

Köller, Olaf; Trautwein, Ulrich

## **Evaluation mit TIMSS-Instrumenten. Untersuchungen in der 8. Jahrgangsstufe an fünf Gesamtschulen**

*Die Deutsche Schule 93 (2001) 2, S. 167-185*



Quellenangabe/ Reference:

Köller, Olaf; Trautwein, Ulrich: Evaluation mit TIMSS-Instrumenten. Untersuchungen in der 8. Jahrgangsstufe an fünf Gesamtschulen - In: Die Deutsche Schule 93 (2001) 2, S. 167-185 - URN: urn:nbn:de:0111-pedocs-275493 - DOI: 10.25656/01:27549

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-275493>

<https://doi.org/10.25656/01:27549>

### **Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

Mitglied der

  
Leibniz-Gemeinschaft

---

Olaf Köller und Ulrich Trautwein

## Evaluation mit TIMSS-Instrumenten

Untersuchungen in der 8. Jahrgangsstufe an fünf Gesamtschulen<sup>1</sup>

---

Nach der ersten Veröffentlichung der Befunde aus der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie TIMSS (vgl. Baumert, Lehmann u.a. 1997) hat in der Bundesrepublik Deutschland eine bemerkenswerte *empirische Wende der Schulentwicklungsforschung* stattgefunden, die sich gegenwärtig in erster Linie auf ein globales Systemmonitoring bezieht, bei dem das gesamte Schulsystem einer Nation bzw. unterschiedliche Organisationsformen – in Deutschland Schulformen – zum Gegenstand empirischer Untersuchungen gemacht werden. In der Bundesrepublik Deutschland wurden in verschiedenen Bundesländern in ausgewählten Jahrgängen und Schulfächern Erhebungen durchgeführt (beispielsweise LAU in Hamburg, QUASUM in Brandenburg und MARKUS in Rheinland-Pfalz), die neben dem globalen Systemmonitoring auch Auswertungen auf Einzelschulebene ermöglichen. Schulen sind somit in der Lage zu überprüfen, inwieweit selbst gesteckte Ziele der Schulentwicklung erreicht wurden.

Die TIMSS-Studie wurde und wird in der Öffentlichkeit vor allem als eine Studie wahrgenommen, die Vergleiche des Leistungsstandes in Mathematik und den Naturwissenschaften zwischen verschiedenen Nationen ermöglicht. Andere Aspekte der Studie kamen demgegenüber in der öffentlichen Wahrnehmung eher zu kurz:

- Erstens wurde unseres Erachtens selten gewürdigt, dass mit den Befunden aus TIMSS wichtige Erkenntnisse über die Effizienz des deutschen Sekundarschulsystems vorliegen, und dies, obwohl in zwei nationalen Berichtsbänden (Baumert, Lehmann u.a. 1997; Baumert, Bos/Lehmann 2000a, 2000b) ausführlich die Befunde in deutschen Schulen dargestellt wurden.<sup>2</sup>
- Zweitens ist u. E. in unzureichendem Maße gewürdigt worden, *dass in TIMSS neben der Leistung in standardisierten Mathematik- und Naturwissenschaftstests auch ein breites Spektrum motivationaler, emotionaler und sozialer Variablen auf Seiten der Schülerinnen und Schüler erhoben wurde*, das

---

1 Diese Untersuchung wurde durch Mittel des hessischen Kultusministeriums unterstützt. Die Verantwortung für den Beitrag liegt bei den Autoren und bildet nicht notwendigerweise die Sicht des hessischen Kultusministeriums ab.

2 Die Tatsache, dass aufgrund der nationalen Analysen gezeigt werden konnte, dass beinahe ein Fünftel der deutschen Schülerinnen und Schüler am Ende der 8. Jahrgangsstufe ein Kenntnisniveau aufwiesen, wie man es entsprechend den Lehrplänen am Ende der Grundschule erwartet (vgl. Baumert, Lehmann u.a. 1997), hat den hohen Wert von TIMSS jenseits irgendwelcher internationalen Ranglisten belegt.

wenigstens zum Teil Analysen erlaubt, inwieweit schulische Ziele jenseits der Fachleistungen erreicht werden. Auf methodische Grenzen gerade bei der Erhebung solcher psychosozialer Variablen haben wir anderenorts hingewiesen (vgl. Baumert/Köller 1998).

- Drittens wurde wenig beachtet, dass die Ergebnisse der Fachleistungstests und der zusätzlichen Messinstrumente eine besondere Chance bieten: Da auf der Basis der TIMSS-Daten quasi national repräsentative Normen, z. B. von Schulformen, bestimmt werden können, *kann sich jede einzelne Schule*, sofern sie dafür offen ist, *an diesen Normen von TIMSS messen lassen*. Man hat so, beinahe im Sinne der standardisierten psychologischen Diagnostik, ein Instrumentarium, das an einer nationalen Stichprobe geeicht ist, und das den Vergleich einer Einzelschule mit den vorhandenen Normen ermöglicht. Solch eine Diagnostik auf Einzelschulebene, die Anhaltspunkte für die eigene Schulentwicklung geben kann, ist also nicht nur für die Schulen möglich, die an der TIMSS-Studie selbst teilgenommen haben. Vielmehr ist auch der Fall denkbar, dass sich Schulen mit in TIMSS eingesetzten Instrumenten in den entsprechenden Altersgruppen untersuchen lassen und anhand der rückgemeldeten Ergebnisse Schlussfolgerungen für zukünftige Entwicklungsprozesse ziehen können. Dies setzt natürlich auf Seiten der involvierten Schulen den Willen voraus, mit Befunden solcher schuldiagnostischen Maßnahmen selbstkritisch umzugehen.

