wohingegen bei Aufgabe E.O05

O05. Jede der sechs Seiten eines Würfels ist entweder rot oder blau angemalt.

Die Wahrscheinlichkeit, beim Würfeln oben eine rote Seite zu haben, ist $\frac{2}{3}$.

Wieviel Seiten sind rot?

dafür kein Sachargument vorgebracht wird. Dies erstaunt um so mehr, als der Lehrer ansonsten seinen Schülern zutraut, alle unbekannten Aufgaben mit Überlegen zu lösen.

4.7.2 Curriculumvalidität aus Lehrerperspektive

Eingangs wurde im Rahmen der TIMSS-Testkritik die Curriculumvalidität diskutiert. Es sei erinnert an die Zitate „The Netherlands concluded that TIMSS did *not* fit in a satisfying way to the intended / perceived / implemented curriculum ...“ und „that there exist not *one* curriculum in the US, but thousands.“ Weiter wurde darauf verwiesen, dass für die Schweiz eine erheblich größere Anzahl von textbooks und curriculum guides analysiert wurde als für Deutschland.

Trotz der geringen Fallzahl wurde versucht, aus den Fragebögen Aussagen über Abweichungen von der Curriculumvalidität zu extrahieren. Für die Beurteilung aus Lehrersicht erhält man:

Tabelle 21: Abweichungen von der Curriculumvalidität nach Lehrerquestionnaires

	7A G1 GER	7C G2 GER	7 KA CH	7 KE CH	7 G1,2 CH
3000 - 2000	37 %	40 %	50 %	43 %	42 %
2000 - 1000	22 %	35 %	37 %	32 %	37 %
1000	5 %	0 %	0 %	3 %	2 %
Legende:	3000: weniger bekannt 2000 unbekannt, aber mit Überlegen lösbar 1000: unbekannt und auch nicht mit Überlegen lösbar				

- Für die Schweiz und Deutschland fallen etwas unter 40 % der Aufgaben in die Kategorien „weniger bekannt“ oder „unbekannt, aber mit Überlegen lösbar“.
- Dabei wird bis auf eine Abweichung (22%) zusätzlich stabil angegeben, etwa 35% der Aufgaben seien tatsächlich unbekannt.

Die für die Schweizer Schüler auch nicht mit Nachdenken lösbar Aufgaben dürften unter anderem auf fehlendes Geometriewissen zurückzuführen sein.

Ein länderspezifisch anderes Bild ergibt sich aus einem Vergleich zwischen curriculumvaliden bekannten Aufgaben und ihrer Lösungshäufigkeit:

Tabelle 22: Ländervergleich gelöster curriculumvalider Aufgaben

	7A G1 GER	7C G2 GER	7 KA CH	7 KE CH	7 G1 CH	7 G2 CH
curriculum- valide Aufgaben	78 %	65 %	63 %	68 %	63 %	63 %
davon bekannt oder gut bekannt	51 %	49 %	42 %	47 %	49 %	49 %
insgesamt wurden gelöst	54 %	55 %	73 %	73 %	66 %	75 %

- Bei etwa gleicher Curriculumvalidität und ebenso gleichem Bekanntheitsgrad „bekannt“ oder „gut bekannt“ liegt die Lösungsrate für die deutschen Schüler bei 54%, die der Schweizer Schüler bei über 70% und somit erheblich über dem Anteil bekannter Aufgaben und auch noch deutlich über dem Anteil curriculumvalider Aufgaben. Dies ist ein deutliches Indiz für eine hohe Bearbeitungsbereitschaft. Die deutschen Schüler lösen im wesentlichen die bekannten Aufgaben, die Schweizer Schüler lösen zusätzlich 20% bis 30% von Aufgaben, die ihnen weniger oder nicht bekannt sind.
- Bei den deutschen Schülern ist der Anteil gelöster Aufgaben etwa gleich dem der bekannten Aufgaben. Daraus darf allerdings nicht gefolgert werden, dass die deutschen Schüler nur bekannte Aufgaben lösen konnten, da „gelöste Aufgaben“ auch nicht curriculumvalide und dennoch gelöste Aufgaben sein können.
- Das Ergebnis relativiert das Gewicht des eingangs zitierten Einwandes, für Deutschland sei die Curriculumvalidität der TIMSS-Aufgaben nicht hinreichend geprüft worden. Anscheinend ist dies für die Schweizer Schüler eine weniger zwingende Voraussetzung für erfolgreiche Aufgabenbearbeitungen.

Es existieren zwei naheliegende erklärende Interpretationen: denkbar ist, dass die Curriculumvalidität nach Angaben des Lehrers sich deutlich vom Maß des Wiedererkennens bei den Schülern unterscheidet. Denkbar ist aber auch, dass

extracurriculare Gründe mit für die Testleistungen zeichnen. Dies könnten soziokulturelle Faktoren wie ein unterschiedliches Maß an Bereitschaft zur intensiven Beschäftigung mit den Aufgaben oder auch lerntechnische Gründe wie unterschiedliche Ausprägungen des Langzeitgedächtnisses sein. Einen Hinweis auf solche extracurricularen Bedingungsfaktoren gab es bereits im Abschnitt zu Bearbeitungsintensität und Bearbeitungsqualität.

4.7.3 Curriculare Schwerpunkte im Ländervergleich

Die Lehrerquestionnaires gestatten auf einfache Weise eine Übersicht über die realisierten curricularen Schwerpunkte.

Zur Analyse auf möglicherweise erkennbare Unterrichtsschwerpunkte wurden die Daten zusammengefaßt in den Gruppen gut bekannt/bekannt, weniger bekannt und nicht bekannt. Danach wurden sie nach den bekannten TIMSS-Inhaltskategorien sortiert

Es ergab sich zu den einzelnen Kategorien:

- Der Bereich $n = \text{Zahlen und Zahlenverständnis}$ hat überall eine sehr hohe curriculare Validität (etwa 88 %).
- Der Bereich $a = \text{Algebra}$ ist länderunspezifisch und im Grad gestreut unbekannt.
- Der Bereich $g = \text{Geometrie}$ ist in der Schweiz zum Testzeitpunkt wenig behandelt. Bei den Klassen 7G1 und 7G2 ist der hohe Prozentsatz von 83 % unbekannter Aufgaben dadurch begründet, dass die Geometrie als Schwerpunktthema in Klasse 8 gesetzt wurde.
- Der Bereich $m = \text{Messen und Maßeinheiten}$ hat nur in der Klasse 7 KE einen sehr hohen Bekanntheitsgrad, sonst sind die Bekanntheitsgrade länderunspezifisch schwankend.
- Der Bereich $d = \text{Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung}$ ist nicht curriculumvalide; lediglich Klasse 7C weist einen Bekanntheitsgrad von 100 % gut bekannt auf. Man könnte auf eine besondere Vorliebe des Lehrers oder sonstige klassenspezifische Umstände wie möglicherweise Sonderprojekte zu diesem Themenkreis schließen.
- Der Bereich $p = \text{Proportionalität}$ ist wie bei d im wesentlichen curriculuminvalid.

In toto gilt: lediglich der Bereich Geometrie ist länderspezifisch prägnant unterschiedlich vertreten; alle anderen Aufgabentypen sind aus Sicht der beurteilenden Lehrer mehr oder weniger regellos streuend curriculumvalide oder invalide.

4.8 Zu den Schülerquestionnaires

Eine Vorab-Sichtung der Schülerfragebögen (einige Transskripte finden sich im Anhang 10.7) ergab bezüglich der Bearbeitungsqualität ein diffuses Bild von teilweise unklaren oder in sich widersprüchlichen Angaben zu Bekanntheitsgrad und Curriculumvalidität. So wurden bei angeblich unbekannten Aufgaben Angaben zur Zeitdimension und zur Übungsintensität gemacht oder auch in der Art eines ‚persönlichen Ausfüllstils‘ mehrere Aufgaben mit gleichen Angaben versehen. Von einer intensiven didaktischen Analyse wurde daher abgesehen.

Die ‚freien‘ Schülerkommentare versprechen dagegen zusätzliche Einsicht in die Einstellung der Schüler zum Test, ihre Einstellung zum aktuellen Problem der Aufgabe sowie allgemeineren Prinzipien von Zustimmung oder Ablehnung des Testverfahrens. Daher werden beispielhaft die Transskripte zu den Klassen 7A GER1 und 7 G2 CH exemplarisch ausgewertet.

Eine freie Kommentierung ist nicht durchgehend zu erwarten. Eine erster klassenbezogener Überblick über die länderspezifische Anzahl freier Kommentierungen ergibt:

Tabelle 23: Anzahl freier Kommentierungen

Klasse:	7A GER 1	7 G2 CH
Gesamtzahl ausgeteilter Aufgaben	312	228
Anzahl freier Zusatzkommentierungen	60	54
in Prozent (davon weiblich/männlich)	19 (88/12)	22 (57/43)

Der Anteil kommentierter Aufgaben ist in beiden Klassen etwa gleich. Auffallend ist dagegen, dass bei den deutschen Schülern fast nur die Mädchen Zusatzkommentare geben, wogegen die Kommentare der Schweizer Schüler etwa gleich auf Mädchen und Jungen verteilt sind. Eine Ursache dafür ist aus den Bearbeitungen nicht erkennbar. Möglicherweise gibt es in den beiden Ländern geschlechtsspezifisch unterschiedliche Bearbeitungsqualitäten in dem Sinne, dass die deutschen Jungen weniger intensiv mit dem Test umgehen als die deutschen Mädchen.

Die Schülerkommentare lassen sich inhaltlich grob aufteilen nach Aussagen zu

- Curriculumvalidität aus Schülersicht,
- Zuordnung des vermuteten „Unterrichtsstoffes“ zu Klassenstufen,
- Hinweise auf unbekannte Fachbegriffe,
- Einschätzung des Bekanntheitsgrades,
- Einschätzung des Schwierigkeitsgrades,
- Entschuldigungen für nicht erbrachte Leistungen,
- Umfeld der Leistungserbringung,
- Schilderung des Rechenganges.

Bei den deutschen Schülern lassen sich deutlich mehr als die Hälfte der Kommentierungen der Kategorie ‚Entschuldigung für nicht erbrachte Leistungen‘ zuordnen. Als Gründe dafür werden (stichwortartig transkribiert) etwa angegeben:

Ich kann diese Aufgabe nicht lösen, da ...

- unbekannt
- zu lange her, wußte nicht mehr, nach einiger Zeit kann man es nicht mehr so gut, vergessen
- noch nicht gehabt und daher nicht in der Lage, neu, aber total
- Stoff einer höheren Klasse
- unterschiedliche Mathematikbücher
- finde sie zu schwer, schon schwieriger, ist mir kein Begriff
- komisch formuliert, unverständlich, zu ungenau erklärt, verstehe die Aufgabenstellung nicht
- kein Rechenweg eingefallen
- hatte Probleme zu kürzen

Lediglich die letzten beiden Formulierungen suchen die Schuld beim Bearbeiter. Insgesamt lässt sich jedoch in diesen Kommentierungen eine schuldabweisende Haltung erkennen.

Die sonstigen Bemerkungen beziehen sich im wesentlichen darauf, dass Aufgaben (ganz) leicht waren oder dass ein Fachterminus (Schwerlinie) noch unbekannt sei. Bemerkenswert ist allerdings noch die Formulierung „... dass man Konzentration braucht, um sich das vorzustellen, und in einer lauten Klasse ist das nicht einfach“. Hier werden die Eindrücke bestätigt, die unter 3.4 zum Testverlauf protokolliert wurden.

Bei den Kommentierungen der Schweizer Schüler finden sich grundlegend anders gewichtete kategoriale Schwerpunkte. Zwar gibt es auch hier mehrfach Bemerkungen wie „ich weiss nicht (mehr), wie das geht“ oder „mir ist kein Rechenweg eingefallen“, aber insgesamt beschäftigen sich die Bemerkungen mit dem mathematischen Gehalt der Aufgaben. Bei fehlenden Vorkenntnissen wird präzise auf den fehlenden Stoff bezogen:

Ich kann diese Aufgabe nicht lösen, da ...

- ich diese Zeichen noch nie gesehen hatte
- ich weiss wirklich nicht, was Schwerlinien sind, und ich kann mir auch nichts darunter vorstellen
- wir das Trapez noch nicht behandelt haben
- wir noch nicht gelernt haben, wie man Winkel berechnet

- wir rechneten noch nie mit Winkeln, doch man kann es durch Überlegen herausfinden
- ich solche Aufgaben eigentlich noch nicht gemacht habe
- wir noch nie solche Aufgaben gelöst haben, aber die Aufgabe war trotzdem nicht so schwer
- ich nicht mehr weiss, ob man zuerst x oder y schreibt
- man für das Ergebnis gar keinen Rechenweg benötigt
- der Rechnungsweg eigene Erfindung ist
- ich mir noch eine Häuschenreihe mehr vorstellen mussste, als ich die Figur spiegelte
- eigentlich hatte ich die Lösung schon, habe es nicht gemerkt und sie durchgestrichen

Diese Kommentierungen beinhalten sachliche Auseinandersetzungen mit den Aufgaben. Eine schuldabweisende Einstellung ist hier nicht erkennbar.

5 Detailanalysen zur Geometrie

5.1 Globale TIMSS - Leistungsprofile

Ein zentrales Forschungsziel der vorliegenden Arbeit war es, anhand eines Teilgebietes der bei TIMSS abgefragten Sachbereiche exemplarisch zu untersuchen, wie weit eine Analyse elaborierter Schülerarbeiten über die TIMSS-Bewertung hinaus gehende Informationen zu Schülerleistungen bereitstellen kann. Erst so wäre der zusätzlich zu erbringende Analysenaufwand zu rechtfertigen. Es bleibt hier offen, wer in einer groß angelegten Untersuchung diese Detailanalysen übernehmen sollte, wie eine Harmonisierung der Interpretationsleistungen der Korrektoren erfolgen könnte, wie solche Interpretationen wissenschaftlich abgesichert werden könnten und wie solche Mikro-Interpretationen von Schülerarbeiten rückkoppelnd auf die Curriculumentwicklung einwirken können.

Baumert et al. stellen die nachfolgend in Auszügen abgedruckte Tabelle der erzielten nationalen Leistungsprofile in den mathematischen Sachgebieten zur Verfügung.

Tabelle 24: Nationale Leistungsprofile nach mathematischen Sachgebieten

Nation	Zahlen und Zahlenverständnis	Geometrie	Algebra	Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung	Messen und Maßeinheiten	Proportionalität
Schweiz (deutschspr.) ^{1,3}						
Korea	0	+	0	0	0	0
Japan	-	+	+	-	-	-
...						
Schweiz (deutsschpr.) ^(1,3)						
Hongkong	0	+	+	-	0	+
Schweden	+	-	-	+	+	0
Schweiz (insgesamt)	+	-	-	+	+	0
Niederlande	0	0	-	+	0	0
Frankreich	0	+	-	+	0	-
...						
Deutschland ^{1,2,3}	+	-	-	+	+	0
England ³	-	0	-	+	0	0
USA ²	+	-	0	+	-	0
...						
Spanien	-	-	+	0	-	0
Südafrika	0	0	+	-	-	+

(signifikante Abweichungen vom nationalen Gesamtmittelwert, „0“ = keine Abweichung,
+ = „positive Abweichung“, „– = negative Abweichung

In der Schweiz und der russischen Föderation wurde je nach Landesteil der 7. oder 8. Jahrgang, in England, Schottland und Kuwait der 9. Jahrgang und in Australien sowie Neuseeland der 8. oder 9. Jahrgang untersucht.

¹ Die 8. Jahrgangsstufe entspricht aufgrund des höheren Alters nicht den internationalen Stichprobenvorgaben.

² Internationale Stichprobenvorgaben nur durch Ziehung von Ersatzschulen erreicht.

³ Internationale Stichprobenvorgaben nicht erreicht. (Baumert et al., 1997, pp107, 108)

Unter den ‚zentralen TIMSS- Ergebnissen im Überblick‘ resümieren Baumert et al.:

„Die relativen Leistungsstärken der deutschen Schülerinnen und Schüler der achten Jahrgangsstufe liegen in der Arithmetik, im Umgang mit Maßeinheiten und in der deskriptiven Statistik. Die relativen Schwächen liegen in den mathematischen Kernbereichen Algebra und Geometrie.“ (Baumert et al., 1997, S. 57). ... „In den Ländern der Spitzengruppe dagegen liegen die relativen Leistungsstärken überwiegend in den mathematischen Kerngebieten.“ (Baumert et al., 1997, S.106)

Zusätzlich betonen sie:

„Obwohl der Schwerpunkt des Geometrie-Unterrichtes in der Mittelstufe international auf der 8. Jahrgangsstufe liegt, bereiten Geometrie-Aufgaben auch einfacher Art Schülern am Ende dieser Jahrgangsstufe noch generell Schwierigkeiten.“ (Baumert et al., 1997, S.71)

Da die Schweizer Schüler – damit übereinstimmend und wie oben bereits erwähnt – erst wenig oder keinen Geometrie-Unterricht hatten, bot sich dieser TIMSS-Sachbereich eben wegen der Curriculuminvalidität und der dennoch insgesamt guten Leistungen für eine eingehende Detailanalyse von Schülerlösungswege an.

Für beide Länder ist denkbar, dass die Schülerleistungen bei TIMSS teilweise durch andere Effekte maskiert wurden. Dies könnte etwa durch die Art, Auswahl oder Präsentation der Aufgaben erfolgen. Möglicherweise (ver-)führen Randbedingungen unterrichtlicher oder situativer Natur zu Aufgabenbearbeitungen, die ein Ankreuzen attraktiver, aber falscher Lösungsvorgaben geradezu implizieren.⁴⁰

Nach den curricularen Vorbedingungen bestand zusätzlich die Hoffnung, damit Hinweise auf mathematikimmanente oder auch extrinsische Ursachen länderspezifischer Schülerleistungen zu erhalten.

⁴⁰ Die darüber hinaus gehende Annahme, es seien ‚naheliegend falsche‘ Lösungen in die Auswahl eingearbeitet worden, könnte den Schluß naheliegen, dass eher die Stabilität des Wissens gegen Irrwege als die Geometriekenntnis an sich getestet werden sollte. Dies wird hier nicht weiter untersucht.

Im Folgenden soll exemplarisch untersucht werden, welche Denkprozesse, Ideen oder Lösungsschritte von den Schülern bei der Bearbeitung von TIMSS-Geometrieaufgaben eingesetzt wurden. Denkbar ist, dass die Beurteilung der Geometrieleistungen nach TIMSS modifiziert werden muss.

5.2 Exemplarische Detailanalysen von Geometrieaufgaben

5.2.1 Auswahlkriterien

Die Schülerleistungen bei den Geometrieaufgaben waren nach TIMSS (TIMSS, 1998c) und im Vergleich mit der vorliegenden Erhebung wie folgt verteilt. Die Aufgaben sind sortiert nach steigendem internationalem Schwierigkeitsgrad (international difficulty index).

Tabelle 25: Schülerleistungen in der Geometrie

Anteil richtig gelöster Aufgaben (Angaben in %)						
	diff. in- dex	nach TIMSS			eigene Daten	
		GER	CH	int. aver.	GER 7	CH 7
M07	457	58	54	67	50	54
K03	478	72	79	64	73	100
J16	548	52	28	47	95	33
R10	551	40	51	47	64	92
M05	565	38	43	43	57	86
O03	581	42	30	42	45	67
I08	597	32	46	38	57	71
P09	617	35	39	35	23	44
L15	623	36	26	31	86	50
K08	639	28	25	27	70	31
						0

Aus der Tabelle und zusätzlich aus den Detailvergleichen der Lösungen ergibt sich:

- Sowohl die deutschen wie auch die Schweizer Schüler erzielen bis auf einige Ausnahmen Ergebnisse meist deutlich über dem internationalen Mittelmaß.
- Die Geometrieaufgaben werden durchgängig entweder vollständig richtig bearbeitet oder fallen unter die Kategorie falsch. Leichte und mittlere Fehler treten kaum auf.
- Dort, wo sie auftreten, sind sie alle von gleicher Kategorie. So ist bei der Aufgabe J 16 (Koordinatendarstellung eines vorgegebenen Punktes im Gitteraster) die Häufung des Fehlertyps 9b (Koordinatenangaben vertauscht)

bei den Schweizer Schülern so markant, dass hier eine didaktische Besonderheit zu vermuten ist. Wahrscheinlich fehlt einfach noch das Wissen um die vereinbarte Reihenfolge bei der Koordinatenangabe. Dies wird von einer Schülerin sogar explizit formuliert.

- Bei den deutschen Schülern häuft sich dagegen in Aufgabe I08 (...liegt einer der angegebenen Punkte auf einer vorgegebenen Geraden...) der Fehlertyp 9e (ebenfalls Vertauschung der Koordinaten). Die Aufgabe ist aber dennoch deutlich häufiger gelöst als bei TIMSS.
- Bei den als falsch kodierten Aufgaben herrscht – sofern sie nicht nach Nr. 0 bis 4 nicht bearbeitet oder lediglich falsch geraten wurden – der Fehlertyp 7 vor. Dies beinhaltet einen im Ansatz korrekten Lösungsweg, der aber aufgabenspezifisch fehlerhaft bearbeitet wird.
- Besonders interessante Aufgabenbearbeitungen konzentrieren sich trotz ihrer subjektiven Natur auf etwa 4 Aufgaben, zum Teil gleich bei deutschen und Schweizer Schülern, zum Teil länderspezifisch.
- Besonders auffällig ist hier die Aufgabe K08 (Drehspiegelung eines Dreiecks), da länderübergreifend wiederum sehr häufig der gleiche Fehlertyp auftritt. Dieser Fehler wurde allerdings durch die Aufgabenstellung geradezu induziert. Dennoch steigt die Lösungsrate gegenüber TIMSS drastisch.

Aus den Lösungshäufigkeiten ist weiter zu entnehmen:

- Bei den deutschen Schülern ist der Prozentsatz gelöster Aufgaben einmal etwas kleiner, einige Male vergleichbar, mehrfach aber auch drastisch höher als bei TIMSS.
- Ähnliches gilt für die Schweizer Schüler. Sie sind in der Regel etwas bis deutlich besser als bei TIMSS.

Die Schüler der Klasse 8 beherrschen die Aufgaben meist sicher – im Gegensatz zu TIMSS. Dass bei der Aufgabe K08 (Drehspiegelung) alle Schüler eine falsche Lösung notieren, sollte anhand der konkreten Aufgabenbearbeitung didaktisch hinterfragt werden. Es ist bezeichnend, dass diese Aufgabe insgesamt den Fehlertyp 8 (Weg/Rechnung vom Ansatz her korrekt, aber ein/mehrere grobe(r) Rechenfehler) erzeugt hat: insgesamt 19 mal wurde die Aufgaben korrekt oder fast korrekt (1 leichter Fehler) gelöst, von den genauso (neunzehn) fehlerhaften Lösungen sind bis auf die geratenen Lösungen alle vom Typ 7 (Weg/Rechnung vom Ansatz her korrekt, aber ein/mehrere grobe(r) Denkfehler, auch: falsche räumliche Drehung / als „Wissenslücke“ gekennzeichnet). Dies gibt ebenfalls deutliche Hinweise auf die Notwendigkeit einer didaktischen Hinterfragung der konkreten Schülerlösungen.

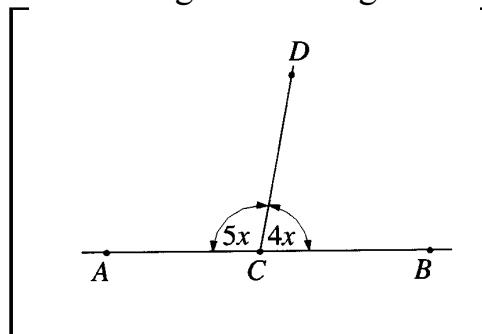
Allerdings kann auch hier nicht ausgeschlossen werden, dass die Aufgabe bewußt darauf angelegt war, die Stabilität des Schülerwissens gegen attraktive Falschaussagen zu testen.

Mit einem parallel zur Aufgabenkodierung eingesetzten Markierungsverfahren⁴¹ war es möglich, aus der Fülle der vorliegenden Schülerlösungen markante Schülerbearbeitungen herauszufiltern und wiederzufinden. Die Auswahl der nachfolgend vorgestellten Geometrieraufgaben stützt sich auch auf dieses Verfahren. Dabei fiel schon vor der vergleichenden Detailanalyse auf, dass die Schüler in vielen Fällen einen Lösungsweg wählten, der sich deutlich vom Standard-Lösungsweg abhob. Zusätzlich wurden auch Gedanken formuliert, die Einblicke auf die abgelaufenen Denkprozesse gewährten.

Die erste Aufgabe M07 wurde auch deshalb ausgewählt, weil sie bereits Gegenstand der didaktischen Forschung war (Meyerhöfer, 2000, S.439).

5.2.2 Winkelmaß (TIMSS-Aufgabe M07)

Die Aufgabe M07 lautet im englischen Originaltext



M07. In this figure AB is a straight line.

What is the measure, in degrees, of angle BCD ?

- A. 20 B. 40 C. 50 D. 80 E. 100

Sie wurde mit gleicher Skizze in das Testheft A S.20, aufgenommen.

M07. In dieser Figur ist AB eine gerade Linie.

Was ist das Maß des Winkels BCD in Grad ?

- A. 20 B. 40 C. 50 D. 80 E. 100

⁴¹ Es handelt sich im wesentlichen um ein überlagertes Zweitschlüsselsystem, mit dem gleichartige Bearbeitungen, gleichartige Fehler oder Fehlschlüsse, markante Ideen oder sonstige Besonderheiten lokalisiert werden können.

Die wörtliche Übersetzung mit ‚Was ist das Maß...‘ könnte gegenüber der Formulierung mit ‚Wieviel Grad mißt/ist ...‘ ein Verständnishindernis sein. Allerdings wurde der fachsprachlich korrekte Ausdruck ‚Maß des Winkels‘ langjährig in Deutschland in die Mathematiklehrbücher aufgenommen. Dennoch lautet die in Deutschland für TIMSS verwendete Version der Aufgabe bei gleicher Skizze:

AB ist in dieser Zeichnung eine Gerade.

Wieviel Grad mißt der Winkel BCD?

A. 20 B. 40 C. 50 D. 80 E. 100 (entnommen Meyerhöfer, 2000)

Der Datei BMIItems.pdf (online, TIMSS 1998b) ist zu dieser Aufgabe zu entnehmen:

Subject : Mathematics, Item Key: D, Content Category: Geometry

Performance Expectation: Solving Problems

International Average Percent of Students Responding Correctly

Upper Grade: 72%, Lower Grade: 67%

International Difficulty Index: 457

Setzt man Lower Grade etwa gleich mit Klasse 7, Upper Grade mit Klasse 8, so ergibt sich folgende Leistungsübersicht nach TIMSS und in der vorliegenden Studie:

- Nach TIMSS bearbeiten in Klasse 7 in TIMSS 67 Prozent der Schüler (internationaler Durchschnitt) die Aufgabe korrekt, in Klasse 8 sind es 72 Prozent.
- In der vorliegenden Studie bearbeiten die deutschen Schüler der Klasse 7 die Aufgabe mit 50% korrekt, die Schweizer Schüler mit 58 %.
- Die deutsche Kontrollgruppe der Klasse 8 erzielt (bei allerdings minimaler Stichprobengröße) etwas mehr als TIMSS (80%).

Man erkennt für die Klassen 7 eine deutlich unter dem internationalen Wert liegende Lösungshäufigkeit. Die Schüler der Klasse 8 beherrschen die Aufgabe dagegen sicher, denn 80 Prozent richtig bedeutet hier aufgrund der geringen Schülerzahl, dass lediglich ein Schüler die Aufgabe falsch angekreuzt hat. Dieser Schüler (Schüler Nr. 066, vgl. auch Transskript dazu im Anhang 10.6) hat jedoch die Aufgabe sehr wohl verstanden und richtig berechnet, dann allerdings den falschen Winkel korrekt ausgerechnet.

Die Aufgabe M07 wurde von Meyerhöfer auf ihre intentionalen Gehalte, ihre sprachlichen Besonderheiten und die Bedeutung der gewählten Formulierung für die Sinnstruktur der Aufgabe hin analysiert und auch darauf, welche einzelnen mathematischen Fähigkeiten eigentlich mit dieser Aufgabe gemessen werden sollen und erfaßt werden können:

„Die Aufgabe stellt vorrangig algebraische, nachrangig geometrische Ansprüche... Äußerlich und in der TIMSS-Statistik stellt sie aber eine Geometrie-Aufgabe dar. Das Wissen um die Größe von 180 Grad für den gestreckten Winkel ist der einzige wesentliche geometrische Aspekt der Aufgabe. Anspruch und Hürde für die Aufgabenlösung liegen einerseits im Übergang von der geometrischen Sichtweise zum algebraischen Problem und umgekehrt, andererseits im Erkennen des Aufgeteiltseins des gestreckten Winkels in neun Teile und nachfolgendem Algebraisieren“. (Meyerhöfer, 2000, S.439)

Diese Hypothesen können hier verifiziert oder falsifiziert und die tatsächlichen Lösungswege klassifiziert werden. Im Wesentlichen haben die Schüler *nicht* den vermuteten algebraischen Weg beschritten.

Es ist nicht sicher, wie Schüler den Aufgabengehalt wahrnehmen. Insbesondere ist nicht vorhersehbar, ob sie die Aufgabe als Rechenaufgabe ansehen und angehen, da sich gerade hier eine Lösung durch Abschätzen anbietet, sofern das Winkelmaß „gefühlsmäßig“ verankert ist. Weiter ist zu untersuchen, wie weit fehlerhaft angekreuzte Lösungen auf mangelnder Geometriekenntnis beruhen oder möglicherweise Sekundäreffekte zum Ankreuzen falscher Lösungen führen.

Zur Bearbeitung und Lösung der Aufgabe sind aus mathematikdidaktischer Sicht einzubringen:

- Verständnis und Interpretation des Textes so weit, dass der gesuchte Winkel aus der Bezeichnung „BCD“ identifiziert wird,
- Kenntnis vom bzw. Umgang mit dem Größenmaß von Winkeln,
- je nach unterrichtlicher Behandlung evtl. Umsetzung der mathematischen Formulierung ‚Maß des Winkels‘ in die vertrautere Formulierung ‚wie gross ist...‘,
- ggf. Schätzen von Winkelmaßen
- ggf. Umsetzung von Bezeichnungen in Gleichungen, also ein Transponieren geometrischer Sachverhalte in algebraische Beziehungen sowie
- ein Retransponieren aus der algebraisch gefundenen Lösung in den geometrischen Raum,
- ggf. eine Plausibilitätskontrolle des gefundenen Winkelmaßes.

Zur Lösung der Aufgabe sind nach Erkennen des gesuchten Winkels zwei Strategien naheliegend:

- eine intuitive Größenabschätzung, ausgehend vom Wissen um den Begriff „rechter Winkel“, Erkennen einer Winkelgröße von ‚etwas weniger‘ als 90 Grad und korrektes Ankreuzen der Lösung,
- alternativ die Berechnung über eine Gleichung, denn die Aufgabenstellung suggeriert durch die Bezeichnung der Winkel mit $4x$ und $5x$, dass zu ihrer Lösung eine Rechnung erforderlich sei.

- Allerdings ist selbst das nicht sicher: ein Schüler hat 4x einfach als Winkelnamen gedeutet. (s.u.)