Wie *zusätzliche Erhebungen unter Nutzung der Instrumente aus nationalen oder internationalen Schulleistungsstudien* aussehen können, soll in diesem Beitrag anhand einer *Ergänzungsstudie zu TIMSS mit fünf hessischen Gesamtschulen* aufgezeigt werden. Dabei wird insbesondere darauf eingegangen, wie spezifische Gegebenheiten an einer Schule (Zusammensetzung der Schülerschaft, besonderes pädagogische Profil) so berücksichtigt werden können, dass ein umfassendes und faires Gesamtbild einer Schule entsteht. Wir geben im Folgenden zunächst einen allgemeinen Überblick über das Projekt und berichten einige schulübergreifende Befunde, bevor wir anhand zweier Fallstudien wichtige Einzelaspekte der Untersuchung illustrieren.

## 1. Teilnehmende Schulen

*Die Initiative zu der vorliegenden Erhebung ging von den fünf hessischen Gesamtschulen bzw. ihren Vertretern aus.* An der Studie nahmen auf Wunsch dieser Schulen die kompletten 8. Jahrgänge teil. Vier der teilnehmenden Schulen sind Versuchsschulen des Landes Hessen, und alle Schulen zeichnen sich durch eine Vielfalt spezieller Charakteristika in Hinsicht auf Struktur, Unterrichtsorganisation und Schülerschaft aus, die an dieser Stelle jedoch nur bruchstückhaft wiedergegeben werden können. Zwei der fünf Schulen nehmen am KMK-Modellversuch zur klasseninternen Differenzierung in Englisch und Mathematik teil, eine dritte setzt zunehmend Formen des „Selbständigen Lernens“ ein, in der vierten Schule lernen die Kinder in jahrgangsübergreifenden Gruppen, und die fünfte Schule setzt ein intensives Integrationsprogramm für Kinder nicht-deutscher Muttersprache um. Neben gemeinsamen Elementen (z.B. Betonung des Projektunterrichts, offener Lernformen jenseits des Standardcurriculums) gibt es auch wesentliche Unterschiede zwischen den Schulen (z.B. Art der Leistungs-differenzierung in Mathematik). Wie in Hessen üblich, bieten alle fünf Schulen nur Unterricht in der Sekundarstufe I an. Sofern Schülerinnen und Schüler die all-

gemeine Hochschulreife anstreben, müssen sie nach der 10. Jahrgangsstufe in eine gymnasiale Oberstufe an einer anderen Schule wechseln.

Aus unserer Sicht ist es für die weitere Interpretation der Befunde aus den fünf Schulen zentral, sie in die *Gesamtschullandschaft des deutschen Sekundarschulsystems* einzuordnen. Im Wesentlichen lassen sich drei Typen von Gesamtschulen unterscheiden (vgl. AG Bildungsbericht am Max-Planck-Institut für Bildungsforschung 1994):

- Schulen des ersten Typus, *Gesamtschulen als Einzelschulen besonderer Prägung innerhalb eines reichhaltigen Schulangebots*, ergänzen in der Regel das Schulangebot einer (städtischen) Region und sichern sich hierbei ihre Attraktivität durch ein besonderes pädagogisches Programm. Liegt die Nachfrage nach solchen Gesamtschulen erheblich über der Kapazität, so verfügen diese Schulen auch über eine leistungsmäßig ausgewogene oder gar positiv selektierte Schülerschaft.
- Schulen des zweiten Typus, die *Gesamtschulen als Teil der schulischen Basisversorgung einer Kommune*, haben sich in den 80er Jahren unter gewissen Bedingungen zu einer attraktiven Lösung für kommunale Schulversorgungsprobleme entwickelt, da auch in strukturschwächeren Gebieten bzw. jungen Gemeinden durch die Einrichtung einer Gesamtschule oder durch die Umwandlung anderer Schulen in eine Gesamtschule eine Basisversorgung gewährleistet werden kann. Diese Schulen binden dann zumeist einen hohen Prozentsatz der Schüler aus dem Einzugsgebiet.
- Der dritte Typus, *Gesamtschulen unter Konkurrenzdruck*, findet sich insbesondere in städtischen Ballungsgebieten in scharfem Wettbewerb mit dem traditionellen dreigliedrigen Sekundarschulsystem. Die Schülerschaft dieser Gesamtschulen ist häufig gekennzeichnet durch den so genannten doppelten *Creaming-Effect*, also dem Fehlen von besonders starken *und* besonders schwachen Schülerinnen und Schülern.

Vier der fünf von uns untersuchten Gesamtschulen sind dem ersten Typus zuzuordnen, der durch eine attraktive Organisation der Schule bzw. des Unterrichts auch für Eltern, die eine hohe Bildungsaspiration für ihre Kinder haben, eine ernst zu nehmende Alternative zum dreigliedrigen Schulsystem darstellt. Ihre besondere Anziehungskraft ist den Schulen dabei nicht in den Schoß gefallen, sondern sie ist Konsequenz einer bewussten pädagogischen Profilbildung der jeweiligen Schule. Eine der untersuchten Schulen beispielsweise wäre vor 15 Jahren sicherlich eher dem dritten Typus zuzuordnen gewesen, hat sich dann aber in Eigeninitiative ein Programm gegeben, wichtige strukturelle und pädagogische Veränderungen durchgeführt und sich dadurch zu einer *Gesamtschule besonderer Prägung* weiterentwickelt. Die fünfte untersuchte Gesamtschule ist als Teil der schulischen Basisversorgung einer Kommune zu sehen. Auch bei dieser Schule ist jedoch eine besondere pädagogische Prägung unübersehbar.

## 2. Ziele und Ablauf der Studie

In Abstimmung mit den teilnehmenden Schulen wurden folgende Ziele der Untersuchung vereinbart:

(1.) *Feststellung des Leistungsstands in Mathematik und (optional) in den Naturwissenschaften*. Da die Original-Instrumente von TIMSS verwendet wur-

den, sollten die Leistungen in den einzelnen Schulen mit denen der national repräsentativen TIMSS-Stichprobe verglichen werden. Darüber hinaus äußerten die einzelnen Schulen Interesse an einer differenzierteren Auswertung der Leistungstestergebnisse (Geschlechtsdifferenzen, Leistungen von Schülerinnen und Schülern mit nicht-deutscher Muttersprache, Effekte des sozialen Hintergrunds, spezifische Stärken und Schwächen je nach Aufgabeninhalt, -typus etc.).

(2.) *Feststellung des motivationalen und psychosozialen Status* der Schülerinnen und Schüler. Im Bewusstsein, dass pädagogische Ziele von Schule sich nicht auf den Leistungsbereich beschränken, wurde vereinbart, zentrale Indikatoren für Lernmotivation und emotionales Wohlbefinden zu erheben.

(3.) *Feststellung von Indikatoren sozialen Lernens*. Da im Rahmen einer Fragebogenstudie Verhaltensbeobachtungen kaum möglich sind, wurde vereinbart, Ergebnisse sozialen Lernens zu erheben. Dafür eignen sich z.B. Skalen, die „Einstellungen gegenüber Minoritäten“ oder „selbst berichtete Normverletzungen“ erfragen.

(4.) *Beurteilung des Mathematikunterrichts*, um Informationen zur Optimierung von Unterrichtsprozessen zu erhalten.

Von Seiten der Mitarbeiter am Max-Planck-Institut für Bildungsforschung bestand das Interesse, systematisch zu untersuchen, mit welchen Befunden in Gesamtschulen zu rechnen ist, die besondere pädagogische Aufgaben im Sekundarbereich übernehmen. Die im Rahmen der Längsschnittstudie „Bildungsprozesse und psychosoziale Entwicklung im Jugendalter (BIJU)“ (vgl. z. B. Baumert u.a. 1996; Baumert/Köller 1998; Baumert, Köller/Schnabel 2000; Köller 1998; Köller, Baumert/Schnabel 1999; Schnabel 1998) gewonnenen Leistungskennwerte haben eine größere Diskussion über die Leistungsfähigkeit von Gesamtschulen ausgelöst, die so weit ging, dass von verschiedenen Seiten gefordert wurde, die Gesamtschulen abzuschaffen. Von unserer Seite wurde wiederholt argumentiert, dass die Schulform per se (dreigliedriges Schulsystem vs. Gesamtschule) *kein* bestimmender Faktor für die Leistungsentwicklung ist, sondern die jeweilige Ausformung einer Schule bzw. Schulart dafür verantwortlich ist, wie sich die Leistungen und psychosozialen Charakteristika ihrer Schüler entwickeln (Baumert/Köller 1998; Baumert, Lehmann u.a. 1997; Köller et al. 1999). Die Untersuchung an den fünf hessischen Gesamtschulen konnte dabei als Bewährungsprobe für diese Annahmen genutzt werden.

Die vorliegende Studie wurde seit 1998 geplant; die Datenerhebung fand im Frühjahr 1999 gegen Ende des 8. Schuljahres statt. Die Auswertung wurde mit dem Überreichen ausführlicher Abschlussberichte für die Schulen im Mai 2000 vorerst beendet. Hinsichtlich der Verwendung der Untersuchungsergebnisse wurde vereinbart, dass die Adressaten der Untersuchungsergebnisse primär die Schulen sein sollten und Veröffentlichungen der Ergebnisse der Zustimmung der Schulen bedürften. Die ausführlichen am Max-Planck-Institut entstandenen Berichte geben den Schulen Interpretationshilfen und weisen auf mögliche Optimierungsmöglichkeiten hin. In einer derzeit entstehenden Buchpublikation werden von Seiten der untersuchten Schulen unter anderem Ausführungen darüber zu lesen sein, welche Unterrichts- bzw. Schulentwicklungsmaßnahmen sich für sie aus den Befunden ergeben haben.