Da aber die TIMSS-Aufgabe als MC-Item keinerlei explizite Rechnung verlangt, kann im Gegensatz zur Annahme von Meyerhöfer vermutet werden, dass Schüler (wertend: vernünftigerweise) den Winkel abgeschätzt haben. Dazu ist es lediglich erforderlich, ihn anhand seiner Bezeichnung mit BCD korrekt zu identifizieren und abzuschätzen, dass er etwas unter 90 Grad liegt (falls ein 90-Grad-Winkel bekannt ist und in die Zeichnung gedacht wird).

Diese Vermutung lässt sich aus den expliziten Schülerlösungen *so* nicht belegen. Die Lösungswege der Schüler sind erheblich verschlungener.

Von den 22 deutschen Schülern der Klasse 7 wird zweimal direkt formuliert „da etwas unter 90 Grad“ – und damit die Aufgabe abgeschlossen. Andere Schüler machen die gefundene Lösung zur Startvorgabe für eine von ihnen modifizierte Aufgabenstellung.

Die Schweizer Schüler verwenden zweimal die Begründung ‚leicht weniger als 90 Grad‘.

Dagegen verwenden alle Schüler der Klasse 8 Deutschland diese Begründung (viermal direkt, einmal indirekt), einmal mit direktem Bezug auf Ratetechniken: ‚zwischen 90 Grad und 50 Grad bleibt nur 80 Grad übrig‘.

Dies legt nahe, dass bei den Schülern der Klasse 7 Deutschland der Winkelbegriff noch nicht ‚gefühlsmäßig‘ gefestigt ist. Bei den Schweizer Schülern ist diese Vermutung schon deshalb wahrscheinlich, weil Geometrie noch nicht oder gerade erst beginnend im Unterricht behandelt wurde. In Klasse 8 Deutschland scheint dagegen das Gefühl für Winkelmaße gefestigt zu sein.

Verstärkt wird diese Vermutung dadurch, dass immerhin 4 deutsche Schüler den Winkel abmessen – was hier als Lösetechnik nicht verboten war. Auch diese Schüler haben die Bezeichnung mit 4x keinesfalls als Aufforderung zur Rechnung gesehen.

Ein weiterer Teil der Schüler ist entweder mit der Bezeichnung der Winkel, nicht aber mit den Winkelmaßen überfordert oder hat die Aufgabe nicht sorgfältig gelesen bzw. die Winkel aus sonstigen Gründen verwechselt: vier deutsche (3 in Klasse 7, 1 in Klasse 8) und zwei Schweizer Schüler beziehen sich auf den Nebenwinkel und geben als Lösung 100 Grad an.

Vier deutsche Schüler führen die durch die Winkelbezeichnung suggerierte Rechnung aus, beantworten dann aber eine so nicht gestellte Frage: sie geben als Lösung 20 Grad an. Die Aufgabenbearbeitungen legen folgende Deutung nahe:

die Schüler änderten die Aufgabenstellung, weil sie aus dem aktuellen Unterricht gerade Verfahren zur Lösung von linearen Gleichungen kennen. Wesentlicher Aufgabenkern dieser Unterrichtssequenzen ist der Rechenauftrag „löse nach x auf“. Vermutlich wurde diese den Schülern zur Zeit präsente Aufgabenstellung übertragen, so dass für sie nicht mehr die Winkelbestimmung vorrangiges Ziel der Aufgabe war. Ganz deutlich wird dies bei Schüler Nr. 096, der formuliert:

„20, weil ‚das‘ etwas weniger ist als die Hälfte und damit weniger als 90 Grad. Und 20 Grad mal 4 ergibt 80 Grad.“

Hier wurde also die korrekte Lösung 80 Grad direkt erkannt und als selbstverständlich mit eingearbeitet, dann aber die Aufgabenstellung in eine andere Fragestellung umgewandelt. Nebenher erkennt man hier eine Lösetechnik, die ich als ‚Lösen durch Probe‘ kennzeichnen möchte (die als Inversionstechnik zur Lösung eines Problems bekanntlich sehr mächtig sein kann).

Gestützt wird diese Deutung der bei den Schülern ablaufenden, den Problemlösungsprozess begleitenden und prägenden Mechanismen durch eine entsprechende Beobachtung im QuaSUM-Test (QuaSUM = Qualitätsuntersuchungen in Schulen im Unterricht in Mathematik) des Landes Brandenburg. Dort wurden in einer Aufgabe verschiedene lineare Graphen und ein Parabelgraph mit Scheitelpunkt S (0|3) als Lösung auf die Frage nach dem Bild zu einer Funktion $f(x)=mx+3$ angeboten (Bieber, 2000, S.9). Nach Mitteilung des Projektbetreuers wurde der Parabelgraph anscheinend deshalb typisch als falsche Lösung gewählt, weil „im Unterricht gerade Parabeln behandelt wurden und kleinere Tests auch nur wenig im Stoff zurückgehen dürfen“.⁴²

Anscheinend registrierten diese Schüler feinfühlig und in starkem Maße die Kontexte des Wissenserwerbs zum aktuellen Zeitpunkt und integrierten sie in ihr mathematisches Verständnis. Eine korrekte Lösung wird dann nicht als solche erkannt oder auch verworfen, wenn sie nicht in das gegenwärtige Unterrichtsgeschehen eingeordnet werden kann. Dies legt die Hypothese nahe, dass in der Mathematik Problemlöseprozesse zumindest von Schülern der Jahrgangsstufen 7 und 8 stark von unterrichtlichen ‚habits‘ geprägt werden. Damit würde die Bedeutung der fachdidaktischen Ausbildung von Lehrern noch einmal unterstrichen.

Für internationale Vergleichsuntersuchungen bedarf es dann mehr als einer nationalen curricularen Adaption von Testitems, um tragfähige ‚Leistungs‘ vergleiche zu erhalten.

⁴² Übernommen aus dem Vortrag von Herrn Bieber anlässlich der Tagung des GDM-Arbeitskreises ‚nationale und internationale Großuntersuchungen‘ in Kassel 16./17.6.2000.

Ein sehr bemerkenswertes Beispiel einer auf unterrichtliche Kontexte uminterpretierten Aufgabenstellung ist die Aufgabenbearbeitung von Schüler 042. Wie sich aus Transskript und Begleitquestionnaire ergibt, wird die richtig gelöste Aufgabe wie folgt kommentiert: „*4x=D ist meiner Meinung nach richtig*“.
Aufgabe ist für mich neu, ich habe den Winkel geschätzt und bin mir daher nicht sicher.“

Ganz offensichtlich wird hier die Rechenanweisung $4x$ als Winkelname interpretiert – was in den Augen eines Schülers, der gerade für ihn neue griechische Buchstaben zur Winkelbezeichnung lernen und akzeptieren musste, kaum befremdlicher erscheinen mag als eine Interpretation der Symbole als Gleichung. Ist das Weglassen des Multiplikationspunktes noch nicht eingeübt worden, so wird die Fehlinterpretation der Rechenanweisung als Name noch verständlicher. Die eigentliche Lösung erfolgt dann durch Abschätzen. D bedeutet einfach die angebotene MC-Alternative D, nämlich 80 Grad.

Hier beschäftigt sich der Schüler intensiv mit der Aufgabe, kann auch das Winkelmaß durch Abschätzen ermitteln, wird aber durch die mathematische Symbolik irritiert. Dennoch versucht er, sie in seinem Lösungsansatz anzuwenden.

Bei der Aufgabenbearbeitung war, wie bereits erwähnt, im Gegensatz zu TIMSS auch ein Messen der Winkel möglich. Davon machten die Schüler wenig Gebrauch; einmal wurde beim Messen entweder der Winkel verwechselt oder auf der falschen Skalenseite des Geodreiecks abgelesen.

Insgesamt lassen sich die Lösungswege der Schüler wie folgt klassifizieren:

- Der korrekte Wert von 80 Grad wurde in beiden Ländern prozentual etwa gleich angegeben, sei es durch Schätzen, Messen oder ohne Begründung.
- Bei den deutschen Schülern wurde viermal der Wert 20 Grad angegeben
- Für beide Länder wurde gleich häufig das Maß des Komplementwinkels angegeben.

Bei allen Aufgaben, wo als Lösung 20 Grad angegeben wurde, ist aus der elaborierten Bearbeitung ersichtlich, dass die Schüler sehr wohl die richtige Lösung 80 Grad erkannt hatten, diesen Wert aber nicht als Lösung der gestellten Aufgabe ansahen.

Einmal werden sogar beide Nebenwinkel korrekt angegeben und das Maß $80:4=20$ sowie $100:5=20$ ermittelt. Der Schüler erbringt dabei mehr, als die Aufgabe abfragen will: er erkennt die Winkel, erkennt die impliziten Rechenaufgaben und interpretiert dann die Fragestellung in eigenem Sinne um. Dass er zweimal 20 Grad als Ergebnis erhält, wird ihn vermutlich in der Ansicht über die Korrektheit seiner Lösung bestärkt haben.

Die erreichten Lösungshäufigkeiten von 45% (GER) bzw. 55% (CH) können hier also nicht als Indikatoren für Schülerfähigkeiten beim Erkennen oder Errechnen von Winkelmaßen verwendet werden.

Die oben im Zitat getroffene Feststellung, dass selbst Geometriaufgaben einfacher Art auch in Klasse 8 noch Schwierigkeiten machen, kann in dieser Studie nicht bestätigt werden. Eher gilt:

- Auch Schüler der Klasse 7 hatten wenig Probleme mit dem eigentlichen geometrischen Gehalt, aber die Art der Aufgabenpräsentation und der Ergebnisfeststellung nach TIMSS maskierten das korrekt erkannte Ergebnis.
- Man kann daher vermuten, dass sie eher mit der Aufgabenpräsentation als mit der eigentlichen Aufgabe Probleme hatten.
- Schüler der Klasse 8 lösten die Aufgabe sowohl in TIMSS als auch hier zu Prozentanteilen, die beim internationalen Durchschnitt oder sogar darüber liegen.

Zur Bewertung der geometrischen Fähigkeiten bei dieser Aufgabe bedarf es somit einer über die MC-Abfrage hinausgehenden Analyse anhand der elaboriert vorliegenden Schülerlösungswege, da die Schüler die Aufgabe anscheinend unter einem anderen Blickwinkel betrachten, als er beim Testdesign vermutet und geplant war.

Die Bedeutung der Aufgabenpräsentation ist Lehrern sicher nicht unbekannt. Somit bleibt zu fragen, ob die TIMSS-Aufgaben bewußt auf Schülerstabilität gegen ggf. aus ihrer Sicht unklare Aufgabenformulierungen zugeschnitten waren.

Als einleuchtende Alternative könnte man sich etwa bei dieser Aufgabe folgendes Procedere vorstellen: ein Mathematiklehrer kommt in die Klasse, zeichnet frei Hand einen Winkel von etwas unter 90 Grad und fragt nach der Größe des Winkels. Diese Aufgabe dürfte wohl mit einer größeren Lösungsrate als bei TIMSS beantwortet werden. Auch der umgekehrte Weg, einen Schüler aufzufordern, einen Winkel knapp unter 90 Grad zu zeichnen, ist sicherlich leicht gangbar.

Der hermeneutischen Interpretation von Meyerhöfer, die Aufgabe stelle vorrangig algebraische Ansprüche, kann nach dieser Studie nicht gefolgt werden.

5.2.3 Kongruente Dreiecke (TIMSS-Aufgabe K08)

Die Aufgabe wird von Baumert et. al. mit folgendem Text vorgestellt:

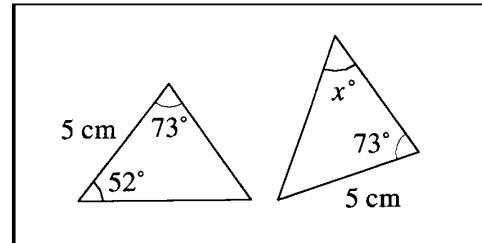
„Eines der für Achtklässler schwierigsten Geometrie-Items erfaßt das Verständnis von Eigenschaften kongruenter Dreiecke.“ (Baumert et al., 1977, S. 71)

Sie lautet:

K08. Diese Dreiecke sind kongruent. Die Maße einiger der Seiten und Winkel sind angegeben.

Welches ist der Wert von x ?

- A. 52 B. 55 C. 62
D. 73 E. 75



Nach TIMSS (TIMSS, 1998c) wurde – auch im internationalen Durchschnitt – etwa zur Hälfte die fehlerhafte Lösung A angekreuzt. Es ist offensichtlich, dass die Schüler dazu eines der beiden Dreiecke in Gedanken drehen, wie es die graphische Darstellung suggeriert. Danach braucht man das rechte Dreieck nur gedanklich anzufassen und etwa 90 Grad gegen den Uhrzeigersinn zu drehen, bis beide 73-Grad-Winkel sich überdecken. Eine gedankliche Drehung des linken Dreiecks im Uhrzeigersinn führt zum gleichen Ergebnis: x überdeckt den 52-Grad-Winkel. Bei einer evtl. auf testwiseness-Strategien beruhenden Aufgabenbearbeitung ist diese Lösung erst recht naheliegend. Die durch die Zusatzbeschriftung „5 cm“ erzwungene zusätzliche Spiegelung wurde offenbar nicht erkannt. Zusätzlich suggeriert die Bezeichnung von je zwei Winkeln in den Dreiecken mit Zahlen oder x , dass ein Zusammenhang jeweils diese beiden Winkelpaare betrifft. Es bleibt zu prüfen, ob der Fehlertyp A unter diesen testpsychologischen Randbedingungen sowie der zeitlichen Restriktion unter TIMSS als Indiz für mangelnde Geometrieleistungen aufgefasst werden kann.

Einen Hinweis dazu gibt auch die vergleichende Betrachtung der Schülerlösungen nach TIMSS in den Klassen 7 und 8 im internationalen Durchschnitt. In beiden Klassenstufen wird länderübergreifend die durch eine reine Drehung gefundene falsche Auswahl A mit etwa 45 bis 48 Prozent gewählt; die richtige Lösung B wird in Klasse 7 mit 27 %, in Klasse 8 mit 35 % Häufigkeit gefunden. Somit gibt es von Klasse 7 nach 8 kaum Zuwachs an richtigen Lösungen. Die Wahl der falschen Lösung A bleibt fast stabil, unabhängig von der Klassenstufe. Dies spricht insgesamt eher für einen die falsche Lösung A präferierenden Hintergrundprozess im Testablauf als für mangelnde Geometriekenntnisse.

Für die korrekte Lösung ist nicht nur eine Drehspiegelung zu erkennen, sondern es muss auch der Innenwinkelsatz für Dreiecke bekannt sein. Sollte dieser zum Testzeitpunkt noch nicht behandelt worden sein, was in der vorliegenden Studie bei den Schweizer Schülern so war, so gehört eine korrekte Lösung eher in den Bereich des Zufalls.

Die vorliegende Studie weist für die deutschen Schüler mit immerhin 70 % statt nur 28 % richtiger Aufgabenbearbeitungen bei TIMSS drastisch bessere Ergebnisse auf. Es scheint daher lohnend, die elaborierten Aufgabenbearbeitungen zu analysieren.

Die deutsche Schülerin in Testheft 110 verweist auf die falsche Lösung 52 Grad ausführlich mit den Worten „ich schätze 52 Grad, da sie ja kongruent sind. Und wie eben beschrieben hab‘ ich kein Geodreieck“.

Sie hat die Aufgabe erst durch Messen mit einem Geodreieck ersetzen wollen. Da ihr das Geodreieck fehlt, versucht sie, Kongruenzen zu erkennen. Sie hat dabei offensichtlich die Rotation gedanklich vollzogen.

Der Schweizer Schüler in Testheft 182 markiert das Wort kongruent und schreibt darunter „noch nie gehört“. Dann kommentiert er: „Ich glaube ich habe mal was gehört, dass man in einem Dreieck 190 Grad haben muss. Das habe ich einmal im Unterricht gehört. Aber vielleicht habe ich nicht gut aufgepasst.“ Er kreuzt aber trotz des falschen Wertes 190 Grad die korrekte Lösung an.

Hier wurde offensichtlich die Drehspiegelung erkannt, der Innenwinkelsatz war aber nicht aus dem Unterricht und auch nicht korrekt präsent. Dass dennoch die richtige Lösung angekreuzt wurde, wird einem Rechenfehler zu verdanken sein.

Die Schweizer Schülerin zu Testheft 185 notiert: „... ich nicht weiss, was kongruent heisst, doch ich habe probiert zu verstehen was es heißt.“

Sie fertigt zusätzlich eine mit Namen für die Dreieckseckpunkte versehene Skizze der Drehspiegelung und berechnet daraus den gesuchten Winkel. Diese Schülerin verfügt also offenbar über den Innenwinkelsatz, erkennt die Drehspiegelung und bemüht sich, einen unbekannten Fachterminus zu deuten.

Gleicher leistet die Schweizer Schülerin zu Testheft 219. Sie schreibt: „ich kenne das Wort kongruent nicht, und rate, was es heißen könnte. Dabei ist mir das Wort gleich eingefallen. Es ist für mich das logischste Wort bei dieser Aufgabe.“

Die Schülerin kreuzt die falsche Lösung (A) an, da sie nur eine Drehung sieht. Das eigentliche Problem ist für sie wohl gar nicht die Aufgabe, sondern die Deutung eines unbekannten Fachbegriffes. Sie stellt sich dieser Aufgabe und „löst“ sie in ihrem eigenen Vokabular. Diese Äußerung kann tiefer gedeutet wer-

den. Die Schülerin versucht, mathematische Fachbegriffe auf ihr Alltagsvokabular abzubilden. Dass ‚gleich‘ das ‚logischste‘ Wort für ‚kongruent‘ sei, könnte dahin interpretiert werden, dass sie unbekannte Fachbegriffe durch für sie ‚vernünftige‘ Begriffe ersetzen will.

Die fehlerhafte Lösung der Schülerin zu Testheft 25 (7A GER) beschreibt den oben vermuteten Denkprozess wie folgt: ‚Man muss das erste Dreieck nach rechts in die richtige Position drehen (im Kopf). So, dass der 73 Grad Winkel da ist, wo der andere ist.‘

Gleichermaßen formuliert die Schülerin zu Testheft 28 (7A GER): ‚Wenn man das rechte Dreieck schiebt, passt es genau auf das linke. Dann muss der Winkel $x=52$ Grad sein.‘

Die korrekte Lösung des Schülers in Testheft 51 (7C GER, Lösung O14) fällt aus dem Rahmen. Sie enthält den lapidaren Vermerk (wörtlich transskribiert): „Diese Ergebniss find ich am schönsten“. Daraus lässt sich zwar kein gedanklicher Lösungsweg extrahieren, aber es ist bemerkenswert, dass ein Schüler Schönheit in eine mathematische Argumentation einbringt wie einst Dirac in der Physik⁴³.

Zusammenfassend lässt sich formulieren: Die Aufgabe suggeriert eine Lösung, die nach einem gedanklichen Drehprozess erhalten werden kann. Die Schüler haben diesen abstrakten Prozess sowohl in TIMSS wie auch in der vorliegenden Studie vollzogen. Dadurch, dass das Ergebnis dieses Gedankenprozesses – kombiniert mit einem Winkelvergleich – als denkbare Lösung vorgegeben war, wurde einigen Schülern offenbar unter der TIMSS-Zeitknappheit der Blick für eine weitere Information – die Seitenbezeichnung – versperrt bzw. wurde diese Angabe nicht mit dem Dreieck assoziiert. Dies führt in TIMSS zu einer hohen Fehlerrate. Bei mehr Zeit wie in der vorliegenden Studie steigt bei den deutschen Schülern die Lösungsrate drastisch von 28 % auf 70%. Die Schweizer Schüler haben Probleme aufgrund mangelnden unterrichtlichen Sachwissens (Curriculuminvalidität bei Kongruenzbegriff und Innenwinkelsatz), befassen sich aber intensiv mit der Sachfrage und entwickeln eigene Gedanken zu noch unbekannten Begriffen. Da es aber an Sachwissen fehlt, steigt die Schweizer Lösungsrate anders als die deutsche nicht wesentlich gegenüber TIMSS (von 25% auf 31%).

Der Grund für das vollständige Versagen der Schüler der Klasse 8 bei dieser Aufgabe (0 % korrekt) bleibt unklar. Die Lösungen enthalten nur zweimal einen Hinweis auf die gedanklich nachvollzogene Drehung, die weitere Transformati-

⁴³ „Es ist viel wichtiger, Schönheit in seinen Gleichungen zu haben, als das Experiment zu beschreiben.“ (Dirac, 1963, S.47)

on wurde nicht erkannt. Die anderen Bearbeitungen sind ohne Text, so dass eine gedanklich ähnliche Aufgabenbearbeitung nur vermutet werden kann. Ähnlich wie in den Bemerkungen zur Aufgabe M07 kann festgehalten werden, dass eine fehlerhafte Lösung der Aufgabe K08 nicht als Hinweis auf mangelnde grundlegende Geometriefähigkeiten gedeutet werden kann. Die didaktische Detailanalyse weist eher intensive geometrische Denkprozesse auf, die allerdings nicht immer vollendet werden. Es erscheint fraglich, ob eine derart konzipierte Aufgabe für die Erfassung einer elementaren geometrischen ‚literacy‘ geeignet ist, wobei die Betonung auf ‚elementar‘ liegt.

Die aus den TIMSS-Ergebnissen gezogene Schlussfolgerung:

„Eines der für Achtklässler schwierigsten Geometrie-Items erfasst das Verständnis von Eigenschaften kongruenter Dreiecke.“ (Baumert et al., 1977, S. 71)

muss hier also anders formuliert werden: die Präsentation der Aufgabe scheint die Schüler unter den TIMSS-Testbedingungen dazu gelenkt zu haben, gedanklich lediglich eine Drehoperation durchzuführen. Dass die zusätzlich erforderliche Spiegelung nicht mehr erkannt wurde, lässt nicht notwendigerweise einen direkten Schluss auf fehlendes Verständnis von Eigenschaften kongruenter Dreiecke zu. Eher mag die Zeitknappheit auf diesen Irrweg geführt haben.

Es bleibt als Frage, welches Geometrieverständnis in eine grundlegenden ‚mathematical literacy‘ eingebracht werden soll.

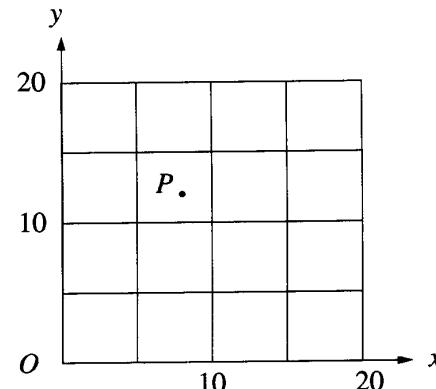
5.2.4 Koordinaten 1 (TIMSS-Aufgabe J16)

Bei der Aufgabe J16 in Testheft D fällt auf, dass acht von zwölf Schweizer Bearbeitungen den gleichen Fehlertyp aufweisen, wogegen die deutschen Schüler die Aufgabe nahezu vollständig korrekt bearbeiten. Somit ist auch diese Aufgabe lohnenswert zur didaktischen Analyse.

Die Aufgabe lautet:

J16. Welche der folgenden sind höchstwahrscheinlich die Koordinaten von Punkt P ?

- A. (8 | 12)
- B. (8 | 8)
- C. (12 | 8)
- D. (12 | 12)



Dabei ist die Lösungshäufigkeit der deutschen Schüler mit 95% erheblich besser als in TIMSS (52 %), die Schweizer Schüler dagegen erzielen vergleichbare Raten: TIMSS 28 %, hier 33 %. Beides bedarf der Erklärung.

Die Aufgabe setzt voraus bzw. verlangt:

- Grundkenntnis der Beschreibung eines Punktes in der Ebene durch Koordinaten,
- Wissen um die Reihenfolge der Koordinatenangaben und
- Interpolation innerhalb eines gegebenen Rasters.

Die Schüler können voraussetzen, dass nur eine der angegebenen Lösungen korrekt ist. Dies eröffnet die Möglichkeit zu testwiseness-Strategien, da die Lage des Punktes mit „x knapp vor zehn, y etwas größer als zehn“ beschrieben werden kann. Die erste der beiden Informationen isoliert die Antworten (A) oder (B) als möglich, die zweite selektiert dann eindeutig die Lösung (A) – immer vorausgesetzt, dass dem Schüler die Reihenfolge bei der Koordinatenangabe bekannt ist. Andernfalls entscheidet er sich für (C). Bei dieser Auswahl sind lediglich die Koordinaten vertauscht.

Implizite Voraussetzung für diesen Weg ist allerdings, dass die Schüler in der Lage sind, mit dem Auge Skalen zu interpolieren. Dies ist keine selbstverständliche Voraussetzung, wie einige der Transskripte in Tabelle 27 zeigen.

Der internationale Schwierigkeitsgrad war 548, die Lösungen waren wie folgt auf die Antwortmöglichkeiten verteilt

Tabelle 26: TIMSS-Lösungshäufigkeiten des Koordinatenproblems (1)

TIMSS-Aufgabe J 16, Angaben in %			
	Deutschland	Schweiz	internationaler Durchschnitt
A (korrekte Lösung)	52	28	47
B (falsch)	7	13	12
C (falsch, da vertauschte Koordinaten)	30	47	32
D (falsch)	2	6	6
Sonstige (keine Lösung etc.)	9	6	4

Die falschen Lösungen B und D werden kaum gewählt.

Man erkennt, dass ein deutlicher Anteil die vertauschten Koordinaten markiert hat, in der Schweiz fast jeder zweite Schüler. (In der vorliegenden Arbeit sind es sogar 8 von 12 Schülern (67 %), die die Koordinaten-Reihenfolge vertauschen.)

Zur Analyse der Schülerlösungsstrategien werden im folgenden einige der deutschen und Schweizer Bearbeitungen im direkten Transskript (ohne Korrektur von Rechtschreibfehlern) oder beschreibend wiedergegeben. Die Nummer vor dem Transskript entspricht jeweils der Testheft-Nummer.

Tabelle 27: Lösungsstrategien beim Koordinatenproblem (1)

Aufgabe J 16	
Schüler Klasse / Geschlecht / Land / richtig/falsch	Freier Kommentar
025 7A w GER r	Die Schülerin zeichnet auf der x-Achse zwischen 0 und 10 und auf der y-Achse zwischen 10 und 15 zusätzliche Skalenstriche ein.
026 7A m GER r	Man muss sich nur ungenau vorstellen, wo 8 und 12 sich kreuzen. Außerdem kann man noch nachmessen.
027 7A m GER r	kein Text
028 7A w GER r	zusätzliche Skalenstriche x: zwischen 5 und 10, y: zwischen 10 und 15, Text: Man kann es ablesen. Indem man sich kleine Hilfen macht.
051 7C m GER r	weil die 8 nahe bei der 10 und die 12 nahe bei der anderen 8
052 7C w GER r	kein Rechenweg. Kommentar: ..., daß ich dabei keine Schwierigkeiten hatte. Zu einfach
053 7C w GER r	ich sehe, dass P auf der y-Achse nahe über 10 liegt und auf der x-Achse kurz vor 10.
054 7C w GER r	... Geht man von der x-Achse aus, liegt P etwas hinter 10. Also muß die erste Zahl der Koordinaten (1.Zahl 2.Zahl) unter 10 liegen. Dann gibt es nur 2 Möglichkeiten A und B. ... Von der y-Achse aus liegt P über 10. Weil die 2. Zahl der Koordinaten bei B die 8 ist und somit unter 10 liegt, ist die Lösung A.
182 7KA m CH f	(Verbindungslien Punkt -> Koordinatenachsen eingezeichnet) Ich bin nicht ganz sicher ob 12/8 oder 8/12
183 7KA w CH f	(Verbindungslien Punkt -> Koordinatenachsen eingezeichnet)
184 7KA m CHf	Ich habe logisch überlegt, welche gar nicht in Frage kommen (z.B. B, D und A). So erhielt ich das Ergebnis.
202 7KE w CH f	(Verbindungslien Punkt -> Koordinatenachsen eingezeichnet) ::: zuerst wusste ich nicht genau was Koordinaten sind, doch dann erinnerte ich mich an den GTZ Unterricht, in dem wir Koordinaten behandelten.
203 7KE m CH f	Man liest immer zuerst auf der y-Achse ab und dann erst auf der x-Achse.
217 7G1 w CH f	Man kann eine Gerade einzeichnen und so erfährt man die Koordinaten; P liegt bei y bei 12 und bei x bei 8.
218 7G1 m CH r	Jetzt muß ich schnell überlegen welche Achse zuerst kommt. Ich glaube die x-Achse. Also ist es (8 12).

Man erkennt unmittelbar folgende Aspekte:

- Einige Schüler haben Probleme, Skalenwerte nur mit dem Auge zu interpolieren. Sie meistern die Problematik durch Einzeichnen von Skalenstrichen

(,...man kann es ablesen indem man sich kleine Hilfen macht‘) oder durch Parallelen zu den Koordinatenachsen.

- Einige Schüler setzen testwiseness-Strategien ein (Ausschluss unmöglicher Lösungen), obwohl ein direktes Ablesen sehr viel effektiver wäre. Daher ist zu vermuten, dass der Einsatz solcher Techniken eher planlos denn als eingeübte Strategie zur schnellen Lösung von MC-Aufgaben erfolgt.
- Den Schweizer Schülern ist ganz offensichtlich der Umgang mit Koordinatendarstellungen sowie die vereinbarte Koordinatenreihenfolge noch fremd. Dies wird mehrfach explizit formuliert.

Der Eindruck, Schweizer Schüler hätten Probleme mit dem Ablesen von Punktkoordinaten, erweist sich somit eher als Nebeneffekt einer punktuellen Curriculuminvalidität. Dies lässt sich auch bei TIMSS anhand des hohen Prozentsatzes von Lösungen mit vertauschten Koordinaten vermuten, die vorliegende Studie kann diese Vermutung anhand der elaborierten Schülerbearbeitungen belegen.

Als Sekundäreffekt lässt sich entnehmen, dass eine scheinbar einfache Koordinateninterpolation die Schüler vor so nicht erwartete Probleme stellen kann. Sie meistern sie durch unterschiedliche Strategien wie etwa das Einzeichnen von Skalen oder Verbindungslien.

Wiederum bleibt als Frage offen, welchen Zweck die in TIMSS vorgegebene Antwortenauswahl verfolgte. Wie bei der Analyse von Aufgabe M07 bereits formuliert könnte die MC-Vorgabe eher dem Test auf Stabilität gegen Irrwege als auf Hinterfragung geometrischer Leistungen dienen.

Es bleibt ungeklärt, zu welchem Prozentanteil die Koordinaten bei TIMSS schlicht aufgrund von Zeitknappheit verwechselt wurden.

5.2.5 Koordinaten 2 (TIMSS-Aufgabe I08)

Nach den Ausführungen des vorigen Abschnitts könnte man vermuten, dass Schüler, die Koordinatenwerte vertauschen, zwangsläufig auch bei anderen Koordinatenaufgaben scheitern. Dies ist allerdings kein zwingender Schluss, da eine vertauschte Interpretation *aller* Koordinaten, auch der der Lösungsvorgaben, zur Markierung der korrekten Lösung führen kann. Dies soll anhand einer weiteren Aufgabe untersucht werden.

Aufgabe I08 fragt den Umgang mit Punktkoordinaten auf eine komplexere Art ab als J16. Diese Aufgabe wird von Baumert ebenfalls als Beispiel einer Geo-

metrieaufgabe von höherem Schwierigkeitsgrad vorgestellt (Baumert et al., 1997, S 72). Sie lautet:

I08. Eine Gerade im Koordinatensystem geht durch die Punkte (3 | 2) und (4 | 4).