### 3. Instrumente

Bei der Planung der Untersuchung war es für alle beteiligten Schulen zentral, die eigenen Ergebnisse in den Mathematiktests an den Befunden der national repräsentativen TIMSS-Stichprobe zu messen.<sup>3</sup> Aus dem Pool der Aufgaben wurde eine Auswahl von vier Testheften eingesetzt, die die Struktur des Gesamttests gut wiedergibt. Die Tabelle 1 belegt dies für den Mathematiktest. Anhand einer Matrix wird gezeigt, wie viele Aufgaben im Vergleich zur TIMS-Studie aus den verschiedenen mathematischen Sachgebieten und kognitiven Anforderungsdimensionen eingesetzt wurden. Grenzen dieser Beschränkung auf vier Testhefte lagen darin, dass sich keine sinnvollen Leistungswerte auf den einzelnen Anforderungsdimensionen und Sachgebieten bilden ließen. Die Gesamtwerte ließen sich allerdings auf dem in TIMSS verwendeten Maßstab ( $M = 500$ ,  $SD = 100$ ) in der international untersuchten Kohorte abbilden, so dass die fünf untersuchten Schulen direkt mit den TIMSS-Schulen verglichen werden können.

Als *Vergleichsstichproben* wurden Bundesländer mit einem ähnlich ausgebauten Sekundarschulwesen ausgewählt, nämlich die Länder Berlin, Brandenburg, Bremen, Hamburg, Hessen, Nordrhein-Westfalen, Saarland und Thüringen. Um einen fairen Vergleich zwischen diesen Ländern und den fünf Gesamtschulen zu ermöglichen, wurden zusätzlich die Studentafeln von der 5. bis zur 8. Jahrgangsstufe analysiert. Zudem wurde berücksichtigt, dass in allen fünf untersuchten Schulen der reguläre Fachunterricht bei der Realisierung von Projekten zum Teil aufgelöst wird. Weiterhin wurde von den Mathematiklehrern der fünf untersuchten Schulen eine Einschätzung der Testaufgaben erbeten. Dabei soll-

Tabelle 1: Mathematische Testaufgaben nach Sachgebiet und kognitiver Anforderungsart; die Werte ohne Klammern geben die Aufgabenzahlen der Untersuchung in den hessischen Gesamtschulen an, die Werte in den Klammern die entsprechenden Zahlen aus der TIMS-Studie:

Sachgebiet	Anforderungsart				Insgesamt
	Wissen	Routineverfahren	komplexe Verfahren	Probleme <sup>*)</sup>	
Zahlen und Zahlenverständnis	6 (10)	5 (13)	9 (12)	7 (17)	27 (52)
Messen und Maßeinheiten	3 (6)	1 (2)	3 (5)	8 (8)	15 (21)
Algebra	5 (8)	5 (10)	0 (1)	8 (10)	18 (29)
Geometrie	3 (5)	2 (6)	4 (6)	3 (6)	12 (23)
Proportionalität	0 (0)	5 (5)	0 (0)	3 (7)	8 (12)
Darstellung und Analysen von Daten/Wahrscheinlichkeitsrechnung	1 (3)	2 (2)	5 (8)	4 (8)	12 (21)
Insgesamt	18 (32)	20 (38)	21 (32)	33 (56)	92 (158)

<sup>\*)</sup> Anwendungsbezogene und mathematische Probleme

3 Auf die Schilderung der Ergebnisse des naturwissenschaftlichen Aufgabenbereichs wird aus Platzgründen verzichtet.

ten die betroffenen Fachlehrer für jede Testaufgabe angeben, in welcher Klassenstufe die jeweiligen Kenntnisse geübt wurden und ob die Lösung der Aufgabe ihren didaktischen Zielsetzungen entspricht. Zusätzlich sollten sie schätzen, welcher Prozentsatz der Schüler die einzelnen Aufgaben voraussichtlich lösen würde.

Eine berechtigte Kritik an der Qualitätskontrolle von Schulen mit Hilfe von Leistungstest liegt darin, dass die Leistung der Schülerschaft natürlich nicht nur von der Qualität des Unterrichts in den jeweiligen Schulen abhängt, sondern auch von den *Voraussetzungen, die Schüler aus dem Elternhaus (und in unserer Studie auch aus der Grundschule) mitbringen*. Um diese außerschulischen Einflüsse abschätzen zu können, wurde eine Anzahl von Fragen zum sozio-ökonomischen Hintergrund und zum so genannten „kulturellen Kapital“ (Bildungshintergrund) der teilnehmenden Schüler eingesetzt. So wurde ein *Prestigewert für den Beruf der Eltern* (Treiman 1977) als Indikator des sozio-ökonomischen Hintergrundes ermittelt und Hinweise auf das kulturelle Kapital der Schülerinnen und Schüler durch die Frage nach dem *Bildungsabschluss der Eltern* sowie nach der *Anzahl von Büchern* im elterlichen Haushalt gewonnen. Beide Indikatoren bieten empirische Hinweise, ob und welche Rolle der familiäre Hintergrund für gelingende Entwicklungsprozesse spielt. Außerdem wurden die *Grundschulempfehlungen* erhoben und erfragt, welche *Sprache* (deutsch vs. eine andere) in den Familien zu Hause normalerweise gesprochen wird. *Alle diese Variablen haben sich als gute Prädiktoren für die Schulleistung von Kindern herausgestellt* (z. B. Lehmann/Peek 1997). Zur Kontrolle weiterer familiärer Belastungsfaktoren wurde nach *Arbeitslosigkeit der Eltern* sowie nach *alleinerziehenden Eltern* gefragt.

Der Erfolg schulischer Erziehungsprozesse drückt sich natürlich nicht nur in Leistungskennwerten aus. Deshalb wurde mit den beteiligten Schulen ferner vereinbart, über den Leistungsbereich hinaus Instrumente zu psychosozialen Variablen einzusetzen (Selbstkonzeptvariablen, schulische Normverletzungen, soziale Einstellungen, s.o.), die idealerweise ebenfalls den Vergleich mit der TIMSS-Stichprobe zulassen. Außerdem waren die Lehrerinnen und Lehrer daran interessiert, wie ihre Schülerinnen und Schüler den Mathematikunterricht beurteilen.

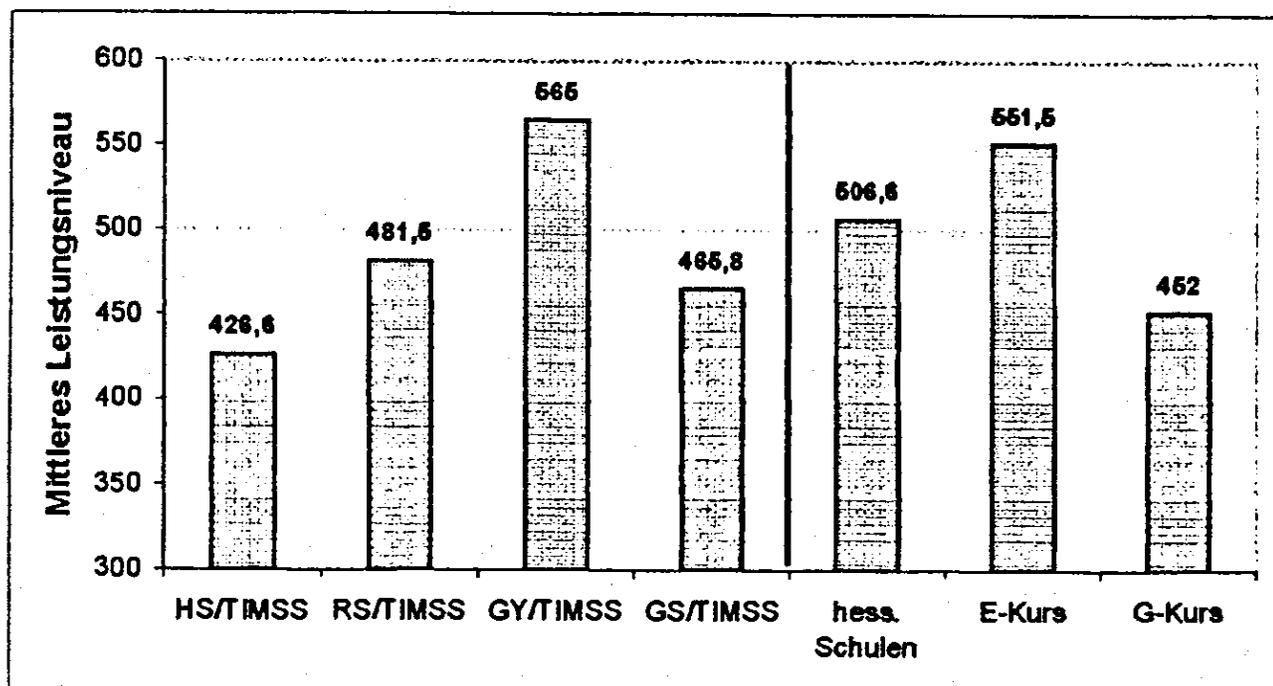
Neben den Ergebnissen der TIMSS-Erhebung liegen zu vielen dieser Variablen auch Vergleichsdaten aus der oben erwähnten BIJU-Studie vor. Die BIJU-Untersuchung ist insofern besonders interessant, als im Bundesland Nordrhein-Westfalen (NRW) Daten einer relativ großen Stichprobe von Gesamtschülern Anfang, Mitte und Ende der 7. Jahrgangsstufe erhoben wurden. Beide Bundesländer, das heißt Hessen und NRW, sind bezüglich ihrer Schulstrukturen mit einem erheblichen Anteil integrierter Gesamtschulen im Sekundarbereich I ähnlich, so dass sich direkte Vergleiche der fünf hessischen Gesamtschulen mit den Daten aus NRW anbieten. Grenzen dieses Vergleichs liegen sicherlich darin, dass an hessischen Gesamtschulen im Gegensatz zu NRW keine gymnasialen Oberstufen eingerichtet sind und die verwendeten BIJU-Daten zudem am Ende der 7. Jahrgangsstufe erhoben wurden. *Trotz dieser Einschränkungen wurden als Vergleichsstichprobe für die vorliegende Untersuchung die Daten von etwa 2.000 Schülerinnen und Schülern aus Hauptschulen, Realschulen, Gymnasien und Gesamtschulen des Landes NRW herangezogen.*