Welcher der Punkte unten liegt ebenfalls auf dieser Geraden?

- | | | |
|--------------|--------------|--------------|
| A. (1 1) | B. (2 4) | C. (5 6) |
| D. (6 3) | E. (6 5) | |

Die Lösungen waren wie folgt auf die Antwortmöglichkeiten verteilt

Tabelle 28: TIMSS-Lösungshäufigkeiten des Koordinatenproblems (2)

TIMSS-Aufgabe I 08, Angaben in %			
	Deutschland	Schweiz	internationaler Durchschnitt
A	15	9	14
B	23	26	27
C (korrekte Lösung)	32	46	38
D	3	4	6
E	5	5	6
Sonstige (keine Lösung etc.)	22	10	9

Es ist ersichtlich, dass insgesamt deutlich weniger als die Hälfte aller Schüler die Aufgabe richtig löste und dass jeweils etwa ein Viertel aller Schüler die falsche Antwort B wählte. Zudem fällt ein deutlicher Anteil von nicht zu bewertenden Aufgaben auf.

Diese Aufgabe ist allein schon vom Umfang des einzubringenden Vorwissens von weit höherem Schwierigkeitsgrad als die vorherige reine Koordinatenableitung. Sie erfordert neben der Koordinatenkenntnis auch Sachwissen zum Begriff ‚Gerade‘ sowie zur Lage von Geraden in Koordinatensystemen. Da diese Wissensaspekte verknüpft werden müssen, ist die Aufgabe deutlich komplexer als die vorhergehende. Dafür spricht auch der hohe Wert des internationalen Schwierigkeitsgrades mit einem Wert von 597. Über die Einordnung als Geometrie-Aufgabe lässt sich diskutieren, denn es ist nicht auf den ersten Blick ersichtlich, welche mathematischen Fähigkeiten durch diese Aufgabe erfasst werden sollen. Infolge der unter TIMSS kurz bemessenen Testzeit war sicher nicht an eine zeitaufwendige graphische oder rechnerische Lösung gedacht. Ebenso wenig dürfte eine Aufstellung der Geradengleichung und noch zeitaufwendigere rechnerische Überprüfung der vorgegebenen Lösungen angestrebt worden sein. Testwiseness-Strategien dürften hier auch wenig greifen. Daher bleibt ein Lösungsweg denkbar über Steigungsdreiecke etwa in der Art ‚von (3|2) nach (4|4) ist ‚eins nach recht und zwei nach oben‘ sowie der Versuch, dieses Steigungs- muster auf die vorgegebenen Punkte anzuwenden. Eine Klassifizierung als

Geometrieaufgabe ist dann kaum nachvollziehbar. Folgt der Schüler diesem Steigungsmuster, so muß er entweder von (4|4) nach ‚rechts‘ gehen oder von (3|2) nach ‚links‘. Da letzteres allerdings eine gedankliche Inversion erfordert („eins nach links und zwei nach unten“), kann vermutet werden, dass zuerst von (4|4) ausgehend nach rechts geprüft wird. Dann erhält man allerdings schon beim ersten Versuch (5|6), also *eine* und damit auch *die* Lösung.

Diese Lösungsstrategie beinhaltet eine komplexe Summe mathematischer Fähigkeiten, nämlich ein ‚Gefühl‘ für Steigungsdreiecke, wie es ein Schüler der Klasse 7 erst erwerben muss und ein nicht unmittelbar unterrichtlich damit befasster Schüler höherer Klassen oft nicht mehr direkt präsent hat. In der vorliegenden Studie, bei der zeichnerische Lösungen nicht verboten waren und ausreichend Bearbeitungszeit zur Verfügung stand, war daher eher ein graphischer Lösungsweg zu erwarten.

Die Schüler wählten ihn tatsächlich. Dazu zeichneten sie die vorgegebenen Punkte sowie eine Gerade durch diese Punkte in ein Koordinatensystem ein und versuchten, einen der als Lösung angebotenen Punkte auf der Geraden zu finden. Drei fehlerhafte Lösungen sind deshalb falsch, weil jeweils die Zeichnung unsauber gefertigt wurde. Ein Schüler zeichnet den Punkt C korrekt ein, vertauscht aber die Koordinaten.

Die Lösungsrate liegt deutlich höher als bei TIMSS, wo im internationalen Durchschnitt 38% der Schüler die Aufgabe lösten.

Zählt man auch solche Aufgaben als gelöst, bei denen Zeichenungsgenauigkeiten Fehlerursache waren, so ergibt sich eine Lösungsrate von insgesamt 73%. Die anderen Bearbeitungen sind falsch geraten oder nicht bearbeitet.

Aus den TIMSS-Lösungsquoten lässt sich also nicht zwingend ableiten, dass lediglich 40% der Schüler über genügend mathematische Fähigkeiten oder speziell Geometriekenntnisse verfügten, um die Aufgabe lösen zu können, sondern es ist ebenso denkbar, dass die Schüler die ihnen präsenten Lösungsstrategien unter den Bedingungen des TIMSS-Tests nicht einsetzen konnten. Dazu seien exemplarisch einige Aufgabenbearbeitungen vorgestellt.

Testheft 057 enthält eine Mehrfachantwort. Die Schülerin zeichnet die Punkte so ungenau, dass mehrere Lösungen möglich werden. Sie betont dies auch noch einmal: „... dass es zwei Lösungen gibt, glaube ich“. Bemerkenswert ist, dass sie sich ihrer Sache offenbar so sicher fühlt, dass sie dafür sogar eine Abweichung von der Grundregel ‚nur eine Lösung‘ in Kauf nimmt.

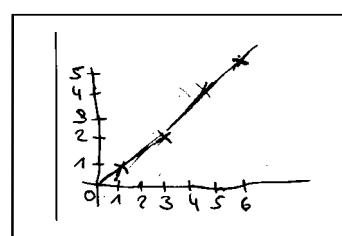


Abbildung 2: Linearität

Eine Schülerin (Testheft 033) findet graphisch die Lösung und bemerkt dann: „...dass sie ganz easy ist. Man muß nicht rechnen, nur probieren.“ Sie beschreibt eine deutliche Diskrepanz in den Anforderungen des graphischen gegenüber dem rechnerischen Lösungsverfahren.

Eine nachdenklich stimmende Aufgabenbearbeitung ist die der Schweizer Schülerin zu Testheft 223. Sie kommentiert ihre Lösungen in der Regel sehr ausführlich. Zur Aufgabe I08 ist erläutert sie:

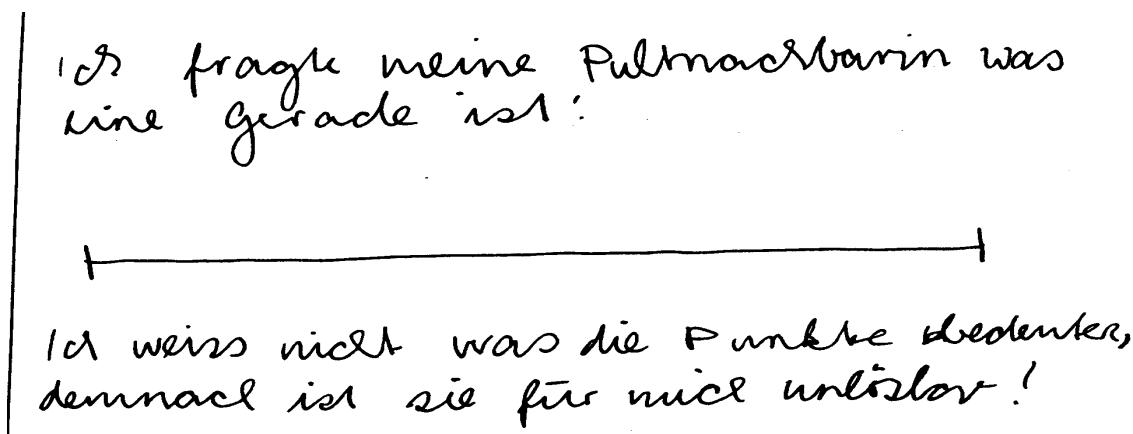


Abbildung 3: Geradenproblem

Sie hat also offensichtlich Probleme mit dem Begriff ‚Gerade‘. Die von der Mitschülerin als Hilfe angebotene und gezeichnete Strecke hilft ihr nicht weiter. Allerdings kann positiv gesehen werden, dass sie den unbekannten Begriff hinterfragt, statt die Aufgabe zu übergehen.

Dies kann als ein Hinweis darauf genommen werden, dass eine didaktische Interpretation von Mathematikaufgaben und ihren Lösungen noch viel kleinschrittiger und von elementaren Begriffen ausgehend anzusetzen ist, als man auch bei behutsamem Einfühlen in die mathematische Begrifflichkeit der Schüler vorherzusehen vermag.

Eine ebenfalls bemerkenswerte Aufgabenbearbeitung ist die des Testhefts 188. Die Schülerin zeichnet die gegebenen Punkte ein und markiert dann ein Steigungsdreieck. (Linie von $(3|2)$ bis $(4|4)$). Danach setzt sie das Steigungsdreieck nach links fort, allerdings unter Vertauschung von x- und y-Wert sowie unter Vernachlässigung der Forderung ‚Gerade‘. Daher kreuzt sie die falsche Lösung (A) an.

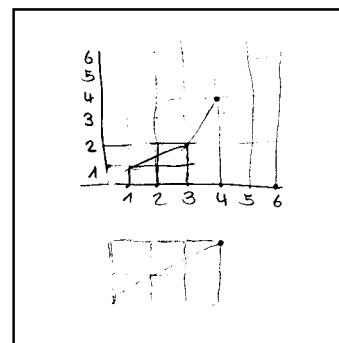


Abbildung 4: Linearität

Nach den oben geschilderten Überlegungen zur Lösungsstrategie ist diese Schülerin trotz falsch markierter Lösung die einzige, die die Idee des Steigungsdreiecks erfasst und – wenn auch fehlerbehaftet – auf das

Problem angewendet hat. Solche Sachverhalte kann ein MC-Test naturgemäß nicht auflösen.

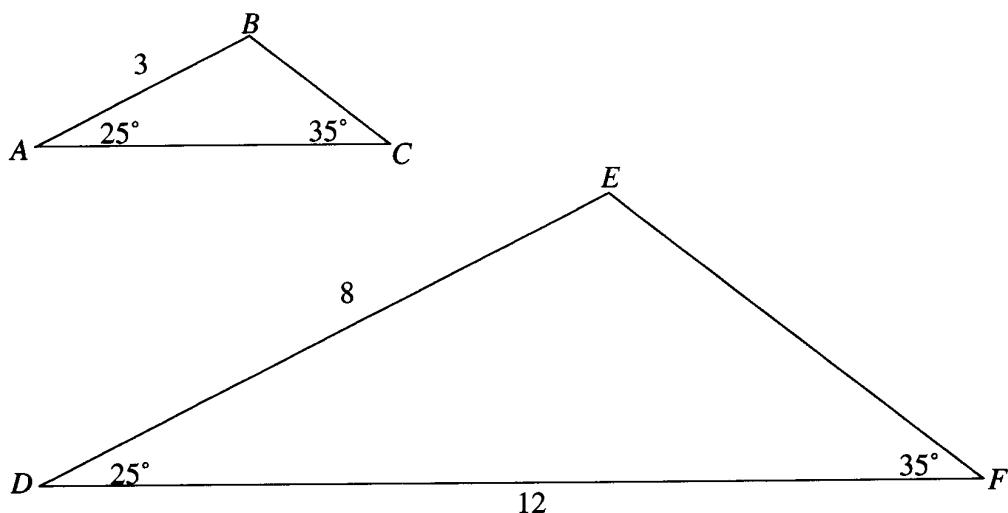
Aus der Tabelle der Lösungshäufigkeiten ergeben sich weitere Forschungsfragen. Es ist bemerkenswert, dass die fehlerhaften Lösungen keinesfalls statistisch verteilt angekreuzt wurden. So wird die der korrekten Lösung (C) ähnliche Antwort (E) sowohl in TIMSS wie auch hier gerade zu 5 % gewählt, obwohl sie bei einer unsauberer graphischen Lösung oder Koordinatenvertauschung ebenfalls in Betracht käme. Dagegen entscheidet sich etwa ein Viertel aller Schüler und bezogen auf die falschen Lösungen sogar ca. 40 % für die falsche Lösung (B). Dahinter lassen sich systematische Denkprozesse vermuten, die allerdings aus den vorliegenden Bearbeitungen nicht aufgelöst werden konnten.

Ein Vertauschen sämtlicher Punktkoordinaten führt übrigens tatsächlich zum Ankreuzen der korrekten Lösung (C).

5.2.6 Ähnlichkeit (TIMSS-Aufgabe P09)

Anhand der Aufgabe P09 soll abschließend der Umgang mit dem Begriff der Ähnlichkeit untersucht werden. Die Aufgabe lautet:

P09. Die Dreiecke ABC und DEF sind ähnlich.



Welches ist die Länge der Seite AC ?

- A. 2 B. 4 C. 4,5 D. 5,5 E. 32

Dieser Stoffbereich ist eher der Klasse 8 oder 9 zuzuordnen, so dass eine Analyse der Lösungsverfahren bei den hier partizipierenden Schülern der 7. Klasse

lohnenswert erscheint. Damit lässt sich erfassen, wie weit der Begriff der Ähnlichkeit „gefühlsmäßig“ in Schülern vorhanden ist, ohne dass eine exakte Definition in den Unterricht eingebracht wurde. Dabei bleibt der didaktische Stellenwert des Ähnlichkeitsbegriffes außer Betracht: man könnte vermutlich auch jedem Schüler der Klasse 4 Ähnlichkeitsabbildungen über die Vergrößerungs- oder Verkleinerungstaste eines Photokopierers erläutern. Natürlich werden dabei Eigenschaften wie Winkeltreue etc. nicht angesprochen.

Der internationale Schwierigkeitsgrad der Aufgabe ist nach TIMSS mit 617 hoch. Die Lösungshäufigkeiten in Prozent ergeben sich wie folgt. (Nicht alle Daten der Tabelle sind verfügbar.)

Tabelle 29: Lösungshäufigkeit bei der Ähnlichkeitsaufgabe

TIMSS–Aufgabe P09, Lösungshäufigkeit in %						
	TIMSS-Daten			eigene Daten		
	GER	CH	internat. Durchschnitt.	GER 7	CH 7	GER 8
Kl. 7	35	39	35	23	44	–
Kl. 8	37	41	38	–	–	60

Die vorgegebenen Antwortalternativen wurden mit folgender Häufigkeit gewählt:

Tabelle 30: Wahl der Antwortalternativen zum Ähnlichkeitsproblem bei TIMSS

TIMSS-Aufgabe P 09, Angaben in %			
	Deutschland	Schweiz	internationaler Durchschnitt
A (2 cm)	2	3	4
B (4 cm)	22	21	23
C (4,5 cm) (korrekte Lösung)	35	39	35
D (5,5 cm)	28	26	27
E (32 cm)	2	3	5
Sonstige (keine Lösung etc.)	11	8	6

Es bleibt bemerkenswert, dass die Ergebnisse einer Kleinpopulation im Wesentlichen den internationalen Durchschnitt widerspiegeln. Damit erhalten Mikrosurveys zwar keine statistische Signifikanz, spiegeln aber zumindest hier ein Leistungsbild der Testpopulation zum internationalen Durchschnitt. Gleiches war schon beim fehlertoleranten Leistungsvergleich in 4.3 zu beobachten.

Die Lösungen A und E werden nur zu einem geringen Teil gewählt. Dagegen werden die zahlenmäßig ähnlichen Lösungsvorgaben B, C und D häufig gewählt – mit einem leichten Überhang bei der korrekten Lösung C.

Diese Feststellung gilt so im wesentlichen für alle partizipierenden Länder. Selbst in Japan, Korea oder Singapur (Spitzengruppe für diese Aufgabe), wo die korrekte Lösung zu 54%, 57% und 50% gewählt wurde, sind B und D noch merkbar angekreuzt worden: Bei Japan und Korea wurden gewählt B zu 25%, D zu 17%, in Singapur B zu 1%, D zu 25%, A und E dagegen nur mit etwa 2%. Dies spricht für einen tiefer liegenden Hintergrundprozess. Im folgenden soll versucht werden, diese Lösungsentscheidungen plausibel zu erklären.

Es gibt mindestens zwei denkbare Strategien zur Lösung der Aufgabe.

Zum einen lässt die Komplexität der Aufgabe vermuten, dass Schüler testwise ness-Strategien einsetzen. Dann genügt zum Ausschluss falscher Lösungen die Betrachtung des kleinen Dreiecks ABC. Unter Weglassung aller weiteren Informationen ist aus der Zeichnung offensichtlich ‚auf einen Blick‘ zu entnehmen, dass die gesuchte Länge der Seite von A nach C im kleinen Dreieck etwas größer als drei sein muss. Deshalb scheiden die Lösungen A (2) und E (32) direkt aus.

Schlägt man nun mit dem Auge einen Kreis um Punkt A mit Radius 3 und versucht, den Schnittpunkt dieses Kreises auf AC zu sehen, so kommt man bei guter Schätzung auf etwas mehr als 4 cm für die Länge der Strecke von A nach C. Allerdings liegt auch 5,5 cm nicht fernab.

Nimmt man ferner an, dass einige Schüler die Aufgabe ‚gefühlsmäßig rechnend‘ bearbeiteten, so liegen folgende Gedanken nahe: ‚um von acht auf zwölf zu kommen, muss ich noch einmal die Hälfte von acht dazutun. Gleiches muss ich dann mit drei machen‘. (Alternativ könnte auch der Faktor 1,5 erkannt worden sein). Eine solche intuitive Erfassung der Streckung liefert das korrekte Ergebnis 4,5 cm, ohne dass sich der Schüler der Problematik der Ähnlichkeit bewusst werden muss.

Schüler, die diesen Weg beschritten und richtig rechneten, erhöhten damit den Prozentsatz richtiger Lösungen gegenüber einem gleichmäßigen Raten.

In der vorliegenden Untersuchung hatten die Schüler Zeit, sich mit der Aufgabe intensiver als bei TIMSS zu befassen, so dass andere Lösungsstrategien zum Einsatz kommen konnten.

Ein Nachmessen mit dem Lineal war nicht möglich, da die Grafik absichtlich nicht maßstäblich dargestellt wurde. Die Strecke von D nach F beträgt im Testheft 13,5 cm, um die Schüler darauf aufmerksam zu machen, dass dieser Lösungsweg ungeeignet ist. Die Strecke von A nach C beträgt 5 cm, um eine ausschließende Entscheidung für die Lösungen C oder D unmöglich zu machen.

Vorliegend wurde diese Aufgabe im Vergleich zu allen anderen Geometrie-Aufgaben bemerkenswert häufig nicht bearbeitet oder (falsch) geraten. Dies stützt die Annahme der Curriculuminvalidität. Die Lehrerbegleitbücher bestätigen diese Vermutung: der Lehrer der Klasse 7A betont ausdrücklich „Ähnlichkeitsbegriff nicht eingeführt in Klasse 7“ und betrachtet die Aufgabe als unlösbar für seine Schüler (ein Schüler hat sie dennoch gelöst). Die anderen Lehrerbegleitbücher weisen die Aufgabe als „unbekannt, aber mit Überlegen lösbar“ aus.

Unter den Bearbeitungen sind sechs so bemerkenswert, dass sie hier einzeln vorgestellt und analysiert werden sollen.

Der Schüler zu Testheft 046 berechnet abweichend von der Aufgabenstellung den nicht angegebenen Winkel bei B. Vermutlich war gerade der Innenwinkelsatz für Dreiecke Unterrichtsgegenstand (Schülerangabe: „...das haben wir vor kurzem erst gemacht“). Die eigentliche Aufgabe wurde nicht erkannt, obwohl die gesuchte Länge sogar ausdrücklich hingeschrieben wurde. Es folgt die bemerkenswerte Kommentierung: „Wofür ist das Dreieck DEF gut?“ (!). Hier ist es schwer, den Gedankengang der Schülerin nachzuvollziehen.

In Testheft 099 wird folgender Lösungsweg angeboten:

„Ich komme auf das Ergebnis, indem ich
als erstes \overline{AB} reiche und aus Punkt A der
Strecke \overline{AB} den Winkel 25° eintrage, dann trage
ich aus Punkt B den Winkel 120° ein (da $25^\circ + 35^\circ$
der oben freien Schenkel des Winkels = 60° sind
 \Rightarrow und $180^\circ - 60^\circ$
 $\Rightarrow 120^\circ$ sind)
= 25° trifft, so entsteht das Dreieck
 $\triangle ABC$ und schließlich jetzt messen wie
lang seit AC ist.“

Abbildung 5: Zur Ähnlichkeit

Der Schüler fertigt eine vollständige Dreieckskonstruktion, basierend auf der angegebenen Seitenlänge und der Berechnung des fehlenden Winkels, und beschreibt sie zusätzlich mit Worten so, wie es vermutlich gerade im Unterricht gefordert wurde. Anschließend misst er die gesuchte Länge schlicht ab.

Dieser Lösung darf man sicher einen Ansatz von Kreativität zusprechen, da der Schüler vorhandenes Wissen in neuem Rahmen gezielt zur Lösung eines Problems einsetzt. Dass er dabei den Weg über die Ähnlichkeit nicht findet, kann ihm nicht angelastet werden. Gerade bei fehlender Curriculumvalidität ist ein abweichend vom Standard-Lösungsweg gefundenes Verfahren eher als kreatives Denken einzuordnen.

TIMSS kann solche kreativen Schülerlösungswege nicht registrieren, da dafür kein Zeit- und Erfassungsrahmen bereitgestellt wurde.

Der Schüler zu Testheft 071 gibt an, eine Aufgabe wie diese sei für ihn neu. Dennoch löst er die Aufgabe korrekt so:

zwei 8er und 0,5 sind 18

$$8 + \frac{1}{2}8 = 12$$

dann nun addiere ich auch die Hälfte von 3 mit 3 und erhält die Lösung 4,5

Abbildung 6: Verhältnisberechnung

Er bildet auf dem oben bereits angesprochenen Weg das Verhältnis 1,5 der beiden Seiten im größeren Dreieck über Addition der Hälfte von 8 und wendet dies auf das kleinere Dreieck an. Der Schüler verwendet hier intuitiv ohne entsprechende unterrichtliche Vorarbeit den 2. Strahlensatz mit entsprechenden Proportionalbeziehungen.

Auch diesem Lösungsweg muß man ein hohes Maß an Kreativität zugestehen, genauso wie dem des Schülers zu Testheft 073, der ebenfalls intuitiv den 2. Strahlensatz erkennt, das entsprechende Streckenverhältnis bildet und sich äußerst geschickt um Bruchverhältnisse „herumarbeitet“:

$$\begin{array}{r}
 3 = 8 \\
 \cancel{2} \cancel{4} \cancel{6} \cancel{8} \\
 \hline
 8 = 24 \\
 4,5 = 12
 \end{array}$$

Ich habe das Verhältnis zwischen der Strecke DF und DE gesucht

Abbildung 7: Weg zur Verhältnisberechnung

Er vermeidet Brüche, indem er erst mit dem ganzzahligen Faktor 3 von 3–8 auf 9–24 umrechnet und anschließend halbiert auf 4,5–12. Diese Rechentechnik und der Kommentar sprechen dafür, dass dem Schüler das eigentliche Ziel des Rechenweges klar ist.

Der Schüler zu Testheft 198 löst das Problem des Bruchfaktors durch eine Abfolge von Division und Multiplikation.

Daneben entnimmt man aus den durchgestrichenen Rechnungen, dass er erst mit Differenzen experimentiert hat, bevor er Proportionalverhältnisse als Lösungsweg entdeckte.

$ \begin{array}{r} 8-3 \cancel{8} \\ \cancel{12}-\cancel{5}=7 \end{array} $ $8 : 2 \cdot 3 = 12$ $3 : 2 \cdot 3 = \underline{\underline{4,5 \text{ cm}}}$
Abbildung 8: proportionale Analogie

6 Ergebnisse

6.1 Globale Befunde

Hier werden die wesentlichen Befunde der vorliegenden Studie vorgestellt. Details finden sich im den Kapiteln über Bearbeitungsqualität, Geometrie sowie bei den Einzelbefunden.

Insgesamt konnte festgestellt werden:

Der in TIMSS beobachtete Leistungsabstand Schweiz/Deutschland wird in der vorliegenden Studie stabil reproduziert. Dies gilt selbst bei der hier vorliegenden Beschränkung auf eine spezielle Schulform und eine erheblich geringere Probandenzahl.

Die Studie ermöglicht nach ihrer Anlage in Stufen immer fehlertolerantere Beurteilungen zur Bearbeitungsqualität der einzelnen Aufgaben. Dabei bleibt der Leistungsabstand Schweiz zu Deutschland erhalten. Dies schließt Rankingunterschiede auf Grund von reinen Rechenfehlern aus.

Die Lehrerkommentare zur Curriculumvalidität unterscheiden sich deutlich von denen der Schüler.

Die Curriculumvalidität ist zumindest in der Schweiz kein entscheidendes Kriterium für die Bearbeitungsgüte. Die Schweizer Schüler lösen auch Geometrieraufgaben, welche sie nach Ansicht ihrer Lehrer nicht lösen können. Allerdings steigt ihre Lösungsrate gegenüber TIMSS nur wenig, wenn ihnen faktisches Wissen wie etwa der Innenwinkelsatz im Dreieck oder die vereinbarte Reihenfolge von Punktkoordinaten fehlt. Auch dann werden punktuell noch Lösungsansätze versucht.

Selbst bei noch unbekannten geometrischen Fachbegriffen (Ähnlichkeit, Kongruenz etc.) versuchen die Schweizer Schüler, darin einen Sinn zu finden, und sie bearbeiten die Aufgaben dennoch. Für die deutschen Schüler kann dies so nicht festgestellt werden.

Die gegenüber TIMSS deutlich erhöhte Bearbeitungszeit führte bei vielen Aufgaben trotz elaborierter Bearbeitung mit eingeforderten Zusatzkommentaren zu einer ebenso deutlich besseren Bearbeitung.

Aus den freien Kommentaren der Schweizer Schüler ist zu entnehmen, dass sie sich intensiv mit den Aufgaben beschäftigen. Die deutschen Schüler kommentieren dagegen nahezu ausschließlich schuldabweisend.

Die Bearbeitungsintensität der Schweizer Schüler ist deutlich höher als in Deutschland. Dies zeigen sowohl das jeweils beobachtete Testumfeld als auch die Bearbeitungsbereitschaft an. Die Schweizer Schüler waren wesentlich eher mit der Bearbeitung ihrer Testhefte fertig als die deutschen und verlangten in der Regel ein weiteres Aufgabenbuch, welches sie auch bearbeiteten, wenn auch wohl aus Zeitgründen nicht immer vollständig.

Zusätzlich verlangten sie direkt oder indirekt, eine Kontrollmöglichkeit über ihre Bearbeitungen zu erhalten.

Rateeffekte sind von marginaler Bedeutung und daneben statistisch gleichverteilt. Auch Testwiseness-Strategien werden höchstens punktuell, keinesfalls aber gezielt eingesetzt.

Zusammenfassend gilt: Der Leistungsunterschied Schweiz/Deutschland ist nach der vorliegenden Studie nicht durch mathematisch-fachlich erkennbare Faktoren, sondern eher durch soziokulturelle Aspekte bedingt.

Es bleibt zu bemerken: die hier vorgestellten Befunde sind aus reinen MC-Tests nicht zu erhalten.

6.2 Einzelbefunde

Die Einzelbefunde wurden i.w. in den Untersuchungen zur Arbeitsbereitschaft und zur Geometrie angesprochen. Sie werden hier zusammenfassend dargestellt.

Der für die Studie gewählte Zeitrahmen von 90 Minuten für 12 Aufgaben erwies sich als ausreichend auch für elaboriert darzustellende Lösungswege sowie die zusätzliche Bearbeitung der Questionnaires.

Dabei benötigten die Schweizer Schüler eine deutlich kürzere Bearbeitungszeit.

Der von TIMSS in den Schülerleistungen Schweiz zu Deutschland konstatierte Leistungsunterschied wird reproduziert.

Auch bei einer gestuft fehlertoleranten Aufgabenbewertung bleibt der Leistungsabstand zwischen beiden Ländern erhalten.

Der statistische Rateerfolg von etwa 50% bei als geraten gekennzeichneten Lösungen spricht gegen den Einsatz von testwiseness-Strategien.

Die Bearbeitungsqualität ist bei den Schweizer Schülern erkennbar besser:

- Sie raten deutlich weniger,

- sie lassen weniger Ketten von Aufgaben aus und
- sie hören selten schon vor Bearbeitung der letzten Aufgabe auf.

Eher gilt:

oft verlangen sie nach Bearbeitung ihres Testheftes nach einem weiteren Aufgabenheft.

Zusätzlich bearbeiten sie mit gutem Erfolg auch Aufgaben, die nach Ansicht der unterrichtenden Lehrer noch nicht Gegenstand des Unterrichts waren.

Dies ist bei den deutschen Schülern so nicht der Fall.

Der Charakter der freien Aufgabenkommentierungen ist für beide Länder deutlich unterschiedlich.

Die deutschen Kommentierungen sind oft schuldabweisend gehalten, die Schweizer Kommentierungen sind eher von sachlicher Art.

Es gibt einen Anteil von 8 % geratener Lösungen (bezogen auf alle Probanden). Der Rateanteil der deutschen Schüler war etwa doppelt so hoch wie bei den Schweizer Schülern. Eine Dunkelziffer nicht ehrlich angegebener ‚Lösungen durch Raten‘ kann nicht abgeschätzt werden. Da jedoch Raten als Lösungstechnik erlaubt war, bestand für die Schüler kein Grund, dies zu verheimlichen.

Für die deutschen Schüler liegen die TIMSS-Aufgaben nach Ansicht der unterrichtenden Lehrer näher am Curriculum als für die Schweizer Schüler. Dies gilt in besonderem Maße für den Teilbereich Geometrie. Dennoch ist die Schweizer Lösungsrate auch dort deutlich höher als die der deutschen Schüler.

Die Curriculumvalidität aus Lehrerperspektive kann von der aus Schülerperspektive deutlich differieren.

Eine Korrelation zwischen Curriculumvalidität aus Lehrersicht und Lösungshäufigkeit ist nicht erkennbar.

Die Schülererfolge bei den Geometriaufgaben sind teils vergleichbar mit TIMSS, teils aber auch drastisch besser als dort. Die Detailanalyse deckt auf, dass das Aufgabendesign eine Maskierung von Schülerleistungen zur Folge haben kann. So interpretieren die deutschen Schüler eine Geometriaufgabe (M07) vermutlich so um, dass sie ihren Mathematik-Erfahrungsraum trifft; daraus resultierend markieren sie eine falsche Lösung, obwohl die originäre TIMSS-Aufgabe durchgängig korrekt gelöst wurde.

Bei einer weiteren Geometrie-Aufgabe (K08) kann die auch nach TIMSS international geringe Lösungsrate ein Zeitproblem sein. Zusätzlich suggeriert die Aufgabenpräsentation eine falsche Lösung. Unter Zeitdruck wurde daher das

Problem anscheinend nur teilweise erkannt (nur Drehung statt Drehspiegelung). Bei genügend Arbeitszeit gilt:

bei den deutschen Schülern steigt hier die Lösungsrate deutlich von 28 auf 70 %, für die Schweizer Schüler liegt Curriculuminvalidität vor; vermutlich daher hat sich die Lösungsrate nicht verbessert. Sie beschäftigen sich aber dennoch intensiv mit der Aufgabe.