## 4. Die Ergebnisse im Überblick

Hinsichtlich der Ergebnisse der TIMSS-Mathematiktests können, insgesamt gesehen, zwei wichtige Feststellungen getroffen werden. Erstens ist der *Leistungsstand an allen fünf teilnehmenden Schulen bemerkenswert*: Die Mittelwerte liegen in einem Bereich, der aufgrund der lernförderlichen Voraussetzungen der Schülerinnen und Schüler (Muttersprache, sozio-ökonomischer Hintergrund, Grundschulempfehlungen, kulturelles Kapital in der Familie) im Einklang mit bisherigen Befunden aus der Schulleistungsforschung (z. B. Lehmann/Peek, 1997) ist. Die Abbildung 1 belegt dies. *Im Mittel erreichen die fünf Schulen einen Leistungsstand, der gutem Realschulniveau entspricht und signifikant ( $p < .01$ ) über den Leistungen der in TIMSS untersuchten Gesamtschulen liegt.* Beim Aufbrechen nach Kursniveau zeigt sich, dass die *Leistungen im Erweiterungskurs (E-Kurs) auf Gymnasialniveau liegen, in den Grundkursen zwischen Haupt- und Realschulniveau.*<sup>4</sup>

Diese Befundlage verdeutlicht, dass eine Vielzahl von zusätzlichen pädagogischen Angeboten, die teilweise Unterrichtszeit in den Kernfächern abziehen, den Erfolg des Mathematik-Unterrichts nicht beeinträchtigen muss.

Abbildung 1: Mathematikleistungen der fünf hessischen Gesamtschulen im Vergleich zu den in TIMSS untersuchten 8. Jahrgängen an verschiedenen Schulformen; HS: Hauptschule; RS: Realschule; GY: Gymnasium; GS: Gesamtschule



Zweitens können Befürchtungen, es bilde sich möglicherweise eine andere Form von Mathematikunterricht an Gesamtschulen heraus, die nicht mit der an anderen Schularten vergleichbar ist, aufgrund zusätzlicher Analysen (nicht in Abbildung 1) zerstreut werden: Es ließen sich keine sinnvoll abgrenzbaren Teilmengen von Testaufgaben finden, die den Gesamtschülern im Vergleich mit

<sup>4</sup> Da eine der untersuchten Schulen eine dreifache Differenzierung vornimmt, wurde in Abbildung 1 der leistungsstärkste Kurs dieser Schule zu den E-Kursen, die beiden anderen zu den G-Kursen gezählt. Die Schüler der Schule ohne jede Differenzierung wurden aufgrund ihrer günstigen sozialen Herkunft dem E-Kursniveau zugeordnet.

Schülern anderer Schularten besonders schwer oder leicht fielen. *Den untersuchten Gesamtschulen gelingt es offenbar in gleicher Weise, die Vorgaben der Lehrpläne (intendiertes Curriculum) umzusetzen (implementiertes Curriculum) und entsprechende Lernerfolge auf Seiten der Schüler (erreichtes Curriculum) zu erzielen.*

Was die zusätzlich eingesetzten Untersuchungsinstrumente zu Ergebnissen sozialen Lernens angeht, ist bemerkenswert, dass die von uns untersuchten Schulen verschiedene kognitive und soziale Ziele vereinbaren können. Offensichtlich kann beides gleichzeitig optimiert werden (vgl. auch Baumert/Köller 2000; Gruehn 1995; Helmke/Schrader 1990).

Im Folgenden wird auf zwei der fünf Schulen in Form anonymisierter Fallbeispiele eingegangen. Dabei werden einige besonders interessante Aspekte herausgegriffen. Die übrigen Schulen werden allein aus Platzgründen hier nicht vorgestellt.

## 5. Die Schule A

Die Schule A hat zu Beginn der 80er Jahre ihr pädagogisches Profil deutlich geändert und reformpädagogische Ansätze (Freinet, Montessori, Pestalozzi, Petersen) als Leitprinzipien „wieder entdeckt“. Die damals vorhandene Sorge, neben dem gegliederten Schulsystem nur noch eine „Restschule“ darzustellen, führte zu einer Umstrukturierung der Schule. Mit Individualisierung, freiem Lernen und der Kombination von „Leistung und Leben“ etablierte sich die Schule als echte Alternative zur Dreigliedrigkeit, die Eltern und Schülerinnen und Schüler ganz unterschiedlicher Herkunft ansprach. Die Schülerschaft, die überwiegend (80 bis 90 Prozent) aus der nahe gelegenen Grundschule auf die untersuchte Gesamtschule wechselt, wird in den Jahrgangsstufen 5 bis 10 unterrichtet. *Die Schule liegt in einem Gebiet, dessen Bevölkerung sozial stark durchmischt ist, mit einem erheblichen Anteil ausländischer Mitbürger und Aussiedlerfamilien.* Entsprechend dem hohen Ausländer- und Aussiedleranteil im Einzugsgebiet gehörten der 8. Jahrgangsstufe des Schuljahrs 1998/99 16,7 Prozent Schüler mit ausländischer Herkunft und 22,7 Prozent Kinder aus Aussiedlerfamilien an (aus Polen und der ehemaligen Sowjetunion). In Folge dieser Zusammensetzung findet man auf Seiten der Schüler eine große Heterogenität bzgl. der sozialen Herkunft und des Bildungshintergrundes. Um insbesondere die Integration der Aussiedlerkinder erfolgreich zu bewältigen, wurde an der Schule ein besonderes Konzept entwickelt, das

- sehr stark auf eine schnelle und intensive Vermittlung der deutschen Sprache abzielt, ohne dass allerdings die Kinder auf ihre Heimatsprache verzichten müssen (*Deutsch als Zweitsprache*).
- die soziale Integration durch verschiedene Maßnahmen beschleunigen soll, ohne dass die Kinder allerdings vollständig ihre kulturelle Identität aufgeben sollen.
- versucht, die Eltern der Aussiedlerkinder in den Prozess der Integration einzubinden.

Bei Schule A handelt es sich um eine *Ganztagschule*, in der bis 14.35 Uhr Unterricht stattfindet (unterbrochen von längeren Pausen), gefolgt von Zusatzange-

boten bis 16.30 Uhr, die nicht verpflichtend sind. Die Zusatzangebote sind breit gefächert (Keramik, Sport, Tanzen, Theater, Werken, Malerei uvm.) und werden nicht nur von Lehrkräften, sondern auch von Kursleitern der Volkshochschule, Musikern, Studenten und Vereinsangehörigen angeboten. Durch diese Einbindung von „anderen“ Lehrkräften erfahren die Schülerinnen und Schüler, dass Schule keineswegs ein Unternehmen ist, in dem nur Lehrer und Schüler Platz finden.

Der Schulalltag beginnt offen, d.h. die Schülerinnen und Schüler können ab 7.30 Uhr die Schule besuchen, obwohl der eigentliche Unterricht erst um 8.45 Uhr beginnt. Dadurch wird zum einen berufstätigen Eltern das Betreuungsproblem genommen, zum anderen können die Kinder diese Zeit frei für sich nutzen, aber auch Angebote wie Proben des Schulchores, Aquaristik und *Hausaufgabenbetreuung* wahrnehmen. Kinder mit Lese-Rechtschreib-Schwäche und andere Kinder mit Schwächen in Deutsch erhalten in dieser Zeit *Fördermaßnahmen*. Wie der Schulalltag offen beginnt, so endet er auch nach dem Regelunterricht mit den Zusatzangeboten am Nachmittag (s.o.). Wichtig zu erwähnen ist weiterhin, dass an der Schule zugunsten von *Projektwochen* und anderen Aktivitäten zeitweise der übliche Wochenstundenplan vollständig aufgegeben wird.

Der Unterricht ist bei verschiedenen Themen fächerübergreifend organisiert, Arbeitstechniken und Unterrichtsmethoden greifen durch das offene Konzept der Schule A auf andere Fächer über. Neben dem üblichen Fächerkanon existiert noch das Fach *Freies Lernen* (FL), das Schülerinnen und Schülern die Möglichkeit gibt, eigene Vorhaben im Schulalltag zu erledigen. Schüler übernehmen hierdurch mit zunehmendem Alter mehr Eigenverantwortung im Lernprozess. Natürlich werden sie dabei nicht frühzeitig allein gelassen, sondern in den ersten Jahrgängen systematisch angeleitet.