Auch Koordinateninterpolationen scheinen bei hinreichender Bearbeitungszeit deutlich besser zu werden (z.T. durch Einbringen von Hilfslinien) als bei knapp bemessener Zeit unter den TIMSS-Konditionen. Bei den Schweizer Schülern sind Koordinaten ein Problem, weil ihnen die vereinbarte Reihenfolge der Koordinaten (noch) nicht bekannt ist.

Bei Aufgabenbearbeitung zur (nicht curriculumvaliden) Ähnlichkeit werden ausgesprochen kreative Ansätze erkennbar. Die Beurteilung solcher Aufgabenbearbeitungen ist nicht im MC-Verfahren erfassbar. Realistische Leistungsbeurteilungen müssen sich daher auch hier auf elaborierte Aufgabenbearbeitungen stützen.

Weitere Auswertungen können möglicherweise Korrelationen zwischen den Lösungsqualitäten in anderen mathematischen Teilbereichen (Analysis, Statistik, ...) und der Geometrieleistung aufdecken oder auch Informationen zur Curriculuminvalidität ergeben. Das eingeführte Bewertungsraster hat sich zumindest hier bewährt..

Aus den ausgewerteten Aufgabenbearbeitungen lassen sich keine spezifisch *mathematikimmanenten* Gründe für die ‚besseren‘ Leistungen der Schweizer Schüler isolieren.

Hinsichtlich der Curriculuminvalidität sind die Schweizer Schüler speziell in der Geometrie benachteiligt gegenüber den deutschen Schülern, da dieser Teilbereich in der Schweiz zum Testzeitpunkt im wesentlichen curriculuminvalid war.

Es gibt jedoch Indikatoren dafür, dass die Schweizer Schüler sich intensiv sachlich mit den Aufgaben auseinandersetzen und nach eigenen Lösungswegen suchen, dies auch bei nicht curriculumvaligen Aufgaben.

Daneben ist zu betonen, dass die Schweizer Schüler durchgängig in der vorhandenen Arbeitszeit nicht nur ein Testheft bearbeiteten, sondern noch ein weiteres Testheft erbaten, wogegen die deutschen Schüler oft Aufgaben des ersten Testheftes nicht bearbeiteten.

Aus der Studie ergeben sich keine signifikant belegbaren Ursachen für diese erkennbar länderspezifisch unterschiedliche Testbearbeitung. Allerdings war das

Studiendesign auch nicht auf eine Erfassung sozialer Skripte (soweit sie existieren) und ihre Auswirkungen auf Mathematikleistungen ausgelegt.

In summa sprechen die Ergebnisse für die Hypothese, dass für die differierenden Testleistungen Schweiz/Deutschland weniger mathematikintrinsische, sondern eher externe Faktoren wie etwa Einsatzbereitschaft, Durchhaltevermögen und Bereitschaft zur selbstständiger Erarbeitung bedingend waren. Dies gilt insbesondere für die Bearbeitung nicht curriculumvalider Aufgaben aus dem Bereich der Geometrie.

7 Ausblick und weitere Forschungsfragen

Die eingangs gestellten Forschungsfragen konnten durch die Studie aussagekräftig beantwortet werden. Zwar ist die Stichprobengröße für eine statistisch repräsentative Aussage zu gering, doch zeigen die allgemeinen Ergebnisse beim Vergleich zu TIMSS so deutliche Analogien, dass man ihr ein hohes Maß an Repräsentativität zubilligen kann.

Nach TIMSS zeigen deutsche Schüler relative Schwächen in den Kernbereichen Algebra und Geometrie. Das musste hier relativiert werden, da vorhandenes Geometriewissen teilweise durch Testform, Fragestellung oder sonstige Bedingungen maskiert wurde.

Der TIMSS-Befund, wonach die relativen Leistungsstärken der deutschen Schüler der 8. Jahrgangsstufe in der Arithmetik, im Umgang mit Maßeinheiten und in der deskriptiven Statistik liegen, konnte wegen der geringen Stichprobengröße (eine Klasse 8) nicht überprüft werden.

Es wäre zu untersuchen, warum die deutschen Schüler gerade im Bereich Algebra die Aufgabenbearbeitung deutlich häufiger als in den anderen Teilbereichen abbrechen oder die Ergebnisse zu erraten versuchen.

Möglicherweise wäre es zusätzlich sinnvoll und praktikabel, eine Bewertungsdimension einzubringen, die für die Wissenskategorien oder Kategorien des Handelns, der Bearbeitungsgüte und für didaktische Besonderheiten kodiert.

Bereits in der Einleitung wurde auf die erhebliche Bedeutung hingewiesen, die den TIMSS- und auch den nachfolgenden PISA-Resultaten in der Öffentlichkeit und auch in der Fachdidaktik zugebilligt wird. Allein aus den TIMSS-Ergebnissen wurden und werden weiterhin Konsequenzen für Verbesserungen des Mathematikunterrichtes abgeleitet und vorgeschlagen.

Die in der vorliegenden Studie gewonnenen Erklärungsparameter für Mathematikleistungen speziell in der Geometrie sprechen dafür, neben Großerhebungen weiterhin detailauflösende Untersuchungen durchzuführen. Der wesentliche Vorteil gegenüber standarisierter Erhebungen auf MC-Basis liegt darin, dass sich die Gründe für ein Versagen in der konkreten Aufgabe analysieren lassen. Dazu bedarf es allerdings in der Sache geübter Mathematik-Pädagogen, die aus den vorgelegten Lösungswegen Rückschlüsse auf die Denkwege der Schüler zu ziehen vermögen.

Nach den hier vorliegenden Rechenwegen und Ergebnissen kann vordergründiges Versagen bei einer Testaufgabe nicht immer mit schlechter oder nur mäßiger mathematischer ‚Begabung‘ begründet werden. Gezielte Einzelstudien könnten

dagegen wertvolle didaktische Zusatzinformationen zur Analyse des Ist-Zustandes und zur weiteren Entwicklung von Unterrichtsqualität geben.

Die Strukturierung der Unterrichtsarbeit gehört zum Aufgabenbereich jedes einzelnen für eine Klasse verantwortlichen Fachlehrers. Parallel dazu scheint es sinnvoll zu sein, sich über so gewonnene Erkenntnisse in den Fachkollegien auseinander zu setzen. Auf Grund der dem Verfasser bekannten Abneigung in den Fachkollegien könnte hier eine wissenschaftliche Begleitung, also Koordinierung ‚von außen‘ wertvoll sein.

Dabei darf nach dieser Studie der Leistungswille der Schüler nicht außer Acht gelassen werden. Es bleibt sicher ein weites mathematikdidaktisches Forschungsfeld, ob dieser bei Tests durch Sanktionen erzwungen werden kann oder wie weit motivative Förderung greift.

8 Literaturverzeichnis

Soweit bei Internet-Quellen Autor und Datum der Veröffentlichung bekannt sind, werden diese angegeben, ansonsten werden Datum und Quelle des Downloads angegeben. Die Zitation der Buchmedien folgt den üblichen Regeln.

Achermann, St.; 16.09.1998: *It takes a village*; Forum IFNEC;
ifnec @admin.humberc.on.ca>

Affolter, W. et al: *Das Zahlenbuch 5*; Zug; Klett und Balmer 1998

Bangert-Drowns, et al.; 07.01.2001; *The instructional effect of feedback in test-like events*; in: Review of Educational Research, 61, 213–238;
<http://www.aera.net/pubs/rer/index.htm>

BASS (*Bereinigte amtliche Sammlung der Schulvorschriften des Landes Nordrhein-Westfalen*), 14. Ausgabe 1999/00: Abschnitt 10-45 Nr2; Frechen;
Ritterbach Verlag 1999

Bauer, H.; 11.03.1998; *Denkschrift (1977) der Deutschen Mathematiker-Vereinigung e.V. zum Mathematikunterricht am Gymnasium*; über
<http://www.mathematik.uni-bielefeld.de/DMV/archiv>

Baumert, J. et al.: *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich: Deskriptive Befunde*; Berlin; Leske & Budrich 1997

Baumert, J. et al.: *TIMSS / III: Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich*; Berlin; Max-Planck-Institut für Bildungsforschung 1999

Baumert, J. et al.; 05.09.2000: *Internationales und nationales Rahmenkonzept für die Erfassung von mathematischer Grundbildung in PISA*;
<http://www.mpib-berlin.mpg.de/pisa/pdfs/KurzFrameworkMaths.pdf>

Beckmann, J.F. und Guthke, J.: *Complex Problem Solving, Intelligence, and Learning Ability*, in: Funke (Hrsg.); *Complex problem solving: The European Perspective*, S. 177-200; Hillsdale, NJ; Lawrence Erlbaum Associates 1995

Bezirksregierung Düsseldorf; 13.05.1999: *Testaufgaben zur Analysis für Grundkurse April 1999*; Eigenauflage Bezirksregierung 1999

Bezirksregierung Düsseldorf; 4.2.1999: *Vergleichsklausuren in Jahrgangsstufe 12*
<http://www.bezreg-duessel-dorf.nrw.de/schule/mathe/material/vklausur/vklaus1.htm>

Bieber, G.: *Qualitätsuntersuchung in Schulen im Unterricht in Mathematik, Mathe-40 und Quasum-Testaufgaben im Vergleich*; Brandenburg; PLIB (Pädagogisches Landesinstitut Brandenburg); Eigendokumentation vom 22.9.1999

Bleuel, J.; 27.03.2000; *Zitieren von Quellen im Internet*;
<http://ourworld.compuserve.com/homepages/jbleuel/ip-zitl.htm> 1996

Bleuel, Jens: *Zitation von Internet-Quellen*; in: Hug, Theo, (Hrsg.): Wie kommt die Wissenschaft zu ihrem Wissen?. Band 1: Einführung in das wissenschaftliche Arbeiten (Buch und CD-ROM); Hohengehren; Schneider Verlag 2001; auch online über
<http://www.bleuel.com/ip-zit.pdf>

BLK (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung) Heft 60: *Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts*; Bonn; BLK Eigenpublikation 1997

Blum, W. und Wiegand, B.: *Wie kommen die deutschen TIMSS-Ergebnisse zu stande?*; in: Blum, Neubrand (Hrsg.): TIMSS und der Mathematikunterricht; Hannover; Schroedel 1998

Büchter A. und Leuders, T.: *Vortrag Lernstandserhebungen in NRW*; Tagung in Kassel der GDM Gesellschaft für Didaktik der Mathematik, Fachbereich ‚Internationale Großuntersuchungen‘; 26./27.1.2004; 2004

Burkard, T.; 09.11.2000: *National curriculum standards*;
burkard@promethean.fsnet.co.uk

Cohen, L. und Manion, L: *Research Methods in Education*; London; Routledge 1998

Demuth, R. und Reiser, O.: *Molekulares Spiegelkabinett*; in: Spektrum der Wissenschaften Verlagsgesellschaft 2/2005; Heidelberg 2005

Digital libraries; 3.5.2002; *Cataloging and indexing of electronic resources*;
Zugänglich z.B. über die DVB (Düsseldorfer Virtuelle Bibliothek)
<http://www.uni-duesseldorf.de/ulb/biber.html>

Dirac, P.A.M: *The Evolution of the Physicist's Picture of Nature*; Scientific American Mai 1963, Verlagsgruppe Georg von Holtzbrinck GmbH 1963

EDK (schweizerische Erziehungsdirektorenkonferenz); 14.04.1999; *Projekt Euridike*; Sekretariat der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren EDK;
<http://edkwww.unibe.ch/d/euridice/euri-Vorwort.html>

Europäische Kommission; 03.10.2000; *Leonardo da Vinci Projekt AEVEM Mathematics*;
http://158.169.50.70/comm/education/tests/liste_de.html#1

Euridice; 3.4.2000;
<http://edkwww.unibe.ch/d/euridice/euri-Vorwort.html>

Fordham Report; 9.11.2000; *State Mathematics Standards, An Appraisal of Math Standards in 46 States, the District of Columbia, and Japan*;
<http://rsvh.addr.com/mva/id141.htm>

Haenisch, H.: *Warum TIMSS nicht geeignet ist, etwas über die Leistungsfähigkeit deutscher Schulen auszusagen*; in: TIMSS und BiJu; Schriftenreihe des Bildungs- und Förderungswerks der Gewerkschaft Erziehung und Wissenschaft im DGB e.V. 1998
F

Hogan, Th.; 03.10.00; *Measurement Implications of „A Nation at Risk“; ERIC Digest*;
<http://ericae.net/db/edo/ED286943.htm>

Hundt, D.: *In Mathe Mangelhaft. Die TIMS-Studie: Deutsche, niederländische und Schweizer Schüler im Vergleich*; Berlin; Eigenverlag BDA (Bundesvereinigung der Deutschen Arbeitgeberverbände) 2000

Jacobs, B.; 17.06.1998 und 07.01.2001; *Aufgaben stellen und Feedback geben*; Medienzentrum der Philosophischen Fakultät der Universität Saarbrücken;
<http://www.phil.uni-sb.de/~jakobs/wwwartikel/feedback/index.htm>

Jacobsen, M.; 25.04.1999; *Achieving Academic Success*; last updated May 1999;
http://www.acs.ucalgary.ca/~dmjacobs/study_skills_siteshtml/Testwise-ness

Jäger, R.: *Vortrag im GDM-Arbeitskreis Vergleichsuntersuchungen im Mathe-matikunterricht*; Kassel 2.12.2000

Kaiser, G.: *Möglichkeiten und Grenzen internationaler Vergleichsuntersuchun-*

gen im Mathematikunterricht; in: Beiträge zum Mathematikunterricht 1999; Hildesheim; franzbecker 1999

Kaiser, G.: *Internationale Vergleichsuntersuchungen im Mathematikunterricht – eine Auseinandersetzung mit ihren Möglichkeiten und Grenzen*; in: Journal für Mathematikdidaktik Jg. 21(2000) Heft 3/4; Stuttgart; Teubner 2000

Keitel-Kreidt, Chr.: *TIMSS zwischen Wissenschaft und Politik. Beispiel für vergleichende Studien zum Mathematikunterricht, Hoffnung und Gefahren*; GEW-Expertenhearing; Kassel 1998

Keitel-Kreidt, Chr, Kilpatrick, J: *The Rationality and Irrationality of International Comparative Studies*; in: Kaiser, G. et al. (Hrsg): International Comparisons in Mathematics Education; London 1998

Klieme, E. et al.; 24.6.2004; *Expertise zur Entwicklung nationaler Bildungsstandards*;
http://www.bmbf.de/pub/zur_Entwicklung_nationaler_Bildungsstandards.pdf 2003

KMK 2003a (Kultusministerkonferenz); 26.4.2004; *Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10), (Beschluss der Kultusministerkonferenz vom 04.12.2003)*; file Bildungsstandards/Mathematik_MSA_BS_04-12-2003.pdf; Bonn 4.12.2003; über <http://www.kmk.org/schul/Bildungsstandards/>

KMK 2003b (Kultusministerkonferenz); 26.4.2004; *Entwicklung und Implementation von Bildungsstandards*; file Argumentationspapier_stand-1.pdf; Bonn Dezember 2003; über <http://www.kmk.org/schul/Bildungsstandards/>

KMK 2004 (Kultusministerkonferenz); 13.5.2004; *Bildungsstandards im Fach Mathematik für den Hauptschulabschluss nach Klasse 9, Entwürfe*; file Mathematik_Hauptschulabschluss_23_04_04.pdf; Bonn 23.4.2004; über www.kmk.org/schul/Bildungsstandards/

KMNRW Kultusministerium des Landes Nordrhein-Westfalen: *Richtlinien und Lehrpläne Mathematik Sekundarstufe II*, in: Heft 4720; Frechen; Ritterbach Verlag 1999

Knoche, N.: *Methodologische Aspekte bei TIMSS, (Un-)Abhängigkeit der Rangreihen vom Rasch-Modell, Zusammenhang von Testmodell und Itemkonstruktion*; Vortrag anlässlich der Tagung der Arbeitsgruppe der GDM

(Gesellschaft für Didaktik der Mathematik) ,Internationale Vergleichsstudien'; Kassel 6.11.1998

Lange, J.de,: *Looking through the TIMSS-mirror from a teaching angle*; Utrecht; Freudenthal-Institut; Eigendruck GEW Materialien und Folien zum GEW-Experten-Hearing; Kassel 1998

Lind, D.: *Über die in TIMSS verwendeten Testmodelle*, in: Beiträge zum Mathematikunterricht; S. 349ff; Hildesheim; franzbecker 1999

Lind, D.: *Welches Raten ist unerwünscht? Eine Erwiderung*; in: Journal für Mathematikdidaktik Jg 25 Heft 1; Stuttgart; Teubner 2004

McREL Institute; 22.11.1999; (Mid-continent Research for Education and Learning, based in Aurora, Colorado, based in Autorenteam; *Standards-benchmarks*;
<http://www.mcrel.org/standards-benchmarks/docs/toc.asp>

Meyerhöfer, W.: *Was mißt TIMSS? Eine objektiv-hermeneutische Analyse von Testaufgaben*, in: Beiträge zum Mathematikunterricht 2000; Hildesheim; franzbecker 2000

Meyerhöfer, W.: *Zum Problem des Ratens bei PISA*, in: Journal für Mathematikdidaktik Jg 25 Heft 1; Stuttgart; Teubner 2004

MJSK Ministerium für Jugend, Schule und Kinder des Landes NRW; 11.6.2004: *Standardsetzung und Standardüberprüfung in Nordrhein-Westfalen*; file standardsicherung_nrw.pdf; Eigenpublikation des Ministeriums, Düsseldorf 19.4.2004; über
www.learn-line.nrw.de/angebote/standardsicherung/downloads

MLA (Modern Language Association); 19.06.03; *Bibliographic Citation Form Guide*;
<http://www.cantonma.org/~connorp/mla.htm>

MNU (Verein zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts e.V.); 23.07.2000; *Mathematische und naturwissenschaftliche Bildung an der Schwelle zu einem neuen Jahrhundert*;
<http://www.mnu.de/Schwelle.htm>

Moser, U. et al.: *Schule auf dem Prüfstand: Eine Evaluation der Sekundarstufe I auf der Grundlage der 'Third International Mathematics and Science Study'*; Zürich; Rüegger 1997

Mouly, G. J.: *Educational Research: The Art and Science of Investigation*; Boston; Allyn and Bacon 1978

MPI (Max-Planck-Institut) PISA; 13.08.2000; *Programme for International Student Assessment*;
<http://www.mpib-berlin.mpg.de/pisa/mathelit.htm>

MSC (Mathematics Standard Committee of San Diego City); 21.11.1999;
Mathematics Content and Performance Standards;
<http://www-internal.sandi.net/standards/Math-eng/mathstand.htm>

MSWF (Ministerium für Schule, Wissenschaft und Forschung des Landes Nordrhein-Westfalen): *Zusammenfassender Bericht über die Auswertung des Verfahrens zur Qualitätssicherung im Abitur 2000*; Eigendruck des Ministeriums 2001

NCEE (National Commission on Excellence in Education); 21.10.1999; *A nation at risk: The imperative for educational reform*; U.S. Government Printing Office; Washington DC 1983; Archiv;
<http://www.ed.gov/pubs/NatAtRisk/risk.html>

Neubrand, M.; 11.03.2001; *Informationen zum PISA-Projekt der OECD*;
<http://www.uni-flensburg.de/mathe/homepg/neubrand/bzmpisa.html>

Neubrand, M. et al.; 08.08.2000; *Framework zur Einordnung des PISA Matematik-Tests in Deutschland*;
<http://www.mpib-berlin.mpg.de/pisa/pdfs/LangfassungMathe.pdf>

OECD (Organisation for Economic Cooperation and Development): *Measuring student knowledge and skills. A new framework for assessment*; Paris; OECD Publication Service; 1999

OECD (Organisation for Economic Cooperation and Development); 08.08.2000;
Measuring student knowledge and skills. A new framework for assessment; OECD Publication Service; Paris; über
<http://www.pisa.oecd.org/docs/assess.htm>

OECD (a); 06.12.2000; *Mathematical curricular strands*;
<http://www.pisa.oecd.org/math/curric.htm>

OECD (b); 06.12.2000; *Key features of PISA*;
<http://www.pisa.oecd.org/pisa/feature.htm>

OECD/CERI: (Organisation for Economic Cooperation and Development):
What works in Innovation: The Assessment of School Performance (94, 14); Paris 1994

Rosenthal, R. und Jacobson, L.: *Pygmalion in the Classroom: Teacher Expectation and Pupil's Intellectual Development*; New York; Holt, Rinehart and Winston 1968; auch über
<http://members.aol.com/svennord/ed/labeking.htm>

Rost, J.: *Testtheorie, Testkonstruktion*; Göttingen; Hans Huber Verlag 1996

Sageder, J.; 27.08.2000; *Bildungstheoretische Modelle*;
<http://paedpsych.jk.uni-linz.ac.at/arbeitsblaetter/wissenschaftpaedagogik/Modelle/Bildungstheorie.html>

Schaak, E.; 06.08.2000; *Bemerkungen zur Bildungsdiskussion in Schweden*;
<http://www.schulforum.ch/6ausland/andere/schweden>.

Schmidt, W.H. et al. 1996a: *Characterizing pedagogical flow: An investigation of mathematics and science teaching in six countries*; Dordrecht; Kluwer 1996

Schmidt, W.H. et al. 1996b: *Many visions, many aims: A cross national investigation of curricular intentions in school mathematics*; Dordrecht; Kluwer 1996

Schulforum Schweiz; 31.10.1999; *Startadresse des Schulforums*;
<http://www.schulforum.ch/index.html>

Schweizer Bundesrat; 27.07.2000; *Konkordat über die Schulkoordination*; Edition Optabyte AG;
http://www.zivilgesetzbuch.ch/sr/411.9/411.9_000.htm

SKBF (Schweizerische Koordinationsstelle für Bildungsforschung) *Schweizerische Eidgenossenschaft und die sie konstituierenden Kantone*;
<http://agora.unige.ch/skbf/welc.dt.html>

SKBF 1999a (Schweizerische Koordinationsstelle für Bildungsforschung): *Mehr fördern, weniger auslesen. Zur Entwicklung der schulischen Beurteilung in der Schweiz*; in: Trendbericht SKBF Nr. 3; Aarau; Eigenverlag SKBF 1999

SKBF 1999b (Schweizerische Koordinationsstelle für Bildungsforschung): *Fördern und Auslesen*; Flyer in: Trendbericht SKBF Nr. 3; Aarau; Eigenverlag SKBF 1999

SKBF 1999c (Schweizerische Koordinationsstelle für Bildungsforschung):
Quelle est l'opinion des Suisses sur l'école; Flyer in: Trendbericht SKBF
Nr. 3; Aarau; Eigenverlag SKBF 1999
Stebler, R.: *Mathematisch-naturwissenschaftlicher Unterricht der Sekundarstufe I in der Deutschschweiz*; in: In Mathe Mangelhaft; Berlin; Veröffentlichung der BDA (Bundesvereinigung der Deutschen Arbeitgeberverbände) 2000

Steffens, K. und Perleth, Ch.: *The structure of cognitive abilities in highly and moderately gifted young people*; in: Cropley, A.J. und Dehn, D., (Hrsg.), Fostering the growth of high ability: European Perspectives; S. 113-125; Norwood; NJ; Ablex 1996

Tennenbaum, R.; 01.08.2000; *Bildung zur schönen Menschlichkeit: Wilhelm von Humboldts Bildungsideal*;
<http://www.schiller-institut.de/seiten/humboldt.htm>

TIMSS Mathematics Achievements; 25.04.1999; *Mathematics Achievements in the Middle School Years Executive Summary*, Seite 1; über
<http://timss.bc.edu>

TIMSS; 10.3.1998a; file *BmathAll.pdf*: *Mathematics Achievement in the MiddleSchool Years*; IEA (International Association for the Evaluation of Educational Achievement); TIMSS International Study Center; über
<http://wwwcsteep.bc.edu/timss> oder <http://timss.bc.edu>

TIMSS; 12.03.1998b; file *BMItems.pdf*; IEA (International Association for the Evaluation of Educational Achievement); TIMSS International Study Center; über
<http://wwwcsteep.bc.edu/timss> oder <http://timss.bc.edu>

TIMSS; 13.05.1998c; file *bism0795.pdf*: *Percent of responses by Math Item Categories - seventh grade*; über
<http://wwwcsteep.bc.edu/timss> oder Adresse <http://timss.bc.edu>

TIMSS; 13.5.1998d; file Technical Report, Volume I: *Design and Development*; über
<http://wwwcsteep.bc.edu/timss> oder Adresse <http://timss.bc.edu>

Törner, G.; 11.03.1998; *Stellungnahme der Deutschen MathematikerVereinigung im Rahmen der Anhörung zu TIMSS bei der Kultusministerkonferenz am 26./27.6.1997 in Bonn*;
<http://www.mathematik.uni-bielefeld.de/DMV/archiv/memorada/TIMSS.htm>

- Törner, G. et al.; 11.03.1998; *Erklärung der Fachverbände DMV, GDM, MNU*
19. 02. 1997;
<http://www.uni-duisburg.de/FB11/PRESSE/TIMSS970219.htm>
- Vögeli-Mantovani, U.: *Mehr fördern, weniger auslesen: Zur Entwicklung der schulischen Beurteilung in der Schweiz*; Aarau; Eigenverlag SKBF 1999
- Weigand, H.G.; 12.04.1998; *Überlegungen zur TIMS-Studie, Ergebnisse, Ursachen, Konsequenzen*; in „*Mathematik in der Schule*“; 35 (1997) Heft 10, Pädagogischer Zeitschriftenverlag;
<http://did.mat.uni-bayreuth.de/~matthias/weigand.htm>
- Wieland, G.: *Das Schweizer Zahlenbuch*; in: Beiträge zum Mathematikunterricht 1999, S. 598 ff; Hildesheim; franzbecker 1999
- Winter, W.: *Wissenschaftliche Arbeiten schreiben*; Wirtschaftsverlag Carl Ueberreuter, Frankfurt/Wien 2004
- Wolfe, R.: *Measurement Obstacles to International Comparisons and the Need for Regional Designs and Analysis in Mathematics Surveys*; in: Kaiser; International Comparisons in Mathematics Education; London; Falmer Press 1999
- Wolter, St. und Bauhofer, W.: *Permanente Erhebung über Bildungsforschung / Enquête permanente sur la recherche éducationnelle*; SKBF/CSRE Schweizerische Koordinationsstelle für Bildungsforschung/ Centre Suisse de coordination pour la recherche en éducation; ISSN 1013-6528; Aarau; Eigenverlag der SKBF/CSRE fortlaufend erscheinend
- Wulftange, J: *Ergebnisse der TIMSS aus der Sicht der Schule*; in: TIMSS und der Mathematikunterricht, S.36; Schroedel 1998

9 Aspekte des Schulsystems in der Schweiz

9.1 Vorbemerkungen

Der deutsche Leser der vorliegenden Studie dürfte zumindest mit dem Schulsystem einiger deutscher Bundesländer hinreichend vertraut sein, aber nicht unbedingt mit dem mehrsprachigen Schweizer Schulwesen. Zur vergleichenden Beurteilung der Schulsysteme sowie zum Verständnis länderspezifischer Gegebenheiten, die möglicherweise als soziale Kontexte oder kulturelle Skripte in Schülerleistungen einfließen, wird daher hier ein einführender Überblick über das Schweizer Schulsystem zusammengestellt.

Es ist einleuchtend, dass sowohl die Mehrsprachigkeit wie auch die geografisch-strukturellen Gegebenheiten Randbedingungen erzeugen, die das Schweizer Schulsystem vom deutschen erheblich unterscheiden. Dabei kann die kantonale Struktur vielleicht noch mit der des deutschen föderalen Bildungssystems in den kulturell autarken Bundesländern gleich gesetzt werden, allerdings bleibt die Vielsprachigkeit noch immer ein wesentlich diskriminierender Faktor. Dies sollte bei der Bewertung soziokultureller Aspekte, wie sie auch in dieser Studie mit untersucht wurden, mit in Betracht gezogen werden. Die mit angedeuteten geostrukturellen Gegebenheiten werden im unten folgenden Beitrag ‚It takes a Village‘ deutlich herausgehoben.

Zur Recherche über das Schweizer Schulsystem wurde neben persönlichen Kontakten gezielt das Internet genutzt, um so eine Übersicht über das Schulsystem wie auch aktuelle gesetzgeberische, didaktische oder sonstige Tendenzen zu erkennen. Das Internet erwies sich hinsichtlich des Schweizer Bildungswesens schnell als ausgesprochen reiche Informationsquelle⁴⁴.

Ohne darin eingeschlossene Wertung seien beispielhaft einige Internetseiten erwähnt

Die Seite <http://www.schulforum.ch/index.html> ist Startseite des Schulforums Schweiz. Sie enthält eine umfangreiche Liste von weiteren Adressverweisen. Die folgende Tabelle zeigt Inhalte dieser Startseite (Stand 1.11.1999). Alle Unterstreichungen sind Internet-Links. Allein schon dies macht den Umfang dieser Linkssammlung deutlich.

⁴⁴ So werden mittlerweile die Forschungsinformationen der SKBF/CSRE (Schweizerische Koordinationsstelle für Bildungsforschung) nicht mehr durch Postversand, sondern nur noch über das Internet übermittelt.

Tabelle 31: Schulforum Schweiz

Startseite des „Schulforum Schweiz“	
Ich sage euch nicht: Folgt mir! Vertraut blindlings meinen Worten! Aber ich sage euch: Folgt niemandem blindlings! Vertraut nicht, sondern denkt! Schafft euch in dem Widerstreit der Meinungen mit gewissenhafter Sorge eigene Überzeugung! (Carl Schurz, 1829 – 1906, dt.-amerik. Politiker, 1887 – 1891 Innenminister der USA und Reformer der US-Verwaltung)	
<u>Landauf – Landab</u>	Aktuelle Informationen über das Geschehen in der aargauischen Schulpolitik: Zeitungsmeldungen, Stellungnahmen, Vorträge usw. -- IFNEC-Mailing-list
<u>Leserbriefe und Ratgeber</u>	Leserbriefe an das Schulforum Ratgeber - ...für Eltern - ...Lehrer - ...Politiker Wie kann man sich gegen falsche Reformen zur Wehr setzen?
<u>Schulreformen</u>	Eine Einführung in 11 Lektionen – Hintergründe und Taktik der Schulreformer. Neue Sprache der Schulreformer. Kritik an Reformen. Schule und Internet -- Artikel zum Thema
<u>Aargau</u>	Bildungspolitische Artikel – Themen und Projekte -- Neue Promotion -- MAR -- Zofingen -- Zur Gesamtschule – Die Bildungsverwaltung -- in English

<u>Andere Kantone</u>	Aktuelles -- Übersicht Was läuft im Kanton Basel-Stadt? in Zürich? in Luzern? Was läuft in anderen Kantonen?
<u>Schweiz</u>	Über Institutionen wie die Erziehungsdirektorenkonferenz (samt Kritik an der EDK) -- TIMSS Schweiz - Romandie -- Bildungsindikatoren -- Artikel zur Bildungspolitik
<u>Ausland</u>	Was läuft in Deutschland? -- in England? – in Kanada? Was läuft in den USA -- in anderen Ländern? Kurzfassung
<u>Kopf, Herz und Hand</u>	Zitate und Sprüche und etwas Spaß muss sein
<u>Netzwerk</u>	Schulische Organisationen in der Schweiz Schulische Organisationen im Ausland -- in den USA Literaturhinweise -- Rolle der Medien
<u>Schulforum Inside</u>	Entstehung des Schulforums -- Startartikel und Antworten Frühere Veranstaltungen -- Unsere Erwartungen an die Schule, Beitritt zum Schulforum – Impressum
<u>Update</u>	Verzeichnis der kürzlich aktualisierten Seiten Zur Eröffnung -- Statistik über die Besucherzahlen Werbung auf www.schulforum.ch

Quelle: Schulforum Schweiz⁴⁵ (1999) <http://www.schulforum.ch/index.html>

Drei Beispiele mögen die Fülle der damit erschließbaren Informationen andeuten:

⁴⁵ Zur Einordnung des Schulforums Schweiz in die Bildungslandschaft sei das Eigenzitat auf ihrer Seite angeführt: „Die überwältigende Mehrheit dieser Institute teilt - oftmals schon aus vitalem Interesse die Haltungen des Schulforums Schweiz zur Bildungspolitik nicht.“ (online, Schulforum, 1999)

Der Link auf IFNEC, IFNEC@ADMIN.HUMBERC.ON.CA, führt zu einer sehr aktiven discussion group, die per email an alle Diskussionsmitglieder pädagogisches Geschehen mit Trends, Neuerungen und politischen Aktivitäten weltweit (mit Fokus auf Großbritannien, USA, Kanada, Schweiz, neuerdings auch China und Rußland) verfolgt, berichtet und zur Diskussion stellt.

Hinter dem Link „Hintergründe“ verbirgt sich eine Übersicht über die Entwicklung der Pädagogik in den letzten Jahrzehnten (ab etwa 1950) in Deutschland und der Schweiz. Zitat aus dieser Seite: „Die Schweiz ist insofern mitbetroffen, als dass sie mit einer Verzögerung von einigen Jahren dieselben Entwicklungen nachvollzieht“.

Unter ‚Netzwerk, Schulische Organisationen in der Schweiz‘ findet sich eine umfassende Sammlung von Bildungsinstitutionen der Schweiz, Bildungsforschungsinstituten, schulnahen Organisationen und einigen sonstigen Informationsorganen. Abschnitt 9.2 gibt dazu einen ausführlichen Überblick.

Daneben sind über das Internet auch viele Bildungsinstitutionen auf kantonaler Ebene erreichbar. So findet man beispielsweise unter

<http://www.bern.ch/direktionen/schuldirektion/schuldirektion.cfm>
die Schuldirektion Bern, unter
<http://www.bern.ch/direktionen/schuldirektion/schulamt.cfm>
das Schulamt Bern, unter
<http://www.sla.unibe.ch/> als Startadresse das
Sekundarlehramt an der Universität Bern⁴⁶ und die
Forschungsstelle für Schulpädagogik und Fachdidaktik (FSF)

jeweils mit umfassenden Informationsangeboten.

Schweizer Schulen sind ebenfalls mit umfassenden Eigendarstellungen im Web vertreten.

Unter <http://edkwww.unibe.ch/d/euridice/euri-Vorwort.html> als Startadresse findet sich eine Publikation angelehnt an die Vorgaben der Europäischen Informationsstelle Eurydice, in der das Schweizer Schulsystem bis in den universitären Bereich hinein beschrieben wird. Die umfangreiche Beteiligung verschiedenster im Bildungssystem tätiger Organisationen an diesem Bericht lässt vermuten, dass die Darstellungen der Publikation nicht einseitig gefärbt sind und ein im Wesentlichen zutreffendes Bild des Schweizer Bildungssystems geben. Die fol-

⁴⁶ Leitung: Prof. Dr. Hans Badertscher, Vizedirektorin: Dr. Tina Hascher,
Vizedirektor Didaktik der Mathematik: Prof. Dr. Armin Hollenstein

genden Darstellungen lehnen sich daher eng an diese Publikation an, einige Stellen sind wörtlich zitiert.

Der in Kopf der Informationen zum Schulforum Schweiz aufgenommene Satz von Carl Schurtz deutet eine Grundhaltung der Designer der website zur Erziehung an, die ähnlich auch in dem folgenden Artikel durchscheint. Er wurde am 16.9.1998 im o.a. Internet-Diskussionsforum IFNEC veröffentlicht. Sein Verfasser, ein Schweizer Lehrer, beschreibt den Unterricht und seine Einbettung in die soziale Struktur im Kanton Aargau. Da sich damit ein weit reichender Einblick in das Umfeld und die Rahmenbedingungen von Schule und Unterricht sowie Unterrichtsorganisation im Rahmen der Schweizer Demokratie gewinnen lässt, sei der Beitrag hier vollständig zitiert.

„It takes a Village.

In Switzerland, even the tiniest hamlet has its own school.

In Aargau, a Swiss canton, with about 90,000 students, we don't have any school buses. Furthermore, because Aargau does not have any large cities, our students are distributed among some 230 villages spread out over 550 square miles (about the area of Metropolitan Toronto). Yet somehow, despite the lack of buses the children all get to school. How?

The answer is that each village has its own kindergarten and primary school (grades 1-5). Even villages with only 70 inhabitants have their own primary schools. In this way, the distance between home and school is kept so small that most pupils can get to school on foot or by bike.

And these are not one-room schools, either. Every grade has its own class, even if there are only nine or ten pupils. Split grades are very rare, occurring only in a few schools in the Jura hills.

Even the secondary schools (grades 6-9) are located in every other village. Secondary students travel to school by foot, bicycle, small motor-scooter, public bus or train (paid for by parents).

Every village has its own elected school board to oversee its school(s). Typically seven in number, the trustees are average people with day jobs who earn something like \$1500 a year for their service. They hire the teachers (usually for a four-year period, renewable), validate the teachers' report cards, set the budget, and field parents' complaints.

Surprisingly, the Swiss policy of locating at least one school in every village costs less than the North American policy of huge, consolidated schools. The trick is to minimise the non-classroom expenses. In Switzerland the goal is to spend almost all the money in the classroom.

To begin with, of course, there are no transportation costs (an expense that eats up almost five per cent of the Ontario education budget). Other savings are achieved by ensuring that almost every adult in the school is a classroom teacher. The Swiss don't have principals, secretaries, librarians, enrichment or remedial teachers, lunchroom supervision (most kids go home for lunch) or crossing guards.

Instead of principals, we have a "rektor" to do the administration and an "inspector" to provide instructional supervision. Both the rektor and the inspector (who comes from another school) continue to teach at least half time, and both positions are term appointments. Other duties are shared among the teachers.

For example, one teacher prepares the timetable every year. Another is responsible for ordering textbooks and supplies, as well as paying the bills. Various teachers take

charge of science materials, the library, the gym equipment, French tapes and books, the video cassettes, the musical instruments and so on. All these appointments are made by the teachers themselves in the course of their meetings. In fact, all major decisions are made by the staff - who teaches which class what subject, which pupil gets punished how, and so forth. Switzerland is a democracy, you see.

In most schools, the classroom teachers teach most subjects, although the top stream secondary schools are on a rotary system with subject specialists. Enrichment and remedial teachers are unknown, as it is considered important to keep primary students at the same level as much as possible.

When the children enter grade 6, they are separated into one of three different streams, although it is possible, and not at all uncommon, for students to switch streams. Children with special needs (such as blindness, mental disability, etc.) attend inter-cantonal residential schools.

Swiss teachers are very well paid, even more than Canadian teachers, and they are highly respected in their communities. Of course they work hard (an average 50 hours a week during the school year) and do a highly professional job, as proven by Swiss students' results on international comparisons. Parental satisfaction is evident in the almost total absence of private schools.

Despite the teachers' high pay and the presence of schools in every village, the Swiss educational system actually costs less than its Canadian counterpart, according to the OECD.

If the government threatened to close a small school in one of the villages, I think there would be a revolution. These schools are important to their communities, and vandalism is almost unknown.

Many of the teachers live in the villages themselves, and everyone benefits from the personal relationships that develop among parents, teachers and students. Most conflicts are solved at the lowest level, just by communicating with each other. It is this kind of face-to-face accountability that is at the heart of effective teaching.

You know what Hillary Clinton says, "It takes a whole village to raise a child." Well, in Switzerland, we all pitch in to contribute to the well-being of our young people. And it works!" (Achermann, 1998)

Dies deutet darauf hin, dass soziale Skripte hier einen nicht zu unterschätzenden Einfluss auf Struktur und Realisierung von Schule haben. Damit könnten die daraus entstehenden Ausprägungen von Unterricht und auch die Mathematikleistungen im Vergleich Schweiz/Deutschland eine nicht unwesentliche Rolle spielen

Einen Hinweis auf das mit den oben zitierten Worten „Switzerland is a democracy, you see“ angedeutete soziale Skript gab Wieland in einem Vortrag zur Genese der Schweizer Version des Schulbuches „Das Zahlenbuch 5“ (Affolter, 1998): das Buch wurde unter intensiver Beteiligung Schweizer Mathematiklehrer begutachtet, erprobt und national bis in die Details adaptiert, und eine solche Adaption unter Beteiligung Schweizer Lehrer und auch Eltern ist nach Wieland Bedingung sine qua non für ein Lehrwerk, welches in der Schweiz eine Chance auf Einführung haben will. (Wieland, 1999)

Die folgenden Abschnitte geben eine Übersichtsskizze über die Schweiz und ihr Bildungswesen.

9.2 Ausgewählte demographische Daten zur Schweiz

Die Schweizerische Eidgenossenschaft wurde 1291 gegründet und ist seit 1848 ein Bundesstaat. Sie setzt sich aus 26 Kantonen zusammen, die in bestimmten Bereichen autonom sind; diese sind ihrerseits in ebenfalls teilautonome Gemeinden unterteilt (im ganzen ungefähr 3000), welche ebenfalls über eine gewisse Autonomie verfügen.

In der Schweiz wohnen auf einer Fläche von 41300 km^2 7 Millionen Einwohner, davon mehr als die Hälfte in Städten mit einer Einwohnerzahl von 10.000 oder mehr.

Es gibt die Landessprachen Deutsch 63,7 %, Französisch 19,2 %, Italienisch 7,6 %, Rätoromanisch 0,6 %, andere 8,9 %. Amtssprachen auf Bundesebene sind Deutsch, Französisch und Italienisch.

Der Ausländeranteil an der Wohnbevölkerung beträgt 18,9 %, das sind 28 % der Erwerbstätigen.

Das Land hat keine Rohstoffe; die Wirtschaft hängt daher wesentlich vom Ausland ab.

Es gab viele Jahre Vollbeschäftigung, seit 1990 erreicht die Arbeitslosenquote jedoch immer neue Rekordwerte (im Juni 1995 betrug diese auf die ganze Schweiz bezogen 4 %, in einigen Kantonen der Westschweiz wurden sogar Arbeitslosenquoten von bis zu 7 % verzeichnet)

Die Grundsätze der direkten Demokratie sind fest verankert auf eidgenössischer Ebene, sogar noch ausgeprägter in Städten und Gemeinden, wenn auch die Strukturen historisch und auch größenbedingt z.T. sehr unterschiedlich sind. Es gibt daher regelmäßige Urnengänge zur Wahl der Behörden oder zur Ausübung des Initiativ- oder Referendumsrechts.

Dezentralisierung und Demokratie sind wesentliche Merkmale des politischen Systems und des Bildungssystems. Die politische und administrative Organisation der Kantone entspricht im Wesentlichen der des Bundes: Legislative, Exekutive, Judikative, finanz- und fiskalpolitische Autonomie.

9.3 Verfassungsmäßige Verankerung

Es gilt der Grundsatz der Bundesverfassung, Artikel 27 Abs.3:

„Die öffentlichen Schulen sollen von den Angehörigen aller Bekenntnisse ohne Beeinträchtigung ihrer Glaubens- und Gewissensfreiheit besucht werden können.“

Da jeder Kanton über seine eigene Schulgesetzgebung verfügt, ist die Festlegung der Aufgaben, welche die Schule zu erfüllen hat, in der Schweiz nicht einheitlich geregelt. Ein Beispiel einer kantonalen Grundsatzregelung ist etwa:

"Die Schule fördert die harmonische Entwicklung der Fähigkeiten des jungen Menschen unter Berücksichtigung der demokratischen Tradition der westlichen Kultur, weckt in ihm die Bereitschaft zu Toleranz, das Verantwortungsbewusstsein gegenüber seinen Mitmenschen und der Umwelt sowie den Respekt vor anderen Sprachen und Kulturen und vermittelt dem Schüler jene Kenntnisse und Fähigkeiten, die ihm den Zugang zu einer Berufsausbildung und zur Fortbildung ermöglichen."

Die Aufteilung der Kompetenzen ist aufgrund des föderalistischen Systems der Schweiz im Bildungsbereich relativ komplex. Der letzte Versuch, einen Bildungsartikel in die Verfassung aufzunehmen, ist 1973 in einer Volksabstimmung knapp gescheitert. Die Kantone sind daher im Wesentlichen souverän, und die kantonalen Gesetzgebungen unterscheiden sich beträchtlich voneinander. Lediglich einige ‚wichtige‘ Bereiche unterliegen zum Zwecke der Einheitlichkeit entsprechenden Bestimmungen des Bundes.

Zu den Bundeskompetenzen gehört die Organisation eines "genügenden Primarunterrichts", der obligatorisch und unentgeltlich ist. Weiter erlässt der Bund Vorschriften über die berufliche Ausbildung, regelt den Turn- und Sportunterricht, ist Träger der Eidgenössischen Technischen Hochschulen in Zürich und Lausanne, des Schweizerischen Instituts für Berufspädagogik (Bern, Lausanne, Lugano) sowie der Eidgenössischen Sportschule in Magglingen, regelt den Zugang zum Medizinstudium sowie zu den Eidgenössischen Technischen Hochschulen, gewährt Beiträge an die kantonalen Universitäten, an die wissenschaftliche Forschung und fördert die Ausbildung und Integration von behinderten Kindern und Jugendlichen:

Im Universitätsbereich: unterstehen neun Bildungsinstitutionen den Kantonen, während die beiden technischen Hochschulen dem Bund unterstellt sind. Jede Universität wird vom entsprechenden kantonalen Erziehungsdepartement verwaltet, verfügt aber über eine weitreichende akademische Freiheit, die im übrigen durch das kantonale Recht garantiert wird. Die beiden technischen Hochschulen jedoch fallen vollständig in die Zuständigkeit des Bundes.

9.4 Interkantonale Zusammenarbeit

Aufgrund der unterschiedlichen Ausgangslagen und der Vielfalt der Zuständigkeitebenen liegt das Hauptproblem in der Kohärenz des gesamten föderalistischen Systems aus 26 Kantonen mit vier Hauptsprachen.

Bis zum Zweiten Weltkrieg waren die kantonalen Systeme sehr isoliert, obwohl die Vorsteher der kantonalen Erziehungsdepartemente (die Erziehungsdirektoren der Kantone) bereits 1897 eine Konferenz gegründet hatten, welche den gegenseitigen Informations- und Erfahrungsaustausch sowie die Koordination der verschiedenen Schulsysteme auf gesamtschweizerischer Ebene zum Ziel hatte.

In den sechziger Jahren nahm das Bedürfnis nach einer verstärkten Koordination der Bildungspolitik zu.

Die Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) wurde umgestaltet, es wurde ein Sekretariat eingerichtet, zusätzlich wurden vier Regionalkonferenzen geschaffen: Westschweiz und Tessin, Nordwestschweiz, Innerschweiz, Ostschweiz).

Wichtiges rechtliches Instrument zur Koordination ist das „Konkordat über die Schulkoordination“ vom 29.10.1970. Es ist auch ein Instrument der freiwilligen Zusammenarbeit und Harmonisierung bei Reformen, Anerkennung von Diplomen, Zusammenarbeit im Bereich der Planung, Forschung und Schulstatistik etc.:

„Die Konkordatskantone bilden eine interkantonale öffentlich-rechtliche Einrichtung zur Förderung des Schulwesens und zur Harmonisierung des entsprechenden kantonalen Rechts.“ (online, Schweizer Bundesrat, , 2000)

Bisher sind 25 Kantone diesem Konkordat beigetreten.

Es gibt eine Angleichung bezüglich folgender Rahmenbedingungen:

- Schuleintrittsalter: vollendete 6. Altersjahr mit Stichtag 30. Juni
- Schulpflicht: neun Jahre, bei mindestens 38 Schulwochen pro Jahr
- Ausbildungsdauer bis zur Maturität: mindestens 12, höchstens 13 Jahre
- Schuljahresbeginn: zwischen Mitte August und Mitte Oktober.

Wegen der direkten Demokratie (Referendum) brachte die Umsetzung dieser Grundsätze beträchtliche Schwierigkeiten mit sich, so dass der Anpassungsprozess über 15 Jahre dauerte.

9.5 Verwaltung

Auf Bundesebene werden die Kompetenzen von zwei verschiedenen Departementen (Ministerien) wahrgenommen: vom Eidgenössischen Departement des Innern (Eidgenössische Technische Hochschulen, Subventionierung der Universitäten, Stipendien, Wissenschaft und Forschung, Sport, Anerkennung der gym-

nasionalen Maturitätsausweise) und vom Eidgenössischen Volkswirtschaftsdepartement (Berufsbildung).

Innerhalb des Eidgenössischen Departements des Innern wurden die entsprechenden Aufgaben der Gruppe für Wissenschaft und Forschung (GWF) und dem Bundesamt für Bildung und Wissenschaft (BBW) übertragen; außerdem wurden zwei weitere Koordinationsorgane geschaffen, der Schweizerische Wissenschaftsrat (SWR) und die Schweizerische Hochschulkonferenz (SHK).

Für das kantonale Schulwesen gilt: Leitung und Verwaltung obliegen der kantonalen Exekutive und deren Erziehungsdepartement (dem Erziehungsministerium des Kantons), zu welchem in mehreren Kantonen noch ein Erziehungsrat hinzukommt. Der für das Erziehungsdepartement zuständige Regierungsrat, der kantonale „Erziehungsminister“, wird vom Volk gewählt und muss sich alle vier bis fünf Jahre zur Wiederwahl stellen.

Die Schulverwaltung auf Gemeindeebene ist je nach Größe der Gemeinde sehr unterschiedlich organisiert; die politischen Behörden werden hier von Schulkommissionen unterstützt, die sich aus gewählten lokalen Politikern, aber auch aus Laien zusammensetzen.

Allgemein lässt sich feststellen, dass in den meisten Kantonen die Schulfragen bis ins Detail von einer zentralen Stelle geregelt werden (Größe der Klassen, verbindliche Lehrpläne, offizielle Lehrmittel usw.), insbesondere im Bereich der obligatorischen Schule. In der Sekundarstufe II hingegen ist dies nicht mehr der Fall.

9.6 Schulaufsicht

Zur Kontrolle der Schulen:

Sie wird realisiert durch vollzeitlich angestellte Beamte und Beamtinnen, die sogenannten Schulinspektorinnen und -inspektorinnen. Diese sind Generalisten und keine Spezialisten für einzelne Fächer und tragen meist die Verantwortung für eine bestimmte Region (Bezirk oder Amt); in einigen Kantonen werden sie teilzeitlich von Inspektoren und Inspektorinnen aus dem Lehrkörper unterstützt. Es gibt Laieninspektorate in den Kantonen Zürich und St. Gallen.

Zur Schulleitung in der Sekundarstufe I:

Diese Funktion besteht nur in Schulen mit einer grossen Anzahl von Klassen. In der Deutschschweiz existiert kein eigentlicher Schuldirektor. In der Westschweiz

hat der directeur d'école eine bedeutend wichtigere Rolle: er nimmt Aufgaben im Bereich der Überwachung und Einführung von pädagogischen Neuerungen wahr.

Zur Schulleitung in der Sekundarstufe II:

Hier ist die Schulleitungsfunktion in der ganzen Schweiz von Bedeutung, da für die Überwachung der Maturitätsschulen kein Inspektorat besteht; die Schweizerische Maturitätskommission prüft, ob die erforderlichen Voraussetzungen für die Anerkennung der Ausweise gemäß den reglementarischen Bestimmungen gegeben sind.

In der Vergangenheit erwartete man von Inspektoren und Schulleiterinnen in erster Linie, dass sie eine Kontrollfunktion (Lehrpläne, Inhalte, pädagogische Methoden, Evaluation, Promotion (das ist die Versetzung in die nächste Klasse) der Schüler und Schülerinnen usw.) ausübten und dass sie bei der Umsetzung von schulischen Neuerungen deren Übereinstimmung mit den kantonalen Vorschriften überprüften.

Diese Aufgaben bestehen teilweise zwar weiterhin, doch das Inspektorat hat sich, in Zusammenarbeit mit weiteren Fachleuten, in Richtung pädagogischer Animation und der Beratung fortentwickelt.

9.7 Finanzierung

- Jede Zuständigkeitsebene ist im steuerlichen Bereich autonom und trägt deshalb die finanziellen Belastungen im Rahmen ihrer Kompetenzen.
- Die obligatorische Schule ist unentgeltlich, erst ab Sekundarstufe II gibt es eine Beteiligung an den Kosten (Schulgeld, Schulbücher usw.)
- Die Ausgaben werden getragen von Gemeinden (Gebäude, Einrichtung, Lehrmaterial, Teil der Besoldung des Lehrkörpers) und den Kantonen (Besoldung).
- Die Kosten für die Ausbildung in der Sekundarstufe II und für die Hochschulstudien (neun Hochschulen sind kantonal) werden größtenteils von den Kantonen übernommen.
- Der Bund beteiligt sich bei Berufsbildung, akademischer Ausbildung – vollständig für die beiden Eidgenössischen Technischen Hochschulen - und für die Forschung.

9.8 Beratende Organe

- Mehrere Kantone verfügen über einen Erziehungsrat.
- In allen Kantonen hat der Lehrkörper sowohl rechtlich als auch praktisch ein Mitspracherecht, nicht nur auf Anhörung, sondern auch – insbesondere in der Deutschschweiz – auf direkte Mitbestimmung (Mitgliedschaft im Erziehungsrat und in den lokalen Schulbehörden).
- Das Mitspracherecht der Eltern ist weniger gut verankert, auch wenn diese in den Schulkommissionen sehr oft vertreten sind.
- Auf Gemeindeebene haben beispielsweise im Rahmen der direkten Demokratie alle Bürger und Bürgerinnen die Möglichkeit, über das Schulsystem zu entscheiden, wodurch der Einfluss der Eltern als eigentlich Betroffene eingeschränkt wird.
- Auf kantonaler Ebene verfügen die Parlamente (Gesetze, Kredite) und die Bürger und Bürgerinnen (Referendum, Initiative) ebenfalls über einen beträchtlichen Einfluss auf die Organisation des Schulsystems, ebenso wie die Gemeinden, die mitsprachberechtigt sind.

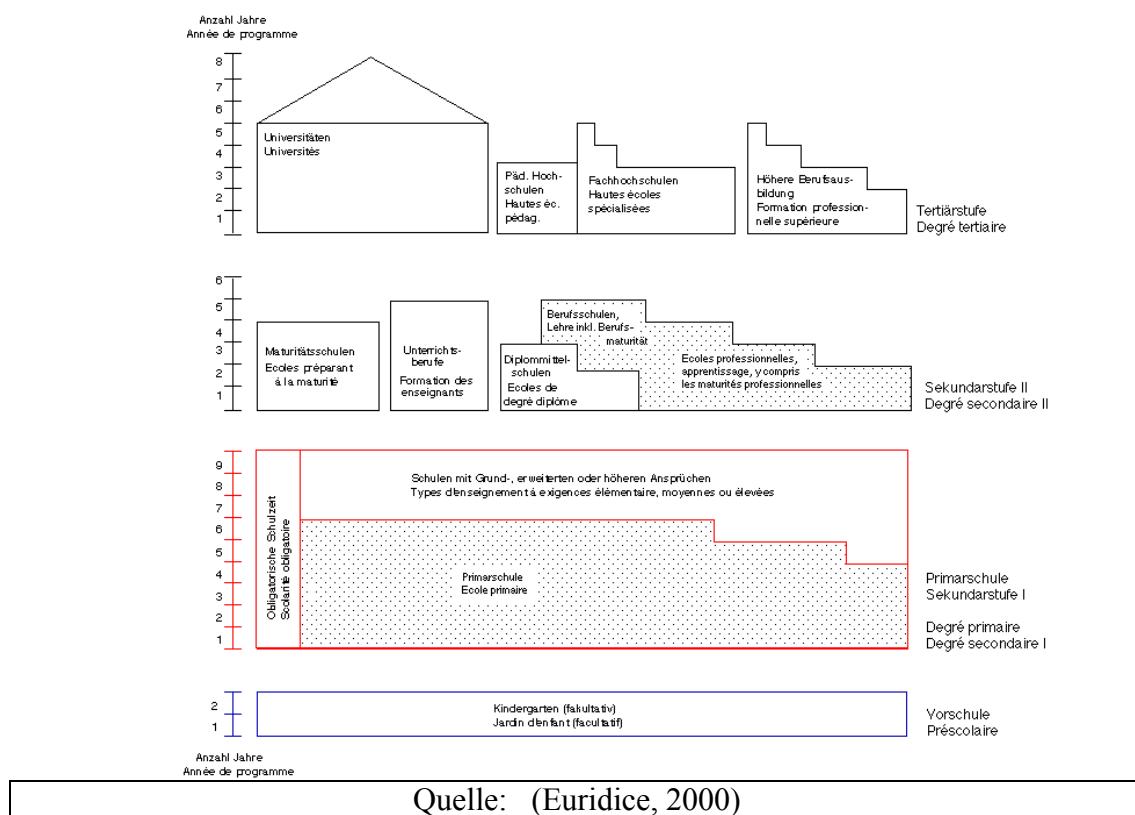
9.9 Obligatorische Schule

In der Schweiz dauert die obligatorische Schule neun Jahre und wird von Kindern im Alter von 6/7 bis 15/16 Jahren besucht; die meisten Kantone bieten außerdem ein zehntes Schuljahr an.

Die obligatorische Schule umfasst in allen Kantonen die Primarschule und die Sekundarstufe I.

9.10 Übersicht Bildungswesen

Abbildung 9: graphische Übersicht Bildungswesen Schweiz



Quelle: (Euridice, 2000)

9.11 Primarschule

Die Kantone tragen die Verantwortung für die obligatorische Schule zusammen mit den Gemeinden. Sie sind zuständig für Organisation und Finanzierung des Primarunterrichts.

In jedem Kanton gibt es eine eigene Schulgesetzgebung, in der die Kompetenzen definiert sind und die Organisation des Unterrichts für alle Schulstufen festgelegt ist.

Zum Schuleintrittsalter:

Das Schuleintrittsalter ist auf das vollendete 6. Altersjahr festgelegt. Stichtag ist der 30. Juni, Abweichungen im kantonalen Recht bis zu vier Monaten vor und nach diesem Datum sind möglich.

Zur Dauer der Primarschule:

In zwanzig Kantonen erstreckt sich die Primarstufe über sechs Jahre, in vier Kantonen über fünf Jahre und in zwei Kantonen über vier Jahre. Am Ende der

Primarschulzeit liegt das Alter der Schüler und Schülerinnen im Durchschnitt bei 12 Jahren und 7 Monaten.

Zur Struktur der Primarschule:

Sie hat in allen Kantonen eine einheitliche Struktur – vorbehaltlich allfälliger Reformen in einzelnen Kantonen. Sie ist nicht weiter unterteilt.

Zur Größe der Schulen:

Die Größe reicht von Einzelklassen in ländlichen Gebieten bis zu Schulen mit mehreren hundert Schülern in Städten. Einige Primarklassen können als Übungsklassen einer Lehrerbildungsinstitution angegliedert sein. 1994/95 umfasste eine Primarklasse im Durchschnitt 20 Schüler. Ungefähr 80 % der Klassen sind nach Jahrgängen zusammengesetzt.

Zur Unterrichtsdauer:

Im gesamtschweizerischen Durchschnitt werden in der ersten Primarklasse 3 3/4 Stunden pro Tag und ungefähr 20 Lektionen pro Woche unterrichtet; in der 5. oder 6. Primarklasse (im letzten Jahr an der Primarschule) sind es insgesamt 5 1/4 Stunden pro Tag und 34 bis 36 Wochenlektionen.

Zur Unterrichtszeit:

Es gibt Unterricht am Vor- und Nachmittag. Je nach Kanton sind der ganze Samstag oder der Samstagnachmittag und ein weiterer Halbtag während der Woche oder der Samstagnachmittag und ein weiterer ganzer Tag während der Woche schulfrei.

Im allgemeinen kehren die Schüler und Schülerinnen über Mittag nach Hause zurück. Falls dies aufgrund der Distanzen nicht möglich ist, organisieren die Gemeinden die Betreuung (Aufsicht und Mittagessen). Seit einigen Jahren werden insbesondere in städtischen Gebieten Anstrengungen unternommen, betreute Mittagstische und/oder Tagesschulen einzurichten.

Zum Schuljahr:

Ein Schuljahr dauert je nach Kanton zwischen 36,5 (Tessin) und 40 Wochen. Die längsten Ferien sind im Sommer (fünf bis neun Wochen).

Aspekte wie Curriculum, Evaluation, Lehrkräfte und Schulstatistik der Primarschulen werden hier nicht weiter verfolgt. Zwar ist die Primarschule die Basis des weiterführenden Schulsystems, und Art sowie didaktische und methodische Organisation des Unterricht auf dieser Stufe sind wohl mit prägend für den weiterführenden Unterricht in der Sekundarstufe I und damit vielleicht auch ein Bedingungsfaktor für die Ergebnisse von TIMSS Population 2, dies war jedoch kein Forschungsgegenstand der vorliegenden Studie.

9.12 Sekundarstufe I

Bei der Organisation der Sekundarstufe I wird die Vielfalt der kantonalen Schulsysteme besonders deutlich. Sie ist Teil der neunjährigen obligatorischen Schulzeit, ihre Dauer somit abhängig von der Dauer der Primarschule. In den meisten Kantonen beträgt sie zur Zeit drei Jahre (Schüler zwischen 12/13 und 15/16 Jahren), in den übrigen Kantonen vier (Schüler zwischen 11/12 und 15/16 Jahren) oder gar fünf Jahre (Schüler zwischen 10/11 und 15/16 Jahren).

Die Aufnahmekriterien sind von Kanton zu Kanton unterschiedlich. Sie werden orientiert an Prüfungen, dem Zeugnis, Gesprächen mit der Primarlehrkraft, auch der Ansicht/dem Wunsch der Eltern, oft auch kombiniert.

In beinahe der ganzen Schweiz (Ausnahmen Genf, Tessin, teilweise im Wallis) ist die Sekundarstufe I in zwei bis vier Schultypen oder Abteilungen mit unterschiedlichen Anforderungen gegliedert und wirkt daher selektiv. Man findet Schultypen mit Grundansprüchen (Realschule; in der Westschweiz je nach Kanton: section pratique, moderne, préprofessionnelle, classes à options). Dieser Schultyp bereitet auf weniger anspruchsvolle ‚Berufslehren‘ vor. Er wird von ungefähr einem Drittel aller Schüler besucht, wobei der Jungenanteil höher als jener der Mädchen ist. Daneben gibt es Abteilungen mit erweiterten Anforderungen. Sie werden besucht von zwei Dritteln der Schüler der Sekundarstufe I. Sie sind in den meisten Kantonen unterteilt in zwei weitere Typen, einen Typ mit gehobenen Ansprüchen (Untergymnasium, section prégymnasiale) und einen Typ mit mittleren Ansprüchen (Sekundarschule oder Progymnasium, section générale).

Einige Kantone haben sich auf dieser Stufe für einen einzigen Schultyp mit gemischten Anforderungen (cycle d'orientation) entschieden. Das Tessin hat eine Gesamtschule (scuola media) geschaffen, in der die Anforderungen nicht auf der Ebene der Klasse, sondern auf jener der Fächer differenziert werden.

Die Schultypen mit gehobenen Ansprüchen bereiten im Allgemeinen auf längere Studien vor (Maturitätsschule, Kollegium, Gymnasium), jene mit mittleren Anforderungen grundsätzlich entweder auf kürzere Studien in der Sekundarstufe II oder häufiger auf anspruchsvollere Berufsausbildungen.

Versuche zur Einführung einer Gesamtschule sind nicht über die Testphase hinausgekommen (Ausnahmen einzig Genf, Tessin und Wallis). Gleichwohl verwirklichen mehrere Kantone einige der Anliegen der Gesamtschule wie etwa die Verbesserung der Durchlässigkeit (Übertritt von einem Schultyp in den anderen), Individualisierung des Unterrichts durch die Verringerung der Schultypen, Einführung von Wahl- oder Niveakursen oder gemeinsamen Unterricht in bestimmten Fächern für Schüler und Schülerinnen aus verschiedenen Schultypen.

Da die Sekundarstufe I obligatorisch ist, eine Ausbildung auf Sekundarniveau bietet und auf die Berufswahl vorbereiten soll, entsteht oft ein Zielkonflikt: die Schüler und Schülerinnen sollen einerseits eine Wahl treffen können, die noch alles offen lässt, d.h. die Berufswahl hinausschieben können, und andererseits sollen sie auf die Anforderungen der Sekundarstufe II vorbereitet werden.

Der Grundsatz der Demokratisierung der Bildung, wonach die gesamte Bevölkerung eine Ausbildung auf Sekundarstufenniveau erhalten soll, steht oft im Gegensatz zum Selektionsprinzip der gymnasialen Ausbildungen der Sekundarstufe II. Die Vielfalt der kantonalen Lösungen zeugt davon, wie schwierig es ange-sichts dieser schulischen und gesellschaftlichen Herausforderungen ist, ein all-seits akzeptiertes Gleichgewicht zu finden.

9.13 Übergänge, Abschlüsse

Die folgenden Daten zu Schultypen und Übergängen folgen den Darlegungen des Vortrags „Mathematisch-naturwissenschaftlicher Unterricht der Sekundarstufe I in der Deutschschweiz“ anlässlich der BDA-Tagung“ TIMSS – Was machen die anderen besser“ am 12.10.1999, veröffentlicht in (Stebler, 2000).

In der Deutschschweiz besuchen die Kinder ab dem Schulpflichtalter von 7 Jahren für fünf bis sechs Jahre die Primarstufe, wobei die durchschnittliche Klassengröße 20 Schüler beträgt. Dementsprechend wird die Sekundarstufe I ab einem Alter von etwa 12 Jahren besucht. Dabei selektieren mehrheitlich die abgebenden Schulen je nach Leistungsstand aus den Ergebnissen der Orientierungsarbeiten in den beiden vorhergehenden Schuljahren, dem Lehrerurteil, der Elternmeinung und dem Schülerwunsch. Die definitive Zuteilung erfolgt nach einer Probezeit (Vögeli-Mantovani, 1999). Die Benennung der weiterführenden Schulen unterscheidet sich deutlich von den deutschen Bezeichnungen Hauptschule, Realschule und Gymnasium im gegliederten Schulsystem; sie sind *Realschulen* für Schulen mit Grundansprüchen, *Sekundarschulen* für Schulen mit erweiterten Ansprüchen (I) und *Progymnasium* für Schulen mit nochmals erweiterten Ansprüchen (II).

Die Zuteilung zu den Schultypen variiert von Kanton zu Kanton und beträgt bei spielsweise in den Kantonen

Tabelle 32: Zuteilung zu den Schultypen Schweiz

	Realschule	Sekundarschule	Progymnasium
Bern	49 %	43 %	8 %
Basel-Land	33 %	41 %	26 %

Der Ausländeranteil bedarf bei einem Verbund aus 26 Kantonen mit vier Hauptsprachen einer eigenen Definition. Stebler spricht daher alternativ von einem „Anteil der Kinder aus fremden Kulturen“ und definiert: „Kulturell fremde Schülerinnen und Schüler sprechen eine von der Unterrichtssprache verschiedene Muttersprache und/oder haben eine ausländische Nationalität.“ (Stebler, 2000).

Für die meisten Real- und Sekundarschüler beginnt nach Abschluss der obligatorischen Schulzeit (Alter ca. 16 Jahre) eine Berufslehre, während die Progymnasiasten ins Gymnasium als separater, sich anschließender Schulform eintreten. Sie erhalten mit 19 Jahren die Matura, die zum Hochschulstudium berechtigt.

Die Hochschulabschlussquote – bezogen auf die 27-jährige ständige Wohnbevölkerung – lag 1998 bei 6 %.

9.14 Curriculum

Die Festlegung des Lehrplans obliegt den Kantonen; sie ist teilweise auf interkantonaler Ebene koordiniert (z.B. Westschweiz). Die Wahl des Lehrmaterials (Lehrbücher) erfolgt grundsätzlich auf Kantonsebene, oft haben die Lehrkräfte diesbezüglich einen gewissen Spielraum. In allen Schultypen gibt es Unterricht in

- der Muttersprache,
- Mathematik,
- einer zweiten Landessprache,
- Naturwissenschaften,
- Geographie, Geschichte und Staatskunde,
- musischen Fächern (Musik, Zeichnen) und
- Turnen.

In Schultypen mit Grundansprüchen werden ausserdem manuelle Fähigkeiten geschult; in den Typen mit erweiterten Ansprüchen wird etwa eine dritte Sprache (meist Englisch) angeboten, Buchhaltung, Maschinenschreiben oder Technisches Zeichnen, in den Klassen des Untergymnasiums zusätzlich noch Latein und Griechisch.

Bestimmte Fächer gibt es als Wahlpflichtfächer oder Wahlfächer. Die Zahl der Unterrichtsstunden liegt zwischen 30 und 35 Lektionen pro Woche, die Unterrichtsmethoden sind sehr vielfältig,

Didaktische Reformen der letzten Jahrzehnte haben sich stärker in der Primarstufe als in der Sekundarstufe ausgewirkt, dort verbunden mit erhöhter Lehrerfluktuation.

9.15 Evaluation

Auf gesamtschweizerischer Ebene findet keine Evaluation statt. Einige Schulen organisieren ‚gemeinsame‘ schriftliche Prüfungen für alle Klassen; einige Kantone führen derartige Prüfungen an allen Schulen durch. Die einzelnen Lehrkräfte sind für eine laufende Beurteilung verantwortlich. Die Notenskala reicht normalerweise von 1 bis 6 (sehr gut), der Durchschnitt liegt bei 4.

Die Promotion (die Versetzung) erfolgt aufgrund der Resultate des ganzen Schuljahres (genügender Gesamtdurchschnitt und oft zusätzlich genügender Durchschnitt in den Hauptfächern Muttersprache, zweite Sprache, Mathematik). Wird ein Schüler nicht in die höhere Klasse versetzt, muss er das Jahr wiederholen oder in einen Schultyp mit weniger hohen Anforderungen wechseln.

Nicht überall ist am Ende der Sekundarstufe I ein Examen vorgesehen. Einige Kantone geben auf der Grundlage einer schriftlichen/mündlichen Prüfung in den Hauptfächern einen Ausweis über den Abschluss der Sekundarschulbildung ab, andere hingegen ohne Prüfung.

Die verschiedenen Schultypen beeinflussen die weiteren Ausbildungs- und Zugangsmöglichkeiten zur Sekundarstufe II: die Schüler des Untergymnasiums können prüfungsfrei auf eine Maturitätsschule überwechseln, sofern sie die Promotionsbedingungen erfüllen; die Schüler in den Schultypen mit mittleren Anforderungen müssen für den Übertritt an ein Gymnasium eine Aufnahmeprüfung absolvieren.

Auch für den Eintritt in die Lehrerseminare werden systematisch Aufnahmeprüfungen durchgeführt, ebenso für bestimmte Berufsausbildungen (z.B. Banklehrer).

Nach Abschluss der Sekundarstufe I besuchten (90,6 % der 17jährigen Schüler eine weiterführende Schule, davon 22,5 % an einer Maturitätsschule, 2,4 % an einer Lehrerbildungsanstalt, 7,2 % an einer Diplommittelschule (DMS) oder an einer anderen allgemeinbildenden Schule und 67,9 % absolvierten eine Berufsausbildung Stand (1994/95).

9.16 Lehrkräfte

Die Anforderungen an die Ausbildung der Lehrkräfte, ihre Arbeitsbedingungen (Gehalt, Anzahl der Unterrichtsstunden usw.) und ihr Status (fest angestellt, provisorisch angestellt, Stellvertreter, Hilfslehrkraft usw.) werden von den Kantonen festgelegt.

Die Verpflichtung zum Besuch von Fortbildungen ist nicht einheitlich geregelt, aber eine Größenordnung von 3 bis 4 Tagen im institutionellen Rahmen und etwa doppelt so lange auf individueller Basis pro Schuljahr ist realistisch. Die Veranstaltungen sind in der Regel gratis und finden mehrheitlich in der unterrichtsfreien Zeit statt. (nach Stebler, 2000)

Demzufolge bestehen von Kanton zu Kanton unterschiedliche Voraussetzungen, auch was die Höhe der Besoldung anbelangt. In der Sekundarstufe I lassen sich grundsätzlich zwei Typen von Lehrkräften unterscheiden: Lehrkräfte in den Abteilungen mit Grundansprüchen sind im allgemeinen Primarlehrkräfte mit einer spezifischen Zusatzausbildung; sie bleiben normalerweise „Generalisten“. In den Abteilungen mit erweiterten Anforderungen hingegen unterrichten an der Universität ausgebildete Spezialistinnen und Spezialisten in einem bestimmten Fach, meistens jedoch in zwei oder drei Fächern.

Die Studiengänge an der Universität sind nicht einheitlich aufgebaut; die Ausbildung der Sekundarlehrkräfte erfolgt entweder nur an der Universität oder besteht aus einer fachbezogenen Ausbildung an der Universität und einer berufsbezogenen Ausbildung an einem pädagogischen Institut. (Dauer zwischen sechs und zwölf Semestern). Die Anforderungen, die in diesen Studiengängen und in den entsprechenden Prüfungen gestellt werden, sind sowohl bezüglich ihres Inhalts als auch vom methodischen und didaktischen Standpunkt aus uneinheitlich.

Zur Zeit werden diese Fragen sehr eingehend untersucht und in Kürze sollen Anstrengungen für eine Koordination unternommen werden, insbesondere im Hinblick auf die gegenseitige Anerkennung der Lehrdiplome innerhalb der Schweiz.

9.17 Statistik Sekundarstufe I

Tabelle 33: Statistik Schweiz Sekundarstufe

Stand: 1994/95	Erweiterte Ansprüche	Ohne Selektion	Grundansprüche	Total
Schüler	166500	29400	88600	284500
Anteil Fremdsprachige ¹ (%)	11	17	28	17
Schüler im 7. Schuljahr ² (%)	60	9	31	100
Lehrkräfte ³ (in Vollzeitbeschäftigung)				20400
Betreuungsverhältnis ³ (Schüler pro Lehrkraft)				12,8
Einrichtungen (Arbeitsstätten ⁴)				6400

¹ Der Anteil Fremdsprachiger ist wegen seines Ausmaßes und aus pädagogisch-didaktischen Gründen eine wichtige Variable.

² Wir vergleichen die prozentualen Anteile der drei Schultypen nur für das 7. Schuljahr. So wirkt sich die kantonal unterschiedliche Dauer der Sekundarstufe I nicht aus.

³ Schuljahr 1993/94; nur öffentliche Schulen

⁴ Betriebszählung 1991. Wegen der Bildungsstruktur und den zur Verfügung stehenden Informationen kann die Anzahl der Einrichtungen nur für die gesamte obligatorische Schulzeit (Primarstufe und Sekundarstufe 1) angegeben werden.
Als Arbeitsstätte gilt ein Gebäude bzw. ein Teil davon, in dem während mindestens 20 Stunden pro Woche unterrichtet wird.

Quelle: Bundesamt für Statistik, Bildungsstatistik
Quelle: (Euridice, 2000)

9.18 Lehrplanweiterentwicklung

Ein Versuch, die Schweizer curriculare Diskussion und die Fortentwicklung der kantonalen Lehrpläne sowie des Schulsystems als Ganzes auch nur im Ansatz zu skizzieren, übersteige deutlich den Rahmen einer Kurzskizze.

9.19 Beispiele aktueller Forschungsfragen

Einen Einblick in aktuelle Projekte (Stand 2000) zur Bildungsforschung erhält man beispielsweise aus der periodisch erscheinenden Veröffentlichung „Permanente Erhebung über Bildungsforschung / Enquête permanente sur la recherche éducationnelle“ der Schweizerischen Koordinationsstelle für Bildungsforschung. (Wolter, Bauhofer, fortlaufend)⁴⁷.

⁴⁷ Der Papierversand der Blätter „Information Bildungsforschung der SKBF/CSRE wird mit Ende des Jahres 2000 eingestellt. „Unter <http://www.skbf-csre.ch> finden sich schon heute umfangreiche Informationen zur Bildungsforschung in der Schweiz Über die Seite „Was gibt's Neues? finden man Zugang zum jeweils neuesten Versand.“. (a.a.O., Mai 2000)

Beispielhaft und ohne eine Gewichtung seien erwähnt:

zu TIMSS:

Schule, Leistung und Persönlichkeit: TIMSS+ / École, compétences et personnalité: TIMSS+ des Amtes für Bildungsforschung der Erziehungsdirektion Bern (ABF) in Zusammenarbeit mit dem Pädagogischen Institut der Universität Zürich, der Abteilung Höheres Lehramt (AHL) der Universität Bern dem IRDP, Neuchâtel, dem Centre de recherches psychopédagogiques (CRPP), Genf und dem Ufficio studi e ricerche (USR) Bellinzona, 1993-1997

Bearbeiter Erich Ramseier, Urs Moser, Kurt Reusser, Peter Labudde, Alex Buff, Maja Huber, Carmen Keller, François Jaquet, Francesco Vanetta, Francesca Pese, Daniel Bain, Fiorella Gabriel et al. Dort werden die Schweizerischen TIMSS-Leistungen in Mathematik und Naturwissenschaften noch einmal nachhaltig hinterfragt.

zur Übersicht über die Schweizerischen Bildungsforschungsaktivitäten:

Les recherches en éducation en Suisse romande et au Tessin – état de situation 1998/99, analyse et priorités / Bildungsforschung in der Westschweiz und im Tessin: Bestandsaufnahme 1998/99, Analyse und Prioritäten des Institut de recherche et de documentation pédagogique (IRDP), Neuchâtel, 1998-1999, Bearbeiter Jaques Weiss, wo „der Schweizer Beitrag an den Universitäten der französischsprachigen Schweiz ... zur Entwicklung des Forschungsgeschehens im Bildungsbereich“ analysiert wird (a.a.O. Informationsblatt 99:061),

zur metadidaktischen Forschung:

Die Auswirkungen eines nationalen Forschungsprogramms auf die Forscherinnen und Forscher / Les effets d'un programme national de recherche sur les chercheurs impliqués des Büros für Bildungsforschung und -evaluation, Bern, 1998-1999, Bearbeiter Thomas Meyer, wo als einem von drei Schwerpunkten der Frage nachgegangen wird, ob es dem nationalen Forschungsprogramm „NFP 33 gelingt, nachhaltige Spuren in der Schweizerischen Bildungsforschungslandschaft zu hinterlassen“ (a.a.O. Informationsblatt 99:063).

Als Beispiele für konkret unterrichtliche didaktische Forschung sowie Trends seien erwähnt

Evaluation zum Wochenplanunterricht auf der Sekundarstufe 1 / Évaluation du „Wochenplanunterricht“ au niveau secondaire I des Departments Erziehungswissenschaften der Universität Freiburg, Abteilung Ausbildung für Lehrkräfte der Sekundarstufe I, 1996-2000, Bearbeiter Alois Higgli, Bernd Kersten et. al., wo

der Frage der Interdependenz von Wochenplanunterricht und didaktischem Verhalten der Lehrkräfte nachgegangen wird,

sowie der Trendbericht der Schweizerischen Koordinationsstelle für Bildungsforschung (SKBF Trendbericht Nr. 3, 1999), welcher folgende aktuelle Entwicklung aufzeichnet:

„Fördern und Auslesen gehören zu den Hauptaufgaben der Beurteilung in der Schule – mit divergierender Zielsetzung. Die Entflechtung der Aufgaben der Beurteilung verlangte 1986 eine These aus dem SIPRI-Projekt der EDK . Selbst wenn diese These schon alt ist, hat sie nichts an Aktualität verloren: Formen und Verfahren der Beurteilung sind nach dem Zweck auszurichten, und verschiedene Zwecke können hier nicht mit den gleichen Mitteln, z.B. Noten, eingelöst werden.

Unterdessen sind Beurteilungskonzepte, die das Fördern ins Zentrum stellen, so weit verfeinert und verbessert worden, dass sie auch in Reglementen Eingang gefunden haben und in die Praxis umgesetzt werden. Darzustellen, wo die Weiterentwicklung der Beurteilung heute steht und wie sie verlaufen ist, sind zwei Hauptanliegen dieses Berichtes. Sichtbare Zeichen dieser Entwicklung sind Lernberichte, Beurteilungsgespräche und die Selbstbeurteilung, die vorerst in der Primarschule große Verbreitung gefunden haben.“ (SKBF, 1999b).

Abschließend als Beispiel einer Untersuchung zum gesellschaftspolitischen Standort von Schule:

„Regard sur l'école suisse“, Dominique Gros, Programme national de recherche 33, Berne, 1999, mit folgender Untersuchung:

„Quelle est l'opinion des Suisses sur l'école? La population suisse en a-t-elle un bon ou mauvais souvenir? Quels sont les accents à mettre? En quoi consistent les forces et les faiblesses de l'école? Ces questions et un tas d'autres ont été posées à 1400 habitants, entre 15 et 65 ans, des trois régions linguistiques suisses. Ce livre présente les résultats de l'enquête.“ (SKBF, 1999c)

Einen Überblick über die Schweizer Ergebnisse in internationalen Vergleichsstudien findet man in der von der Schweizer Koordinationskonferenz Bildungsforschung (CORECHED) in Auftrag gegebenen Studie „Ausgewählte Schweizerische Ergebnisse in internationalen Leistungsvergleichsstudien“.

9.20 Internet-Links zu Bildungsorganisationen in der Schweiz

Tabelle 34: Bildungsorganisationen in der Schweiz

Zeit-Fragen	ist eine regelmässig erscheinende Zeitung für den deutschsprachigen Raum, die sich vorwiegend mit gesellschafts- und bildungspolitischen Themen in Deutschland, Österreich und der Schweiz befasst. Sie bietet keine tagespolitischen, sondern ausschliesslich grundsatzpolitische Artikel zu diesen Themen an und bezieht Stellung für eine freiheitliche, bürgerliche Wertesordnung.
9.20.1.1.1 Bildungsforschungsinstitute	
Agora-Server der Universität Genf	Bildungsforscher: auf diesem Server finden Sie eine virtuelle Bibliothek, Beschreibungen von Forschungsprojekte, und Kontaktadressen. Lehrer: ein Diskussionsforum steht Ihnen zur Verfügung um Erfahrungen auszutauschen. Auch die Mediothek und Softwarebeschreibungen können Sie bei der Gestaltung des Unterrichts aktiv unterstützen.
Die Schweizerische Konferenz der kantonalen Erziehungsdirektoren	ist der Zusammenschluss der 26 kantonalen Regierungsmitglieder, die für Erziehung, Bildung, Kultur und Sport zuständig sind. Die EDK verfügt über ein Sekretariat mit Sitz in Bern sowie einzelne Institutionen für besondere Dienstleistungen. Die EDK ist wichtiger bildungspolitischer Partner des Bundes.
Das nationale Forschungsprogramm 33 (NFP33)	behandelt die Thematik "Die Wirksamkeit unserer Bildungssysteme angesichts der demographischen und technologischen Entwicklung, und angesichts der Probleme der mehrsprachigen Schweiz". Gegenstand der Forschungsarbeit sind Bildungsinstitutionen und -prozesse und das in ihnen und durch sie Bewirkte. An Beispielen soll gezeigt werden, wie Forschung dazu beitragen kann, die Flexibilität und das Innovationspotential des Bildungswesens zu erhöhen, die Fähigkeit zur Evaluation und Selbstevaluation zu fördern und das erzieherische Handeln wirksamer zu gestalten.
Die Schweizerische Koordinationsstelle für Bildungsforschung SKBF	untersucht und erklärt die Wirklichkeit des Bildungsgetriebens, wie etwa Lernprozesse, Unterrichtsmethoden, Schülerbeurteilung und Schulstrukturen. Die SKBF übernimmt die wissenschaftliche Begleitung und Auswertung von Versuchen und Reformen im Bildungswesen. Damit trägt sie zur Klärung der Frage bei, ob angestrebte Ziele mit den eingesetzten Mitteln tatsächlich erreicht werden (evalutive Funktion). Zu den Aufgaben der SKBF gehören auch Planung, Entwicklung und Beratung in Erziehungs- und Schulfragen (Planungs- und Entwicklungsfunktion). Die SKBF koordiniert rund 125 meist kleinere Institutionen im Bereich der Bildungsforschung und Schulentwicklung

Die Schweizerische Fachstelle für Informationstechnologien im Bildungswesen SFIB	verfolgt das Ziel, Beiträge zu sinnvollem Computereinsatz im Unterricht zu leisten. Sie erbringt sie im Auftrag des BIGA und der EDK für allgemeinbildende und berufsbildende Schulen sowie für Sonderschulen in der ganzen Schweiz. Die Fachstelle verbreitet Informationen zum Thema Computereinsatz in der Schule. Sie arbeitet als Informationsdrehscheibe und macht ihre Arbeiten den Interessierten zugänglich.
Centre de compétence en matière d'informatique pédagogique au bénéfice de l'enseignement et des services du département de l'instruction publique genevois (DIP),	il offre un lieu d'échanges de réflexion et de formation réunissant des développeurs, chercheurs enseignants et utilisateurs dans le cadre de l'optimisation des nouvelles technologies de l'information et de la communication (NTIC).
Institut romand de recherches et de documentation pédagogiques. L'IRDP	est une institution de l'ensemble des cantons romands et du Tessin et constitue un lieu de débat éducatif en Suisse latine avec un objectif: une école toujours meilleure, des classes enfantines à l'Université. Il possède deux champs d'action principaux. Installé à Neuchâtel depuis sa création, en 1970, l'IRDP relève de la Conférence des chefs des Départements de l'instruction publique de la Suisse romande et du Tessin. Il est reconnu par la Loi fédérale sur la Recherche, qui lui attribue des subventions importantes.
L'Istituto Svizzero di Pedagogia per la Formazione Professionale Sezione di lingua italiana	è una divisione dell'Ufficio Federale dell'Industria, delle Arti e Mestieri e del Lavoro (UFIAML). Esso è presente nelle tre regioni linguistiche, con le sedi di Zollikofen, Losanna e Lugano
Zeitschrift "Panorama"	der Schweizerischen Gesellschaft für angewandte Berufsbildungsforschung SGAB und Schweizerischer Verband für Berufsberatung SVB, in Zusammenarbeit mit dem BIGA, Abteilung Arbeitsmarkt; erscheint zweimonatlich; Auflage/tirage: 3000 Ex. Abonnement: SFr/frs 60.- p.a. Einzelheft/numéro isolé SFr/frs 15.-
Der Verein Jugend und Wirtschaft	will gesamtschweizerisch als aktiver, eigenständiger und neutraler Vermittler zwischen Schule und Wirtschaft – die zentrale Bedeutung und die Komplexität des wirtschaftlichen Geschehens bewusst machen - Verständnis dafür wecken, dass wir Teil der Wirtschaft sind, und dass die Wirtschaft Teil unseres ganzen Lebens ist.
The International Baccalaureate Organisation (IBO)	a nonprofit educational foundation based in Switzerland, offers: the Diploma Programme for students in the final two years of secondary school the Middle Years Programme for students in the 11-16 age range. Aims: As described in our mission statement, the organisation's academic programmes aim to awaken the intelligence of young people and teach them to relate the content of the classroom to the realities of the world outside.
Abteilung Berufspädagogik	Die Mitarbeiterinnen und Mitarbeiter der Abteilung Berufspädagogik im Amt für Berufsbildung beschäftigen sich mit der Aus-, Fort- und Weiterbildung von Berufs-

	<p>schullehrern und -lehrerinnen, mit ihrer didaktischen Unterstützung, mit Schulentwicklung im Bereich der beruflichen Aus- und Weiterbildung und mit anderen Aufgaben zur Förderung der Innovation in der Berufsbildung. Die Informationsangebote sollen Kontaktmöglichkeiten für alle an der Berufsbildung Interessierten sein. Die zentrale Aufgabe ist es, einen Beitrag an die Entwicklung der Berufsbildung zu leisten, insbesondere an diejenige der Berufsschulen und der beruflichen Weiterbildung</p>
Das Schweizerische Institut für Berufspädagogik (SIBP)	ist die eidgenössische Aus- und Fortbildungsstätte für gewerblich-industrielle Berufsschullehrerinnen und Berufsschullehrer. Das 1972 gegründete Institut ist eine Abteilung des Bundesamtes für Industrie, Gewerbe und Arbeit (BIGA). Es hat seinen Sitz in Zollikofen bei Bern und Lugano.
Der Server "berufsbildung.ch"	bietet eine ganze Reihe von Links zu Berufsausbildungsstätten und -zentren im In- und Ausland.
Das Heilpädagogische Seminar Zürich HPS	ist ein Ausbildungsinstitut, das durch die Kantone Zürich, St. Gallen, Solothurn, Thurgau, Aargau, Schaffhausen und Graubünden getragen wird. Das HPS bietet Aus-, Fort- und Weiterbildungen für heilpädagogische Fachleute im Vorschul-, Schul- und Erwachsenenbereich an. Es stellt dem heilpädagogischen Feld und seinen Nachbargebieten eine Vielzahl von Dienstleistungen zur Verfügung.
Die Zentralstelle für die Weiterbildung der Mittelschullehrerinnen und -Lehrer - WBZ	publiziert jährlich ein Fortbildungsprogramm mit rund 100 Kursprojekten. Diese Angebote betreffen alle Fächer der allgemein- und berufsbildenden höheren Mittelschulen sowie fachübergreifende Themen, didaktische, pädagogische und Schulentwicklungsfragen. Die WBZ stellt die Administration und die Rechnungsführung, die Bedürfnisanalyse und die Evaluation sowie die Kaderbildung im Zusammenhang mit den Fortbildungsprojekten sicher. Außerdem berät sie fortbildungsinteressierte Lehrpersonen, Fachgruppen und Schulen und unterhält eine kleine Bibliothek und Dokumentationsstelle.
DKZ	Server des Bundesamts für Industrie, Gewerbe und Arbeit Dokumentations- und Kommunikationszentrum zur Berufsmaturität. Eine Dienstleistung der Abteilung Berufsbildung des BIGA
EducETH	ist ein Prototyp eines Education Servers, der 1995 an der ETH Zürich entwickelt wurde. EducETH richtet sich auf den Bereich der Sekundarstufe II aus. Im Rahmen des Projektes "School goes Internet" (Departement Informatik, Prof. C.A. Zehnder) wird EducETH bis auf weiteres von der ETH Zürich gewartet und betreut. Jedes betreute Fach untersteht einem speziellen Fachmaster, der inhaltlich die Leitung und Verantwortung für ein Fach übernimmt.
Interface	Die Schweizer Zeitschrift für den Einsatz des Computers im Unterricht

	Le magazine suisse pour l'utilisation de l'ordinateur dans l'enseignement L'organo svizzero per l'uso del computer nell'insegnamento.
Braintrack	is a worldwide university-index on the Internet. It contains internet-addresses of universities, polytechnics, colleges and other higher educational institutions all over the world
Schule aktuell on-line	Die Zeitschrift "Schule aktuell" richtet sich an Berufs-, Fach- und Mittelschüler und deren Lehrkräfte und erscheint jährlich. Schule aktuell wird allen betreffenden Schulen kostenlos zur Verfügung gestellt. Zudem unterhält der Verlag eine etwas desorganisierten Server mit Bildungsinformationen aus allen möglichen Bereichen.
Schulnetz Schweiz	Das Schulnetz Schweiz will Schulen im deutschsprachigen Raum den Zugang und die sinnvolle Nutzung des Internets erleichtern. Dabei richten wir besonderes Augenmerk auf Elementarstufe und Sekundarstufe I. Auf unserem Server stellen wir Schulen gratis Platz für ihre Präsentation zur Verfügung. Interessierte nehmen mit dem Administrator des Schulnetzes Kontakt auf. Die Seiten können von der Schule selber auf dem Server aktualisiert werden.
SLZ	Server der Schweizer Lehrerinnen- und Lehrerzeitung (Organ des Dachverbandes LCH)

Quelle und links unter <http://www.schulforum.ch/8netz/links-ch/links-ch.html>

10 Anhang: Ergänzende Daten zur Studie

Hier sind für wissenschaftliche Detailfragen Grundelemente der Studie zusammengetragen wie etwa verwendete TIMSS-Aufgaben mit ihren mathematischen Ansprüchen, Aufbau der Questionnaires, Struktur der Antwortkodierungen etc. Ihre Lektüre ist zum Verständnis der Studie nur bei weitergehenden Fragen erforderlich; die Auflistung erfolgt eher aus wissenschaftlicher Korrektheit.

10.1 Verwendete TIMSS-Items

Es wurden die nachfolgend kurz umrissenen TIMSS-Aufgaben verwendet. Die originalen Aufgaben sind über das Internet in Englisch erhältlich.

Die in Spalte 2 verwendeten Abkürzungen für die mathematischen Teilbereiche wurden bereits in 2.4.1 erläutert. Zur leichteren Übersicht werden sie hier noch einmal wiederholt:

Die TIMSS-Aufgaben sind nach *sechs inhaltlichen Dimensionen* strukturiert. Es bedeutet:

- n Zahlen und Zahlenverständnis (‘fraction and number sense’)
- a Algebra (‘algebra’)
- m Messen und Maßeinheiten (‘measurement’)
- g Geometrie (‘geometry’)
- d Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung (‘data representation, analysis, and probability’)
- p Proportionalität (‘proportionality’).

Auch in Wiederholung: der difficulty index ist Maß dafür, in welchem Maße die Aufgaben über alle partizipierenden Länder korrekt gelöst werden konnten.

Tabelle 35: Gehalt der verwendeten TIMSS-Aufgaben

Beispiele mathematischer Mikrobereiche in TIMSS-Aufgaben und zur Lösung erforderliche(s) Wissen/Fähigkeiten/Fertigkeiten						
Aufgabe	Test-art	Grad	A rt	diffic. Index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
I 01	m	2	a	628	Bedeutung einer Variablen	Sachverhalte mathematisieren ‚rückwärts‘
I 03	m	1	m	603	Flaschen abfüllen: Division mit Einheitenumrechnung	Aufteilen durch Division mit Einheitenumrechnung
I 04	s	4	a	591	zwei Folgen weiterführen: ausprobieren oder kgV	Zahlenfolgen weiterschreiben oder kgV anwenden

Aufgabe	Test-art	Grad	Art	diffic. Index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
I 05	m	4	n	448	Umgang mit Dezimaldifferenzen	im Kopf Dezimalstellen rechnen oder korrekt schriftl. subtrahieren
I 06	s	1	n	427	Brüche ordnen/vergleichen : größeren angeben	z.B. Regel kennen, dass Vergrößerung nur des Zählers den Wert eines Bruches erhöht
I 08	m	4	g	597	Punktkoordinaten im Kopf testen (Proportionalitätsgefühl)	Geradengleichung und Koordinaten kennen und zeichnen oder sehr gutes Gefühl für Steigungsdreieck
I 09	m	3	d	490	Wahrscheinlichkeiten aus Bruchanteilen erkennen	Gefühl für Bruchgrößen und Wissen um Verknüpfung rel. Häufigkeiten/Wahrscheinlichkeiten
J 12	s	2	n	593	schriftliches Dividieren	Regeln Bruchdivision (kürzen nicht gefordert)
J 14	m	2	n	585	stellenrichtige schriftliche Division	Kommastellenregelung bei Dezimaldivision oder Gefühl für Div. durch Zahl kleiner als 1
J 16	m	1	g	548	Koordinaten ungefähr ablesen	Umgang mit Koordinaten und Ablesen von Skalen
J 18	m	2	a	594	Relation erraten: Verdoppeln minus 1	Regel- bzw. Mustererkennung bei Folgen
K 01	m	3	n	464	graphische Repräsentation von Bruchanteilen erkennen	Bruchanteildarstellung in Flächen, Bruchgefühl oder davon unabhängig Einteilungsgefühl ('etwas mehr als die Hälfte')
K 02	s	4	n	487	Dezimalzahlen addieren	schriftlich addieren mit Dezimalstellen
K 03	m	3	g	478	Objekt in Gedanken räumlich rotieren	starkes räumliches Anschauungsvermögen
K 04	m	2	a	606	Ungleichung interpretieren	Terme rückübersetzen und Bedeutung einer einfachen Ungleichung erfassen und weiterschließen von der Hälfte auf das Ganze oder Rechenregeln anwenden
K 06	m	3	n	604	Prozentaufschlag abschätzen	Prozentrechnung: aus prozentualer Erhöhung auf absolute Zahl rückrechnen
K 08	m	2	g	639	Dreiecke im Kopf rotieren und äquivalente Stücke ablesen	Wissen über kongruente Dreiecke und ebenes Rotieren einer Figur im Kopf und Innenwinkelsatz Dreieck und daraus Berechnung Restwinkel
K 09	m	2	n	563	Brüche im Kopf addieren	gleichnamig machen vor Addition von Brüchen, evtl. im Kopf

Aufgabe	Test-art	Grad	Art	diffic. Index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
						oder geschicktes Abschätzen durch Umwandeln in Dezimalzahlen oder intelligent guessing mit Zahlgrößengefühl
L 11	m	4	a	640	Bruchreihenanteile erkennen und addieren	komplexe Sprachinformation verstehen und (gut abschätzend oder Bruchreihe entwickelnd) addieren
L 13	m	1	a	326	Mustererkennung in Musterreihen	Muster in Symbolfolgen erkennen, evtl. durch Zahlinformation speichern und wiedererkennen
L 14	m	2	p	693	Proportionalitäten mit Brüchen weiterberechnen	Proportionalität erkennen bei umgekehrt vorgegebener Formulierung ($x \text{ prop } y!$) und Prop.faktor errechnen und in beiden Richtungen anwenden
L 15	m	4	g	623	Winkelsumme im Viereck	Innenwinkelsumme kennen und 4. Winkel berechnen
L 16	s	2	a	615	Lineare Gleichung in x lösen	einfache Termumformungen
M 04	m	3	n	615	den größten einer Reihe von Brüchen erkennen	Umwandlung einfacher Brüche in Dezimalzahlen, evtl. auswendig, dadurch Sortieren/Ordnen der Brüche
M 05	m	2	g	565	180-Grad-Drehung im Kopf nachvollziehen	math. Sprache verstehen und Drehen in der Ebene im Kopf
M 06	s	4	p	634	aus Verhältnisangaben auf absolute Zahlen schließen + Bruchrechnung	Verhältnisangaben in Bruchanteile umrechnen und Bruchanteile einer Menge
M 07	m	4	g	457	Winkel abschätzen	Bezeichnungsweise von Winkeln kennen und (Winkel abschätzen oder über Bruchanteile aus gegebenen Verhältnissen berechnen)
M 08	s	2	n	575	Dezimalzahlen schriftlich multiplizieren	Kommasetzungsregeln bei Multiplikation von Dezimalzahlen
N 13	s	2	a	576	Bruchterm für gegebenes x berechnen	Berechnung eines Bruch-Variablenterms für speziellen Wert von x, Kürzen nicht gefordert
N 15	m	1	m	492	Winkel abschätzen	30-Grad- Winkel abschätzen, dabei Abgrenzung von 45 Grad
N 16	m	4	n	580	Bruchanteile als Zahlenarätsel, rückwärts zu rechnen	komplexe sprachliche Information aufnehmen und (daraus 'rückwärts' rechnen oder (Gleichung aufstellen und lösen))

Aufgabe	Test-art	Grad	Art	diffic. Index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
N 17	m	4	n	611	Proportionalrechnen mit Bruchanteilen in komplexer Fragestellung	komplexe sprachliche Information aufnehmen, daraus proportional mit Bruchanteil/Dezimalzahl Verbrauch berechnen, negative Frageformulierung erkennen, Dezimalzahlen subtrahieren
N 18	m	4	d	541	Wahrscheinlichkeit des Ziehens	Wahrscheinlichkeitsdefinition als günstig/möglich kennen, Fälle abzählen und als Bruch schreiben
O 02	m	2	n	680	prozentuale Steigerung errechnen aus absoluten Zahlen	Differenz als Aufschlag auf Grundwert erkennen und vom Zahlenanteil auf Prozentanteil umrechnen, Vermeidung der Falle Aufschlag = Prozentzahl
O 03	m	1	g	581	stufig liegende Nebenwinkel an geschnittenen Parallelen	Stufenwinkel an Parallelen und Transfer über Nebenwinkel, Vermeidung der Wechselwinkellösung aus falsch verstandener Fragestellung
O 05	m	4	d	587	Anzahlen aus Wahrscheinlichkeiten	Verknüpfung Häufigkeit /Wahrscheinlichkeit 'rückwärts' durch geschicktes Erweitern der vorgegebenen Häufigkeit
O 09	cs	4	n	571	Division durch einen Bruch erkennen	Anzahl Rundläufe durch Division ermitteln, dabei Division durch einen Bruch, evtl. Ausweichen über Aufteilen von 1/4 auf 1
P 09	m	2	g	617	Seitenberechnungen aus Verhältnissen bei ähnlichen Dreiecken	Ähnlichkeit als Proportionalstreckung erkennen, Streckfaktor als unechten Bruch berechnen und anwenden, dabei Richtung der Streckung beachten
P 10	m	1	a	540	Einfachen Term verstehen/vereinfachen	Verständnis einer Termsumme: Bedeutung eines Termes
P 16	s	2	n	637	Dezimalzahl in zunde gekürzte Bruchzahl umwandeln	Dezimalbrüche als Brüche schreiben und Brüche kürzen
Q 01	m	3	a	595	sprachliche Angaben als Term mathematisieren	komplexe sprachliche Formulierung eines Termes als Term 'übersetzen'
Q 02	m	2	a	568	Bruchterme subtrahieren	Subtraktion gleichnamiger Variablen-Bruchterme
Q 03	m	3	m	636	Längenangaben in verschiedenen Einheiten vergleichen	Zeitangaben vergleichen, dabei Umrechnung verschiedener Einheiten auch durch Division

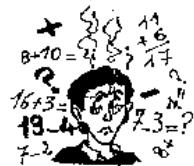
Aufgabe	Test-art	Grad	Art	diffic. Index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
Q 06	m	2	n	610	Wasserverbrauch im Jahr als Multiplikation zweier Zahlen°°!!	Anzahl Wochen pro Jahr wissen und durch Multiplikation Jahresverbrauch ermitteln
Q 07	m	2	a	519	Einfache Gleichung mit Termen lösen incl. Division	Werte in Variablenleichung einsetzen und eine Variable durch Division/Seitenvergleich ermitteln
Q 08	m	3	n	587	Anordnung von Dezimalzahlen und Brüchen gemischt	Frage richtig lesen und Bruch in Dezimalzahl umwandeln und Vergleich von Dezimalzahlen oder ohne Umwandlung durch geschicktes Ausschließen
Q 09	m	2	n	558	Bruchrechnung mit Addition und Multiplikation	Multiplikation und Addition ungleichnamiger Brüche incl. gleichnamig machen vor Addition, Vorrang Punktrechnung wird durch Klammern vorgegeben
R 07	m	4	n	583	Division von Dezimalzahlen erkennen (Papierdicke)	stellenrichtige Dezimaldivision oder Kopfrechnen mit Komma-verschiebung
R 08	m	4	d	565	gekrümmten Gräben extrapolieren und ablesen	Graph verstehen, extrapolieren und ablesen
R 09	m	1	a	603	falsche Termbeziehung (als Gleichung) erkennen	negative Frage nach grundlegenden Rechenregeln
R 10	m	3	g	551	Flächenzerlegung eines Trapezes in äquivalente Dreiecke	gedankliche Zerlegung eines Trapezes in kongruente Teilflächen, dabei interne Zerlegung eines Quadrates in zwei Dreiecke (oder gedanklich umgekehrt)
R 11	m	4	a	584	sprachliche Anzahlbeziehungen erkennen	komplex formulierte Anzahlbeziehungen verstehen und Teilprodukte summieren und Restfaktor durch Division ermitteln oder Gleichung aufstellen und lösen
R 13	s	4	n	642	Bruchanteile eines Wertes, negativ formuliert wg negativer Fo. und dadurch möglicher geschickter Rechenweise	Zahlanteil berechnen durch Multiplikation Zahl mal Bruch und Kürzen
R 14	s	4	p	617	Bruchanteile einer Menge berechnen	Bruchanteil aus Einzelanteilen, Anwendung auf errechnete Gesamtmenge als multiplikativer Faktor

Aufgabe	Test-art	Grad	Art	diffic. Index	Aufgabe beinhaltet	erforderliche Fähigkeiten / Fertigkeiten
T 01a	e	4	a	627	Äpfel in Kisten, bekannt Gesamtmenge und Mengendifferenz	komplexe Angabe zu Aufteilung einer Gesamtmenge verstehen, Aufteilen mit richtiger vorgegebener Differenz durch Raten oder Starten vom Mittelwert und Addition/Subtraktion
T 01b	e	4	a	631	Rechenweg angeben, auch 'guess and check' akzeptiert	anderer denkbarer Rechenweg wie 'Gesamtmenge minus Differenz):2 = gesuchte kleinere Menge usw.'
T 02a	e	4	p	699	sprachlich komplexe Aufgabe zu Proportionalität° bei Teilflächen	komplexe sprachliche Information verstehen, daraus Berechnung einer Boxenzahl, dann Berechnung der Gesamtstückzahl
T 02b	e	4	p	815	sprachlich komplexe Aufgabe zu Bruchanteilen bei Teilflächen	unter Verwendung von a) Berechnung von Teilsummen, daraus Bruchanteil berechnen
U 02a	e	4	m	621	Rechteckseiten proportional ändern	Streck- und Verkürzungsfaktoren auf Rechteckseiten anwenden und neues Rechteck zeichnen
U 02b	e	4	m	737	Flächenverhältnis bei Änderung der Seitenlängen mehrere Wege	alte und neue Rechtecksfläche berechnen, und korrekte Verhältnisbildung in vorgegebener Reihenfolge
V 02	e	4	d	675	komplexe Berechnung - Analyse von Mietangeboten	Fragestellung korrekt erkennen, komplexe sprachliche tabellarische Informationen auswerten, dabei unbekannte Einheiten akzeptieren, komplexe Rechnungen
V 03	m	2	p	603	Mischungsverhältnis aus absoluten Mengen	Anteile bei mehr als zwei Teilmengen als Verhältnis
V 04	s	3	m	610	Parallelogrammfläche mit Hilfslinien zum Rechteck °!	Parallelogrammfläche aus Rechtecksfläche herleiten, Abschneiden einer Ecke und Ansetzen zum Rechteck erkennen und auswerten oder Formel für Parallelogrammfläche anwenden

10.2 Testheft A: Schülerinformationsblatt

Das Testheft bestand aus einer Deckseite mit den Schülerinformationen zum Test. Danach folgten alle Aufgaben eins bis zwölf jeweils auf einer linken, die

Lösungsalternativen und Kommentarabfragen auf der rechten Seite von je zwei DIN A 4-Seiten.



Liebe Schülerin, lieber Schüler,

recht herzlichen Dank dafür, dass du bei diesem Mathematik-Wettbewerb mitmachst, bei dem 4 Klassen aus der Schweiz und 4 aus Deutschland gegeneinander antreten.

Wie Du vielleicht weißt, wurde vor einiger Zeit ein internationaler Mathematik-Vergleichstest mit dem abgekürzten Namen TIMSS mit Schülern aus über 40 Ländern durchgeführt.

Dabei gab es viele Aufgaben, wo die Lösungen nur aus einer Auswahl angekreuzt werden mussten, so dass man nicht unterscheiden konnte, ob sie gut gerechnet, geschickt geraten oder einfach nur so geraten waren.

In dem Aufgabenbuch, was jetzt vor dir liegt, sind einige der TIMSS-Aufgaben. Auch hier kannst du die Lösung ankreuzen, allerdings wollen wir diesmal zusätzlich wissen, **wie** du die Lösung erhalten hast. Es ist also **sehr wichtig** – genauso wichtig, wie die Aufgaben richtig zu lösen -, **dass du deinen Lösungsweg, deine Rechnung oder deine Überlegungen dazu hinschreibst !!!**

- Zeit hast du genug: für nur 12 Aufgaben hast du bis zu 90 Minuten.
- Platz hast du auch genug: eine Doppelseite für jede Aufgabe.

Die Aufgaben sind bunt gemischt. Die Reihenfolge der Bearbeitung bleibt dir überlassen, aber bitte möglichst keine vergessen!

Die Auswertung aller Wettbewerbshefte wird eine ganze Weile dauern. Danach kann dir deine Lehrerin/dein Lehrer sicher sagen, wer Sieger wurde – wenn es einen klaren Sieger gibt.

Wir – das Auswerteteam - wollen auch aus deinem Wettbewerbsheft lernen, wie man Mathe- matikunterricht vielleicht noch besser und interessanter machen kann. Genau deshalb kommt es so sehr darauf an zu sehen, **wie** du die Aufgaben gelöst hast. Zur Auswertung brauchen wir zusätzlich noch einige Angaben von dir:

Datum: _____

Mädchen

Junge

Name	Klasse	Geburtsdatum	oder Alter in Jahren und Monaten
------	--------	--------------	-------------------------------------

Wenn Du Deinen Namen nicht nennen möchtest, gib doch bitte einfach einen frei erfundenen Namen ein. Was wir aber wohl zur Auswertung und zum Vergleich brauchen, ist entweder dein Geburtsdatum oder dein Alter in Jahren und Monaten. Kreuze auch bitte das Geschlecht an.

Kommentierungsblatt zu jeder Aufgabe:

richtig ist: A B C D E

- Eine Aufgabe wie diese ist für mich neu.
- Ähnliche Aufgaben kenne ich aus dem Unterricht.
- Das ist schon länger her (mehr als sechs Monate).
- Das ist schon etwas her (drei bis sechs Monate).
- Das haben wir vor kurzem erst gemacht (weniger als drei Monate).
- Ich konnte sie damals gut.
- Ich kann sie noch gut.
- Ich habe sie nie gut gekonnt.

- Das Ergebnis kann man leicht erraten.
- Ich rate das Ergebnis, weil mir dazu kein Rechenweg einfällt.

- Zu der Aufgabe möchte ich noch sagen, daß

10.3 Lehrerbegleittext

⇒⇒⇒ Lehrerbegleitbuch ⇐⇐⇐

Liebe Kollegin, lieber Kollege,

recht herzlichen Dank für Ihre Bereitschaft, mit Ihren SchülerInnen die beiliegenden Aufgaben zu bearbeiten. Sie dienen als Forschungsgrundlage für die Frage, ob es landesspezifische Unterschiede Schweiz/Deutschland in den Techniken gibt, die die SchülerInnen zur Aufgabenlösung einsetzen.

Die originalen TIMSS-Aufgabenbearbeitungen sind leider nicht frei zugänglich und wären aufgrund des hohen multiple-choice-Anteils auch größtenteils wenig aussagekräftig für diese Fragestellung.

In diesem Test werden viele der freigegebenen TIMSS-Aufgaben erneut verwendet; allerdings sollen über die TIMSS-Auswertetechnik hinaus Lösungswege begutachtet, klassifiziert und bewertet werden.

Umso wichtiger ist es daher, dass Ihre SchülerInnen möglichst alle Gedanken, Ideen, Lösungsschritte, Zwischenrechnungen etc. hinschreiben !!! – **soweit die Aufgabe es zuläßt.**

Bitte wiederholen Sie dies mehrmals während der Bearbeitungszeit und schreiben Sie es bitte zusätzlich auf die Tafel.

Zum Testablauf:

Es gibt 6 Aufgabenbooklets A, B, C, D, E (und F als Sondertests). Jedes booklet hat andere Aufgaben; die Themenbereiche und Schwierigkeitsgrade sind in A bis E etwa gleichverteilt. Buch F ist anders konzipiert und strukturiert.

Bitte verteilen Sie die booklets gleichmäßig auf die SchülerInnen.

Lassen Sie genügend Zeit, die Schülerhinweise zu lesen (Sie finden Sie auf der nächsten Seite zu Ihrer Information noch einmal abgedruckt) und betonen Sie noch einmal die Wichtigkeit, Lösungswege hinzuschreiben, und auch, dass genügend Zeit (maximal 90 Minuten) zur Verfügung steht.

Bitte erwähnen Sie auch, dass ‚Abgucken‘ nicht nur nichts ‚bringt‘, sondern auch dem Testzweck entgegenläuft. Die Leistungen werden ja ohnehin anonym erbracht.

Erklären Sie bitte, dass mit Absicht für jede Aufgabe eine ganze Doppelseite Platz reserviert wurde, damit auch viele Überlegungen hinpassen.

Sollten Ihre SchülerInnen deutlich früher fertig sein, spricht nichts dagegen, dann einfach aufzuhören.
⇒⇒⇒⇒⇒

Zur Abschätzung der Curriculumvalidität sowie zur Beurteilung des Schwierigkeitsgrades der Aufgaben für **Ihre** SchülerInnen und aus **Ihrer** Sicht wäre es sehr wichtig, wenn Sie die Aufgaben kurz durch Ankreuzen der jeweils darunter gedruckten Bewertungstabellen einordnen könnten. Es sind zwar 6 booklets mit je 12 Aufgaben, aber da jeweils mehrere davon von ähnlichem Typ sind, ist es vielleicht doch keine so riesige Arbeit. Erst durch Ihre Angaben lässt sich das Gewicht der Aufgabenbearbeitungen richtig einordnen.

Dieses Lehrerbegleitbuch reserviert für jede Aufgabe eine Seite mit einem vorgefertigtem Beurteilungsraster, welches lediglich ein Vorschlag ist. Sie können natürlich gerne Ihre weiteren Kommentare oder Meinungen hinzufügen.

Recht herzlichen Dank für Ihre Mitarbeit auch dabei.

Hier jetzt – wie oben angekündigt - kleingedruckt noch einmal der Text, den die SchülerInnen auf der ersten Seite lesen können.

Dieser Text, der die Lehrer darüber informiert, wie die Schüler in den Test eingeführt werden, ist der Text zu 10.2 (Schülerinformationsblatt) und wird deshalb nicht erneut abgedruckt.

10.4 Lösungsblatt und Kodierschlüssel zu Testheft A

Exemplarisch wird das vollständige Lösungsblatt für das Testheft A abgedruckt. Es folgt ebenfalls exemplarisch der verwendete Kodierschlüssel dazu.

Die Schlüssel für die anderen Testhefte sind analog aufgebaut.

Tabelle 36: Lösungsblatt zu Textheft A

Lösungsblatt zu Testheft A			
1.	K01	D	
2.	P10	B	4m
3.	K03	A	
4.	K09		$\frac{115}{24} = 4 \frac{19}{24}$
5.	V03	D	5 9
6.	J18	B	3
7.	V02		Gebäude A: 9600/Jahr oder 800/Monat ist günstiger B: 9900/Jahr oder 825/Monat
8.	J14*	C	15,35
9.	Q03	B	1500 Minuten ist die längste Zeit (25 Stunden)
10.	M07	D	80 Grad
11.	O02	C	25 %
12.	L16		x = 7

Tabelle 37: Ergebniskodierung zu Testheft A

Ergebniskodierung zu Testheft A	
20	Weg/Rechnung korrekt und korrekt angekreuzt/angegeben
19	Weg/Rechnung teilweise falsch, aber Ergebnis korrekt und korrekt angekreuzt
18	kaum/kein Weg/Rechengang vorhanden, aber korrekt angekreuzt/angegeben
17	Weg/Rechnung nicht nachvollziehbar, aber Ergebnis korrekt und korrekt angekreuzt
16	Rechenfehler heben sich weg, daher korrektes Ergebnis und korrekt angekreuzt
15	Weg/Rechnung falsch, aber korrekt angekreuzt/angegeben
14	explizit als geraten gekennzeichnet und korrekt angekreuzt auch: als unklar bezeichnet und korrekt geraten
13	keiner der Fälle 14 - 20, aber Ergebnis korrekt angekreuzt/angegeben, z.B. auch / abgemessen statt gerechnet / abgeschätzt / ausprobiert / ... /
12	Weg/Rechnung und Ergebnis korrekt, aber nicht angekreuzt
11	Weg/Rechnung und Ergebnis korrekt, aber falsch angekreuzt/angegeben
10	Weg/Rechnung korrekt, aber ein leichter Rechenfehler, daher falsch angekreuzt/angegeben
9	Weg/Rechnung korrekt, aber mehrere leichte Rechenfehler, daher falsch angekreuzt
8	Weg/Rechnung vom Ansatz her korrekt, aber ein/mehrere grobe(r) Rechenfehler
7	Weg/Rechnung vom Ansatz her korrekt, aber ein/mehrere grobe(r) Denkfehler, auch: / falsche räumliche Drehung / als „Wissenslücke“ gekennzeichnet /
6	Weg/Rechnung vom Ansatz her falsch und daher falsch angegeben/angekreuzt
5	Rechengang sinnleer, Ergebnis falsch angegeben/angekreuzt
4	kein/e Weg/Rechnung vorhanden und Ergebnis falsch angegeben/angekreuzt

3	explizit als geraten gekennzeichnet und falsch angekreuzt
2	als „nicht verstanden“ oder „kann ich nicht“ gekennzeichnet und nicht bearbeitet
1	keiner der Fälle 2 - 12 , aber Ergebnis falsch angekreuzt/angegeben, z.B. auch / nicht angekreuzt mehrfach angekreuzt / nach Ansatz Bearbeitung abgebrochen / ... /
0	Aufgabe nicht bearbeitet

10.5 Transskripte der Lehrerquestionnaires

Tabelle 38: Transskripte der Lehrerquestionnaires

Transskript der expliziten Textanmerkungen in den Lehrerquestionnaires					
1: Kategorie		2: Testheft		3: TIMSS-Nr.	4: Klasse
5: Kodierung im Lehrerfragebogen		6: Lehrerkommentar			
1	2	3	4	5	6
a	A	P10	7a G1	4109	Schreibweise 4m unbekannt, stattdessen 4 * m
a	A	L16	7a G1	2009	Stoff der Klasse 8, Schreibweise 10x, 15x unbekannt
a	B	N13	7a G1	4009	Schreibweise 7x, 5x unbekannt
g	B	R10	7a G1	3009	Begriff „Trapez“ nicht bekannt
p	B	T2ab	7a G1	2009	Aufgabenstellung unklar formuliert
a	B	R09	7a G1	4009	Schreibweise ohne Multiplikationszeichen nicht bekannt, reelle Zahlen nicht bekannt
a	C	Q01	7a G1	4309	3n Schreibweise ohne Multiplikationszeichen ist nicht bekannt
a	C	K04	7a G1	2009	Äquivalenzbegriff noch nicht eingeführt
p	D	M06	7a G1	2009	Verhältnisbegriff noch nicht behandelt
a	D	Q07	7a G1	0209	„LW“ Schreibweise unbekannt, Auflösen von Formeln erst in Klasse 8
a	E	Q02	7a G1	3109	Schreibweise 2x unbekannt, Zusammenfassen gleichartiger Glieder ausführlich erst in Klasse 8
m	E	V04	7a G1	2009	Flächenberechnung von Vierecken (außer Quadrat, Rechteck) erst in Klasse 8
p	F	F05	7a G1	1009	Begriff „Schwerlinie“ nicht bekannt (Seitenhalbierende?)
p	A	V03	7 KA	2009	„Verhältnis“ stiftet ev. Verwirrung andere Fragestellungen für das gleiche Problem bekannt
n	A	O02	7 KA	3149	soeben mit %-Rechnung gestartet
A	B	R09	7 KA	3109	Begriff „reelle“ Zahlen noch unbekannt
g	B	P09	7 KA	2009	Ähnlichkeit noch unbekannt
n	B	Q08	7 KA	4219	kaum geübt in dieser Form
a	C	K04	7 KA	2009	„aequivalent“: Ausdruck unbekannt
n	D	I06	7 KA	0009	Scherzfrage ??
p	D	M06	7 KA	2009	„Verhältnis“
g	D	K08	7 KA	2009	„kongruent“ noch unbekannt
n	D	K06	7 KA	3249	soeben mit Prozentrechnung begonnen
g	E	M05	7 KA	2009	Abbildungen trotz Lehrplan noch nicht durchgenommen
g	E	I08	7 KA	2009	Koord.system unbekannt
a	F	F06	7 KA	2009	„rational“ → unbekannt
m	F	F09	7 KA	2009	s.o. (dort wurde das Wort „rational“ unterstrichen (Anm.

Transskript der expliziten Textanmerkungen in den Lehrerquestionnaires							
1: Kategorie		2: Testheft		3: TIMSS-Nr.		4: Klasse	
5: Kodierung im Lehrerfragebogen			6: Lehrerkommentar				
1	2	3	4	5	6		
					des Verf.)		
a	D	Q07	7KE	3329	Haben in Physik solche Formeln ausgewertet.		
n	F	F08	7KE	1009	ist noch nicht behandelt worden		
g	F	F03	7G1,2	2009	griechische Buchstaben noch nicht eingeführt		
n	F	F08	7G1,2	1009	Doppelzeichen unbekannt - -		
Die handschriftlichen Kommentare wurden im Druckbild so weit wie möglich nachgebildet.							

10.6 Transskripte der Schülerlösungswege bei Aufgabe M 07

Tabelle 39: Transskripte zu Aufgabe M07 GER Kl.7

Aufgabe M 07 booklet A Seite 20 Kommentare deutscher Schüler Klasse 7			
Schüler		Kommentar	Bekanntheitsgrad
013 w	r	Ich messe die Gradzahl mit dem Geodreieck ab. Ich habe 80 Grad raus.	vor kurzen gut
014 w	r	Mir ist diese Aufgabe nicht klar. Ich habe auch sehr große Probleme mit Geometrie.	ähnlich bekannt geraten
015 w	r	kommt ungefähr hin	vor kurzem
016 w	f	20 Grad $5 \times E4x = 9x$ $180^\circ : 9x = 20$	
39 m	r	80, weil es sehr nah an einem rechten Winkel liegt und der hat 90 grad	ähnlich bekannt damals gut, leicht zu erraten
40 w	-	nicht bearbeitet	für mich neu
41 m	-	nichts angekreuzt, Text: $360 : 9 = 40$, $4E40 = 160$	für mich neu
42 w	r	4x=D ist meiner Meinung nach richtig. <i>Interpretation (Verf.): 4x = Name Winkel, D ist Punkt der Auswahl, also 80 Grad</i>	für mich neu, Winkel geschätzt, daher nicht sicher
093 w	f	doppelt angekreuzt 20° und 100°	keine Angabe
094 w	r	Man muß wissen wo der Winkel 90° ist	noch sagen: leicht zu lösen
095 m*	-	keine Ahnung	ähnlich bekannt
096 m	f	20° , weil das etwas weniger ist als die Hälfte und damit weniger als 90° . Und $20^\circ E 4$ ergibt 80°	ähnlich bekannt
097 m	f	Es kommt 20 heraus, da rechts das Ergebnis 80° ist und links das Ergebnis 100° ist. Die Angaben 80° und 100° habe ich durch Anlegen des Geo-Dreiecks rausgekriegt und dann $80:4$ und $100:5$ gerechnet.	etwas her damals gut leicht erratbar
120 m	f	20 Grad ich rechne die Winkel aus, dann rechne ich die 80 Grad, die dort herauskommen, mal	-
121 w	r	Das Maß des Winkels beträgt 80 Grad. Ich habe es mit dem Geodreieck gemessen.	ähnlich bekannt, damals gut, leicht

Aufgabe M 07 booklet A Seite 20 Kommentare deutscher Schüler Klasse 7			
Schü-ler		Kommentar	Bekanntheitsgrad
			erraten
122 m	f	100 Grad Ich habe einfach den Winkel im Punkt C ab- gemessen.	ähnlich bekannt etwas her
123 m	f	100 Grad	ähnlich bekannt, länger her, damals gut, geraten
145 w	f	100 Grad	neu, ich die Grad- zahl gemessen habe
146 m	r		damals gut
147 w	r		ähnlich bekannt, noch einigermaßen, erraten
148 m	r	Ich habe das Geodreieck an den Punkt C angelegt und den Winkel ausgemessen	ähnlich bekannt, vor kurzem, noch gut
149 w **	f	40 Grad Gekritzeln und ich rate	nie gut, Aufgabe dumm

Tabelle 40: Transskripte zu Aufgabe M07 CH Kl.7

Aufgabe M 07 booklet A Seite 20 Kommentare Schweizer Schüler Klasse 7			
Schü-ler		Kommentar	Bekanntheitsgrad
172 w	f	100 Grad eine Senkrechte gezeichnet, korrekt Winkel AC Senkrechte=90 Grad, Winkel Senkrechte CD =10Grad, mit Bogen markiert, =100 Grad daran geschrieben (also für Winkel ACD korrektes Ergebnis angegeben)	damals gut
173 m	f	50 Grad ohne Kommentar	leicht erraten
174 w	r	90 Grad-Hilfslinie korrekt eingezeichnet	leicht erraten
194 m	r	Es ist ganz sicher weniger als ein 90-Grad-Winkel. Der Winkel AB ist 180. $5x+4x=9x$ $x=20$ Grad $4E20$ grad = 80 grad	ähnlich bekannt etwas länger her noch gut
195 w	r		neu, geraten
196 m	r	ich habe das Geodreieck benutzt	ähnlich bekannt länger her damals gut
208 w	r		neu, geraten
209 m	-	nicht bearbeitet	neu
210 w	r	man muß sehr gutes Vorstellungsvermögen haben	neu
243 w	r	Ich denke, dass es 80 Grad sind, weil es leicht weniger ist als 90 Grad	leicht erraten
244 m	f	wir noch nicht gelernt haben, wie man Winkel berechnet	neu, geraten
245 m	f	100 Grad	leicht erraten

Tabelle 41: Transskripte zu Aufgabe M07 GER Kl.8

Aufgabe M 07 booklet A Seite 20 Kommentare deutscher Schüler Klasse 8			
Schü-ler		Kommentar	Bekanntheitsgrad
065 m	r	BCD ist kleiner als 100 Grad aber nicht 90 Grad. BCD + ACD sind 180 Grad. Also BCD gleich 80 Grad	ähnlich bekannt
066 m	f	180:9=20 5 E 20 = 100	ähnlich bekannt
067 m	r	3 Skizzen: 1x gestreckte Gerade 180 Grad, 1x zwei Winkel 90 Grad, 1x 100/80 Grad	vor kurzem erst
068 w	r	BCD ist etwas weniger als 90°. Also 80°	ähnlich bekannt, etwas her, damals gut, leicht erratbar
069 w	r	Zwischen 90° und 50° bleibt nur 80° über.	ähnlich bekannt, vor kurzem, noch gut, leicht erratbar

10.7 Beispielhafte Transskripte von Schülerkommentierungen

Tabelle 42: Allgemeine Transskripte der Schülerkommentare

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 A GER 1					
Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transskribiert.					
1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
A	K01	n	3310	7A G1 GER 013 w	man Konzentration braucht um sich das vorzustellen, und in einer lauten Klasse ist das nicht einfach.
A	P10	a	4109		
A	K03	g	3200	7A G1 GER 013 w	man sich auch konzentrieren muss und sich das mit geschlossenen Augen vorstellen muss.
A	J18	a	3310	7A G1 GER 014 w	ich mir die Lösung nur denken kann, da fast alle Aufgaben in diesem Heft entweder schon lange her sind oder noch nicht besprochen worden ist.
			7A G1 GER 015 w	ich sie nur geraten habe ich wusste nicht mehr wie das geht	
A	V02	d	3000	7A G1 GER 015 w	... ist für mich neu, aber total
A	I14	n	4000	7A G1 GER 013 w	ich leider vergessen habe wie genau das geht.
A	L16	a	2009	7A G1 GER 013 w	ich mich nicht mehr gut erinnern kann

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 A GER 1

Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.

1: Aufgabenbuch 2: TIMSS-Aufgabennummer 3: „content category“ = Bereich
4: Lehrerkodierung 5: Klasse, Schule, Land, Schülerkennziffer, Geschlecht

1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
				7A G1 GER 014 w	Kann ich nicht mehr. Ich bin mir auch nicht sicher ob wir diese Aufgaben je gemacht haben! Insgesamt fand ich diesen Test ziemlich schwer da ich sehr viel nicht verstanden habe! Ich schätze mal das sie mich jetzt für total Doof halten!
B	N13	a	4009	7A G1 GER 017 w	wir eine solche Aufgabe noch nie hatten. Bzw. war die Aufgabenstellung noch nie so formuliert worden. Sie ist mir kein Begriff.
B	R10	g	3009		
B	N16	n	4200	7A G1 GER 017 w	Ich rate das Ergebnis, weil mir dazu kein Rechenweg einfällt geändert in: „, weil mir dazu nicht der richtige Rechenweg einfällt.
B	T2ab	p	2009	7A G1 GER 017 w	ich sie wirklich (!) nicht verstehre.
				7A G1 GER 018 w	mir die Aufgabe unbekannt ist.
				7A G1 GER 019 m	sie sehr schwer ist. Sie ist unverständlich.
				7A G1 GER 020 w	wir das noch nicht im Unterricht durchgenommen haben.
B	R09	a	4009	7A G1 GER 017 w	wir so etwas noch nie zuvor bearbeitet haben. Außerdem ist mir „reelle Zahl“ kein Begriff.
B	I09	d	2000		
B	R07	n	4000	7A G1 GER 017 w	... leicht erraten. ergänzt zu: „wenn man eine Division vornimmt. obwohl diese (siehe meine Rechnung) falsch ist.“
B	I03	m	4010		
B	P09	g	1009	7A G1 GER 017 w	ich um das zu sagen in der Zeichnung nachmessen musste, was uns eigentlich untersagt wurde. Außerdem ergab die Messung nur einen Näherungswert von 5,2 cm.
				7A G1 GER 018 w	ich mit so einer ähnlichen Aufgabe immer große Probleme hatte!
B	Q08	n	4000		
B	L11	a	2000	7A G1 GER 017 w	Text geändert zu „nicht der richtige Rechenweg.“ +: Gesamtmeinung zum Test: Vieles hatten wir noch nicht oder manche Formulierungen tauchten bei uns im Unterricht noch nie zuvor auf. Wie zum Beispiel der Begriff „reelle Zahl“.
C	Q01	a	4309	7A G1 GER 023	ich keine Ahnung habe was die Zahlen be-

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 A GER 1

Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.

1: Aufgabenbuch 2: TIMSS-Aufgabenummer 3: „content category“ = Bereich
4: Lehrerkodierung 5: Klasse, Schule, Land, Schülerkennziffer, Geschlecht

1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
				m	deuten, da sie Stoff einer höheren Klasse sind.
				7A G1 GER 024 w	wir diese Aufgabe noch nicht im Unterricht durchgenommen haben.
C	U2b	m	3009	7A G1 GER 022 m	Verhältnis hatten wir noch nicht aber ich denke es ist 1 cm.
C	K04	a	2009	7A G1 GER 023 m	auch diese Aufgabe, Stoff einer höheren Klasse ist.
				7A G1 GER 024 w	ich den Ausdruck „äquivalent“ nicht kenne.
D	M06	p	2009	7A G1 GER 025 w	mir kein Rechenweg eingefallen ist. Ich glaube auch dass wir <u>das</u> noch nich hatten.
				7A G1 GER 028 w	ich die Aufgabe noch nicht hatte.
D	Q07	a	3209	7A G1 GER 025 w	mir kein Rechenweg eingefallen ist.
				7A G1 GER 028 w	mir kein Rechenweg eingefallen ist
D	N18	d	1000	7A G1 GER 025 w	mir kein Rechenweg eingefallen ist.
D	M08	n	4000	7A G1 GER 025 w	ich sie besonders schwer finde und das Ergebnis daher falsch ist.
E	I05	n	4000	7A G1 GER 029 w	sie sehr leicht zu lösen war und ich damit keine Probleme hatte.
				7A G1 GER 033 w	diese Aufgabe sehr leicht für mich zu lösen war und ich nicht viel bzw. gar nicht überlegen musste.
				7A G1 GER 030 m	sie ziemlich leicht ist.
E	Q02	a	3109		
E	M05	g	4300	7A G1 GER 031 w	ich sowas noch nicht gehabt habe es aber trotzdem lösen konnte.
				7A G1 GER 032 m	nicht klar ist, in welche Richtung man drehen soll.
E	J12	n	4000	7A G1 GER 029 w	sie auch sehr einfach war.
				7A G1 GER 033 w	sie ebenfalls leicht war. Ich glaube, dass ist der Stoff der 5 oder 6. Klasse.
E	L14	p	4020	7A G1 GER 033 w	wir solche noch nicht hatten und ich deshalb nicht in der Lage bin sie zu lösen.
E	R11	a	4000	7A G1 GER 029 w	sie relativ einfach war und man kaum überlegen musste.

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 A GER 1

Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.

1: Aufgabenbuch 2: TIMSS-Aufgabenummer 3: „content category“ = Bereich
4: Lehrerkodierung 5: Klasse, Schule, Land, Schülerkennziffer, Geschlecht

1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
				7A G1 GER 033 w	sie schon schwieriger war. Jemand der nicht sehr gut logisch denken kann (ich) ist sie schon schwieriger. Ich glaube deshalb, dass ich sie falsch gelöst habe
E	O05	d	1000	7A G1 GER 031 w	sie zu ungenau erklärt ist.
E	V04	m	2009	7A G1 GER 031 w	ich die eine Seite abgemessen habe aber nicht weiß ob es richtig ist.
				7A G1 GER 033 w	ich sie nicht kann.
E	I08	g	2000	7A G1 GER 033 w	sie ganz easy ist. Man muss nicht rechnen, nur probieren.
E	I01	a	2000	7A G1 GER 033 w	sie komisch formuliert ist.
F	F02	a	4000	7A G1 GER 038 w	ich bei b die Aufgabe nicht lösen kann weil mir dieses unbekannt ist oder ich es vergessen habe
F	F04	n	4000	7A G1 GER 034 w	Sie war sehr einfach
				7A G1 GER 038 w	sie sehr einfach war weil unser Lehrer so etwas sehr oft mit uns gemacht hat
F	F05	p	1009	7A G1 GER 034 w	wir noch nie über Schwerlinien gesprochen haben. Nur der Schwerpunkt ist mir bekannt.
				7A G1 GER 038 w	mir der Begriff „Schwerlinie“ unbekannt ist. Darum kann ich diese Aufgabe nicht lösen. Ich könnte mit aber vorstellen das es Seitenhalbierende sind.
				7A G1 GER 037 m	ich nicht genau weiß, ob Schwerlinien das gleiche sind wie Seitenhalbierende.
F	F06	a	4000		
F	F07	d	4200	7A G1 GER 034 w	Ich nicht mehr wusste, wie man den Punkt mit Zirkel und Lineal konstruierte. Ich habe es mit dem Geodreieck bzw. mit der Skala gemacht.
F	F08	n	4000	7A G1 GER 034 w	es sehr eintönig war nur – und nie + zu schreiben.
F	F09	m	4000	7A G1 GER 038 w	wir so etwas nie gemacht haben und ich nur den Grundaufbau kann.
F	F10	g	4000	7A G1 GER 034 w	sie bei mir in der 7. Klasse noch nicht vorkam. Ich nehme an, dass es an den unterschiedlichen Mathematikbüchern liegen kann.
				7A G1 GER 038 w	es schon sehr lange her ist und ich fast alles vergessen habe
F	F11	n	4000	7A G1 GER 034 w	ich bei der Aufgabe b) Probleme hatte zu kürzen

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 A GER 1

Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.

1: Aufgabenbuch 2: TIMSS-Aufgabennummer 3: „content category“ = Bereich
4: Lehrerkodierung 5: Klasse, Schule, Land, Schülerkennziffer, Geschlecht

1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
				7A G1 GER 036 w	ich sie zu schwer finde!
				7A G1 GER 038 w	dieses Thema sehr schwer ist und man es nach einiger Zeit nicht mehr so gut kann
F	F12	a	1000	7A G1 GER 036 w	sie mir nicht klar ist. Ich verstehe die Aufgabenstellung nicht.

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7C GER 2

Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.

1: booklet 2: TIMSS-Aufgabennummer 3: content category
4: Lehrerkodierung 5: Klasse, Schule, Land, Schülerkennziffer, Geschlecht

1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
A	K01	n	4310	7C G2 GER 145 w	von 18 Kästchen 10 angemalt sind also muß es im Diagramm mindestens über die Hälfte sein. Und da 9 die Hälfte ist darf es nur kurz über die Hälfte sein.
				7C G2 GER 146 m	sie sehr einfach ist.
A	P10	a	2000	7C G2 GER 145 w	ich geraten habe
				7C G2 GER 148 m	Ich vermute sie herauszukriegen.
				7C G2 GER 146 m	ich mir nicht sicher bin.
A	K03	g	2000	7C G2 GER 145 w	ich alle Körper im Kopf umgedreht habe.
				7C G2 GER 146 m	es sehr lange her ist, seit wir das gemacht haben.
A	K09	n	5410	7C G2 GER 145 w	ich alle Brüche auf einen Nenner brachte und dann normal addiert habe.
				7C G2 GER 146 m	sie einfach ist.
A	V03	p	4310	7C G2 GER 145 w	ich blau und gelb zusammen gerechnet habe und aus rot und blau gelb den Bruch gebildet habe.
				7C G2 GER 149 m	sie scheiße ist
				7C G2 GER 146 m	sie einfach ist.
A	J18	a	2000	7C G2 GER 145 w	ich die Zahl verdoppelt habe minus 1
A	V02	d	4310	7C G2 GER 145 w	zeds pro Jahr pro Quadratmeter ausgerechnet habe

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7C GER 2					
Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.					
1:	booklet	2:	TIMSS-Aufgabennummer	3:	content category
4:	Lehrerkodierung	5:	Klasse, Schule, Land, Schülernennziffer, Geschlecht		
1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
				7C G2 GER 146 m	es so lange her ist, daß ich es nicht mehr weiß.
A	I14	n	5410	7C G2 GER 145 w	ich das Komma verschoben habe und dann normal geteilt habe.
				7C G2 GER 146 m	ich es erraten habe.
A	Q03	m	4310	7C G2 GER 145 w	ich alle auf Sekunden umgerechnet habe.
				7C G2 GER 14 m	es lange her ist, dass wir so etwas gemacht haben
A	M07	g	3130	7C G2 GER 145 w	ich die Gradzahl gemessen habe.
				7C G2 GER 149 m	sie dumm sind
				7C G2 GER 146 m	ich es erraten habe. ich kann sie noch einigermaßen
A	O02	n	5310	7C G2 GER 145 w	ich ausgerechnet habe um wie viel es stieg
				7C G2 GER 146 m	ich mir nicht mehr sicher bin, wie der Rechenweg ist.
				7C G2 GER 145 w	es lange her ist, dass ich so etwas gerechnet habe
A	L16	a	2000	7C G2 GER 145 w	ich sie nicht gekonnt habe / unlösbar
				7C G2 GER 146 m	ich nur versucht habe es zu erraten, weil es ziemlich schwer war.
				7C G2 GER 149 m	ich überhaupt keine Lust habe diesen billigen Test zu machen deshalb mach ich mit allen Fehlern Schluss!

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 G2 CH					
Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.					
1:	booklet	2:	TIMSS-Aufgabennummer	3:	content category
4:	Lehrerkodierung	5:	Klasse, Schule, Land, Schülernennziffer, Geschlecht		
1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
A	P10	a		7 G2 CH 244 m	wir solche immer wieder repetiert haben.
A	K03	g		7 G2 CH 244 m	sie zwar neu ist für mich, aber sie ist ziemlich einfach.
A	V02	d		7 G2 CH 244 m	ich so etwas noch nie gesehen habe, und ich auch keine Ahnung habe, wie man das überhaupt beginnen soll.
A	I14	n		7 G2 CH 244 m	ich es fast verlernt habe. (Zeitergänzung „ca. 1 Jahr“)
A	M07	g		7 G2 CH 244 m	wir noch nicht gelernt haben, wie man Winkel

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 G2 CH					
Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.					
1:	booklet	2:	TIMSS-Aufgabennummer	3:	content category
4:	Lehrerkodierung	5:	Klasse, Schule, Land, Schülerkennziffer, Geschlecht		
1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
					berechnet.
A	O02	n		7 G2 CH 244 m	ich es nicht mehr sicher kann.
B	R10	g		7 G2 CH 246 w	wir das Trapez noch nicht behandelt haben
C	K02	n		7 G2 CH 249 w	das Zusammenzählen nie besonders schwer war
				7 G2 CH 251 m	sie ziemlich einfach war
C	I04	a		7 G2 CH 249 w	es sicher auch auf schnellere Weise ginge
				7 G2 CH 251 m	ich keinen besseren Lösungsweg fand.
C	L15	g		7 G2 CH 249 w	ich solche Aufgaben eigentlich noch nicht gemacht habe
				7 G2 CH 250 w	mir ist kein Rechenweg eingefallen!
				7 G2 CH 251 m	zum Glück war sie nicht schwerer, denn wir rechneten noch nie mit Winkeln, doch man kann es durch überlegen herausfinden.
C	Q06	n			
C	R14	p		7 G2 CH 249 w	ich schon länger keine solche Aufgabe mehr gemacht habe
C	Q01	a		7 G2 CH 251 m	Eigentlich hatte ich die Lösung schon, habe es nicht gemerkt und sie durchgestrichen.
C	R08*	d		7 G2 CH 249 w	ich den Rechenweg nicht mehr genau weiss
				7 G2 CH 250 w	mir ist kein Rechenweg eingefallen!
				7 G2 CH 251 m	(wieso einfach wenn's auch kompliziert geht)
C	M04	n		7 G2 CH 249 w	es schon länger her ist, dass ich eine solche Aufgabe gelöst habe
				7 G2 CH 251 m	Text „„kann man leicht erraten“ geändert in „„könnnte““
C	U2a	m		7 G2 CH 249 w	sie für mich nicht sehr schwer ist
C	U2b	m		7 G2 CH 249 w	sie für mich nicht sehr schwer ist
C	O03	g		7 G2 CH 249 w	ich noch keine solche Aufgabe gehabt habe
				7 G2 CH 250 w	dieses Thema nehmen wir gerade durch und ich komme nicht 100% nach!
				7 G2 CH 251 m	wir noch nie mit Winkeln rechneten
C	R13	n		7 G2 CH 249 w	ich solche Aufgaben nicht schwer finde
				7 G2 CH 251 m	ich sie schnell hatte
C	K04	a		7 G2 CH 249 w	ich noch keine solche Aufgabe hatte
D	I06	n		7 G2 CH 253 m	man für das Ergebnis gar keinen Rechenweg benötigt.
D	J16	g		7 G2 CH 254 w	ich nicht mehr weiss, ob man zuerst x oder y schreibt!
D	N18	d		7 G2 CH 252 w	D kann es nicht sein, denn sonst müsste es gleich viele Chips mit einer geraden Zahl, wie mit einer ungeraden Zahl haben.
D	M08	n		7 G2 CH 254 w	ich sie mündlich nicht kann
D	N15	m		7 G2 CH 253 m	ich Winkelberechnen noch nicht gelernt habe.

Transskript der Zusatzbemerkungen in den Schülerquestionnaires 7 G2 CH					
Die Kommentare wurden originalgetreu (ohne Rechtschreibkorrekturen) transkribiert.					
1	2	3	4	5	Zu der Aufgabe möchte ich noch sagen, dass:
D	K06	n		7 G2 CH 252 w	ich sehr lange an dieser Aufgabe gerechnet habe.
E	V04	m		7 G2 CH 255 w	wir es vor kurzer Zeit gelernt haben. Wir haben aber noch nicht geübt.
				7 G2 CH 256 w	Wir haben erst gerade angefangen mit Geometrie.
				7 G2 CH 257 m	der Rechnungsweg eigene erfundung ist.
E	I08	g		7 G2 CH 255 w	ich nicht mehr genau weiss, wie man sie löst.
E	P16	n			
E	I01	a		7 G2 CH 255 w	ich zwar wüsste, wie man sie vielleicht lösen könnte, doch ich komme nicht weiter.
				7 G2 CH 256 w	ich erreichte meine Lösung durch Ausprobieren.
				7 G2 CH 257 m	ich bei den verschiedenen Lösungen nicht nach komme, was genau gemeint ist.
F	F01	n			
F	F02	a		7 G2 CH 260 w	ich nicht weiss, wie man das rechnen könnte
F	F03	g		7 G2 CH 259 m	ich wirklich bei allen Nummern geraten habe, da mir kein Rechenweg eingefallen ist. Ich nehme aber an, dass man mit den je zwei Winkelangaben rechnen muss.
				7 G2 CH 260 w	ich diese Zeichen noch nie gesehen habe!
F	F04	n		7 G2 CH 260 w	wir noch nie solche Aufgaben gelöst haben, aber die Aufgabe war trotzdem nicht so schwer.
F	F05	p		7 G2 CH 259 m	wir bisher noch nichts über die Schwerlinien gehört haben, wir haben nur verschiedene Dreiecke angeschaut
				7 G2 CH 260 w	ich weiss wirklich nicht, was Schwerlinien sind, und ich kann mir auch nichts darunter vorstellen.
F	F06	a			
F	F07	d		7 G2 CH 259 m	ich mir noch eine Häuschenreihe mehr vorstellen musste, als ich die Figur spiegelte
				7 G2 CH 260 w	ich mit den minus-Zahlen nicht ganz nachkomme.
F	F08	n		7 G2 CH 259 m	wir noch keine Minuszahlen in Klammern hatten
F	F09	m		7 G2 CH 259 m	wir noch nicht viel mit negativen Zahlen gerechnet haben
F	F11	n		7 G2 CH 260 w	ich bei der 2. aufgabe nicht genau weiss wie sie zu rechnen ist.
F	F12	a		7 G2 CH 260 w	ich nicht weiss wie rechnen

10.8 Zum Themenkomplex ‚testwiseness‘

Unter <http://www.upei.ca/~xliu/measurement/glossary.htm#T> findet sich folgende Definition von testwiseness:

„Testwiseness: the ability of a test taker to identify flaws in items that give away the correct answers; skill at taking tests and outwitting poor item writers.“

Unter http://www.shsu.edu/~counsel/test_taking.html werden folgende Strategien zur Lösung von multiple choice tests vorgeschlagen:

- ✓ Anticipate the answer before you look at them.
- ✓ Read over all of your options.
- ✓ Eliminate highly implausible answers.
- ✓ Some examiners give away answers in their tests. By answering one question, you may be able to deduce the answer to another.
- ✓ If you must guess, keep in mind the following tips for multiple choice tests:
 - ✓ Sometimes lengthy or highly specific answers will be the correct answer.
 - ✓ Be aware of words like "always", "never", "only", "must", and "completely". These are usually the wrong answers since there are many exceptions to rules. These are extreme words which are more than likely to be the wrong choice.

In 1.6.2 wurde ein weiterer Weg zur Bearbeitung von MC-Aufgaben beschrieben: auf Grund der Vorgaben können recht einfach drei der vier Lösungen ausgeschlossen werden, ohne dass der Proband die eigentlich geforderte Aufgabe lösen muss. Auch hier gilt: MC-Abfragen greifen nur dann, wenn in der Antwortskala die Option ‚keines von allem‘ mit eingearbeitet wird.

10.9 Details zur ‚mathematical literacy‘

In der Projektbeschreibung des Max-Planck-Instituts für Bildungsforschung in Berlin zum PISA-Projekt wird die ‚mathematical literacy‘ klassifiziert nach drei Kategorien:

- mathematische Inhalte,
- Prozesse der Mathematik und
- Situationen, in denen Mathematik angewendet werden kann.
(nach MPI PISA, 2000)

Die Aspekte der mathematischen Grundbildung, die dem PISA-Test zugrunde lagen, werden wie folgt beschrieben:

Für den OECD/PISA-Test sind vier Aspekte maßgeblich. Die beiden Hauptaspekte sind ‚mathematische Kompetenzen‘ und ‚mathematische Leitideen‘. Die beiden Nebenaspekte sind curriculare Teilbereiche der Mathematik und ‚Situationen und Kontexte‘.

Die Hauptaspekte dienen dazu, den Bereich, der mit dieser Erhebung erfasst wird, abzustecken und Leistungen zu beschreiben. Die Nebenaspekte werden verwendet, um zu gewährleisten, dass der Bereich angemessen breit abgedeckt und ein ausgewogenes Spektrum von Aufgaben ausgewählt wird.

Mathematische Kompetenzen:

- Ausgangspunkt ist eine Liste allgemeiner mathematischer Kenntnisse, Fertigkeiten und Fähigkeiten:
- die Fähigkeit, mathematisch zu denken
- die Fähigkeit, mathematisch zu argumentieren
- die Fähigkeit zur mathematischen Modellierung
- die Fähigkeit, Probleme zu stellen und zu lösen
- die Fähigkeit, mathematische Darstellungen zu nutzen
- die Fähigkeit, mit den symbolischen, formalen und technischen Elementen der Mathematik umzugehen
- die Fähigkeit, zu kommunizieren
- die Fähigkeit, Hilfsmittel einzusetzen und zu gebrauchen.

Kompetenzklassen:

Im Rahmen von OECD/PISA ist nicht vorgesehen, Items zu entwickeln, mit denen diese Fähigkeiten jeweils getrennt erfasst werden. Wenn man echte Mathematik betreibt, kommen gewöhnlich viele (wenn nicht alle) dieser Fähigkeiten gleichzeitig zum Einsatz, so dass jeder Versuch der Messung einzelner Fähigkeiten zu künstlichen Aufgaben und damit zu einer unnötigen Aufspaltung des Bereichs der mathematischen Grundbildung führen würde.

Um den Aspekt der *mathematischen Kompetenzen* durch die Konstruktion von Items und Tests operationalisieren zu können, ist es hilfreich, die Fähigkeiten zu drei größeren Kompetenzklassen zusammenzufassen. Diese drei Kompetenzklassen sind:

- Kompetenzklasse 1: Wiedergabe, Definitionen und Berechnungen
- Kompetenzklasse 2: Querverbindungen und Zusammenhänge herstellen, um Probleme zu lösen
- Kompetenzklasse 3: Einsichtsvolles mathematisches Denken und Verallgemeinern (Bauermert, Klieme et al., 2000)

Vergleicht man diese Kompetenzklassen nach PISA mit der dreistufigen kategorialen Klassifizierung der Abituraufgaben, wie sie nach den Richtlinien des Landes NRW in Mathematik vorgeschrieben sind (vgl. 2.3.3), so stellt man Identität fest.

Für die Abiturvorschläge in Mathematik muss zur Vorlage bei der Dienstaufsicht für jede einzelne Teilaufgabe der prozentuale Anteil der Stufen bewertet werden. Für den Gesamtvorschlag werden sogar in den Richtlinien und von der Fachauf-

sicht je nach Grund- oder Leistungskurs differierend prozentuale Anteile der Stufen oder Kompetenzklassen in engen Grenzen vorgegeben.

Die Mathematik-Fachlehrer am Gymnasium in NRW arbeiten daher in der Oberstufe schon seit sehr langen Jahren mit dem dreistufigen Kompetenzklassenmodell.

An der Umorientierung der Aufgaben und Zielvorgaben des Mathematikunterrichtes auch in der Sekundarstufe Eins sind mehrere Aspekte bemerkens- und überdenkenswert:

- Die Zielvorstellungen des traditionellen Mathematikunterrichtes werden umgestellt von einer Liste bestimmter Fähigkeiten und Fertigkeiten auf die Erzeugung einer ‚mathematischen Grundhaltung‘. Allgemeine mathematische Kenntnisse, Fertigkeiten und Fähigkeiten oder konkreter die Fähigkeit, mit den symbolischen, formalen und technischen Elementen der Mathematik umzugehen, sind nach dieser Sichtweise lediglich Ausgangspunkt einer wesentlich breiteren Zielorientierung mit der Vorstellung, gewisse fundamentale Ideen der Mathematik im Schüler zu verankern.
- Die tradierten Ziele des Mathematikunterrichtes werden durch das Forschungsziel neu definiert. Die Forschung übernimmt hier zusätzlich die Aufgabe, im Sinne von Mouly den Fortschritt voranzutreiben:
- „Research is best conceived as the progress of arriving at dependable solutions to problems through the planned and systematic collection, analysis, and interpretation of data. It is a most important tool for advancing knowledge, *for promoting progress*, and for enabling man to relate more effectively to his environment, to accomplish his purposes, and to resolve his conflicts. (Mouly, 1978)
- Es scheint vorhersagbar, dass Testitems, die dieser Auffassung von Mathematikunterricht entsprechen, weit von tradierten Testitems oder üblichen ‚Klassenarbeiten‘ entfernt sind.
- Es bleibt zu untersuchen, wie und wie weit prinzipiell Fähigkeiten der Kompetenzklassen zwei und drei durch zeitlich lokalisierte Tests und statistische Auswertungen erfassbar sind. Zumindest erscheint dieses Vorhaben als so anspruchsvoll, dass es kaum durch MC-Tests wie in TIMSS realisiert werden kann. Zur Zeit werden auf Initiative der Europäischen Kommission nicht nur im Bereich der Mathematik, sondern z.B. auch für Physik, written expressions, key skills, marketing etc. entsprechende Forschungsprojekte realisiert. Finanziert über die Programme SOKRATES und LEONARDO DA VINCI soll gezeigt werden, dass Tests zur Bewertung von Wissen und Kompetenzen entwickelt werden können, die prinzipiell auch online absolviert werden können. Für den Bereich der Mathematik ist dies das Projekt AEVEM: „Bewertung von Mathematikkenntnissen mittels eines Multimedia-PC. Ziel des Projekts ist die Entwicklung eines Systems zur Bewertung von Mathematik-

kenntnissen mittels eines Multimedia-PC. Evaluiert werden soll anhand von Tests, die mit hierarchisch gegliederten Modulen (vom niedrigsten bis zum höchsten Kenntnisstand) in Verbindung stehen. Der Test kann über ein Server-Zentrum, das Fragen generiert, und über das Internet verbreitet werden.“ (online, Europäische Kommission, 2000)

- Es besteht jedenfalls Bedarf zu spezifizieren, welcher Grundbestand an mathematischem Sachwissen z.B. zur Lösung konkreter individueller und gesellschaftlicher Probleme, zur Steigerung der Abstraktionsfähigkeit, zur Beherrschung von Symbolsprachen, zur Sicherheit im Umgang mit Gedankenmodellen oder auch zum Einblick in die Wissenschaftssystematik vermittelt werden soll.
- Andererseits impliziert die Festlegung eines Basiscurriculums noch nicht, dass mit seiner erfolgreichen Vermittlung das geschaffen wäre, was mit dem Begriff „mathematical literacy“ umschrieben wird und in Folge der TIMS-Studien als Bildungsziel eingefordert wird.

Einige Worte des Dankes

Mein ganz besonderer Dank gilt Herrn Prof. Dr. Günter Törner für seine Bereitschaft, mich als externen Doktoranden aufzunehmen und zu betreuen. So erhielt ich die Gelegenheit, mich wissenschaftlich forschend mit einem Thema zu beschäftigen, welches nicht nur gesellschaftspolitisch aktuell ist.

Professor Törner war immer für Fragen und Probleme ansprechbar, und immer wieder zeigte er neue Blickrichtungen oder andere Strukturierungen auf, um die Problematik weiter auszuleuchten.

Herrn Prof. Dr. Leppig danke ich dafür, dass er sich bereit erklärt hat, die Dissertation als Zweitkorrektor zu bewerten.

Ein herzlicher Dank gebührt Herrn Prof. Dr. Armin Hollenstein von der Forschungsstelle für Schulpädagogik und Fachdidaktik in Bern. Erst durch seine vermittelnde Hilfe entstand der Kontakt zu Schulen in Bern, wo die Schweizer Tests durchgeführt wurden. Seine Ablaufplanung in dem mir aus sachlichen Erwägungen gesetztem engen Zeitfenster war hervorragend.

Ich danke den Schulleitern der Berner Schulen dafür, dass sie mir ihre Schulen freundlich geöffnet haben. Einige kollegiale Gespräche haben mir deutlich länderspezifische Unterschiede in schulischen und außerschulischen Bedingungen aufgezeigt.

Allen beteiligten Lehrerinnen und Lehrern in der Schweiz und in Deutschland danke ich für die Arbeit, die sie sich zusätzlich zur Testbetreuung mit der sorgfältigen Bearbeitung der umfangreichen Lehrerquestionnaires gemacht haben.

Eidesstattliche Versicherung

Hiermit versichere ich,

Reinhard Woschek

aus Mönchengladbach,

an Eides statt,

dass ich die vorliegende Arbeit

in allen Teilen allein

und ohne Hilfe Dritter

verfasst habe.

Mönchengladbach, den 9.5.2005

Woschek