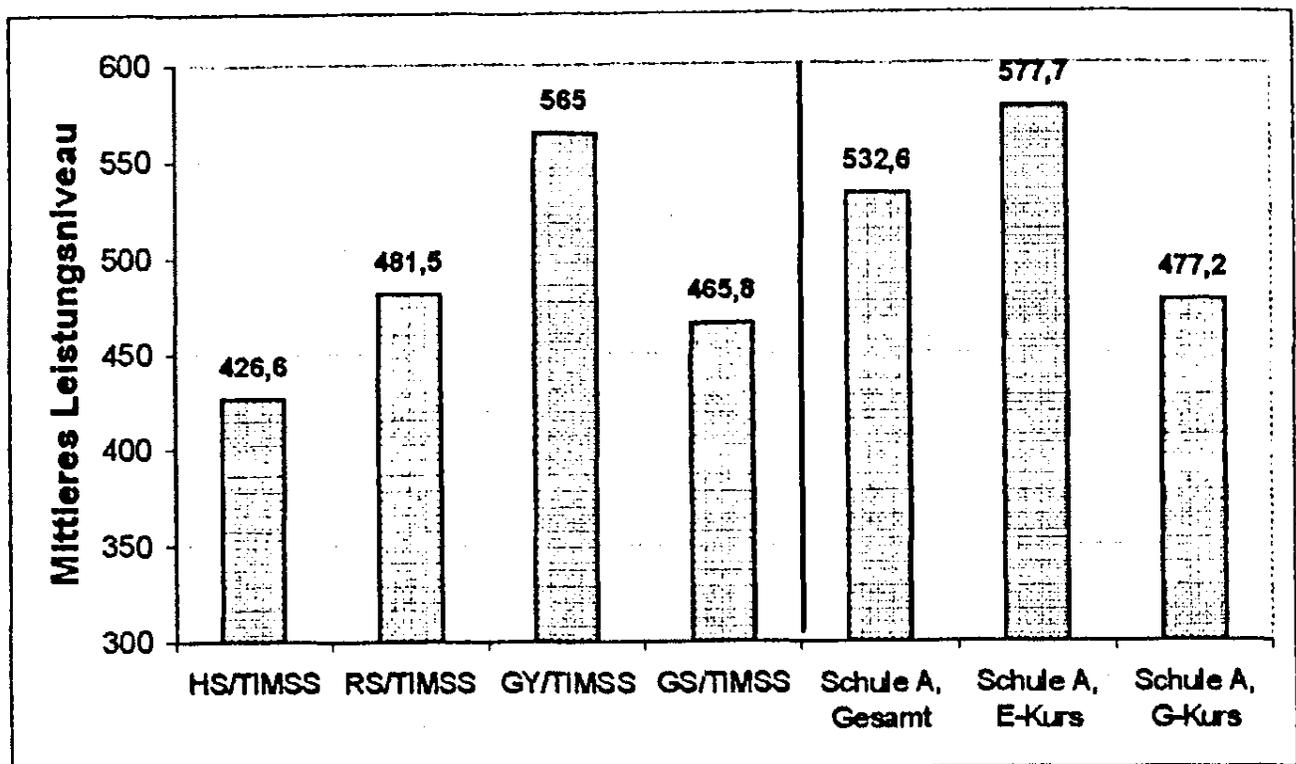
*Differenziert wird ab der 7. Jahrgangsstufe, und zwar zunächst in Englisch und Mathematik, ab der 8. Jahrgangsstufe in Deutsch und ab der 9. Jahrgangsstufe in den Naturwissenschaften und Französisch. Zwei Niveaustufen, E- und G-Kursniveau werden unterschieden. Um die Kosten einer teilweisen Auflösung des Klassenverbandes zu minimieren, werden aus jeweils zwei Klassen eines Jahrgangs (Partnerklassen) die beiden Kurse gebildet. Jede Jahrgangsgruppe hat für zwei Jahre ihren eigenen Bereich, den sie verantwortlich gestaltet und pflegt. Durch diese stabile soziale und räumliche Basis soll die Integration von Schülern sehr unterschiedlicher sozialer Herkunft in der bewusst heterogen zusammengesetzten Klasse ermöglicht werden.*

Im Folgenden wollen wir zu *drei ausgewählten Aspekten* die Ergebnisse berichten. Zunächst wird auf die Mathematikleistungen eingegangen, danach auf die besondere Problematik der Förderung nicht-deutschsprachiger Schülerinnen und Schüler, und abschließend auf die Konsequenzen der Leistungsgruppierung für psychosoziale Schülervariablen.

### *5.1 Mathematikleistungen im Vergleich zu den TIMSS-Schulen*

Die Mathematikleistungen sind in der Abbildung 2 dargestellt.

Abbildung 2: Mathematikleistungen der Schule A im Vergleich zu den in TIMSS untersuchten 8. Jahrgängen an verschiedenen Schulformen; HS: Hauptschule; RS: Realschule; GY: Gymnasium; GS: Gesamtschule



Den Vergleichsrahmen bilden erneut 8. Klassen der vier Schulformen, die an der TIMS-Untersuchung teilgenommen haben. Insgesamt zeigen die untersuchten Schülerinnen und Schüler der Schule A einen Wissensstand, der zwischen dem mittleren Realschul- und Gymnasialniveau in der TIMS-Untersuchung liegt. Getrennte Analysen für das E- und G-Kursniveau zeigen, dass im Grundkurs Realschulniveau, im Erweiterungskurs leicht überdurchschnittliches Gymnasialniveau erreicht wird. Bemerkenswert sind die Differenzen zwischen G- und E-Kursschülern. Sie korrespondieren mit einem Wissenszuwachs, wie er üblicherweise in zwei Schuljahren erreicht wird<sup>5</sup>. Dies validiert zum einen die in der untersuchten Schule vorgenommene Einteilung der Schüler, macht aber auch deutlich, dass es selbst innerhalb einer Schule mit besonderem pädagogischen Profil unmöglich ist, deutliche Leistungsdisparitäten zu vermeiden.

## 5.2 Schulleistungen von Schülerinnen und Schülern aus Migrantenfamilien

Zu den integrativen Herausforderungen der Schule A gehört die Einbindung von Schülerinnen und Schülern mit sehr unterschiedlichem sozioökonomischem und soziokulturellem Hintergrund. Ein besonderes Augenmerk gilt dabei Schülerinnen und Schülern aus Migrantenfamilien. Zu dieser Gruppe zählen wir im Folgenden Kinder aus Familien, in denen im Alltag nicht Deutsch gesprochen wird. Dort, wo die Gesamtschule fester Bestandteil der Sekundarstufe I ist, hält sie für viele Migrantenkinder Wege zu einem höheren Bil-

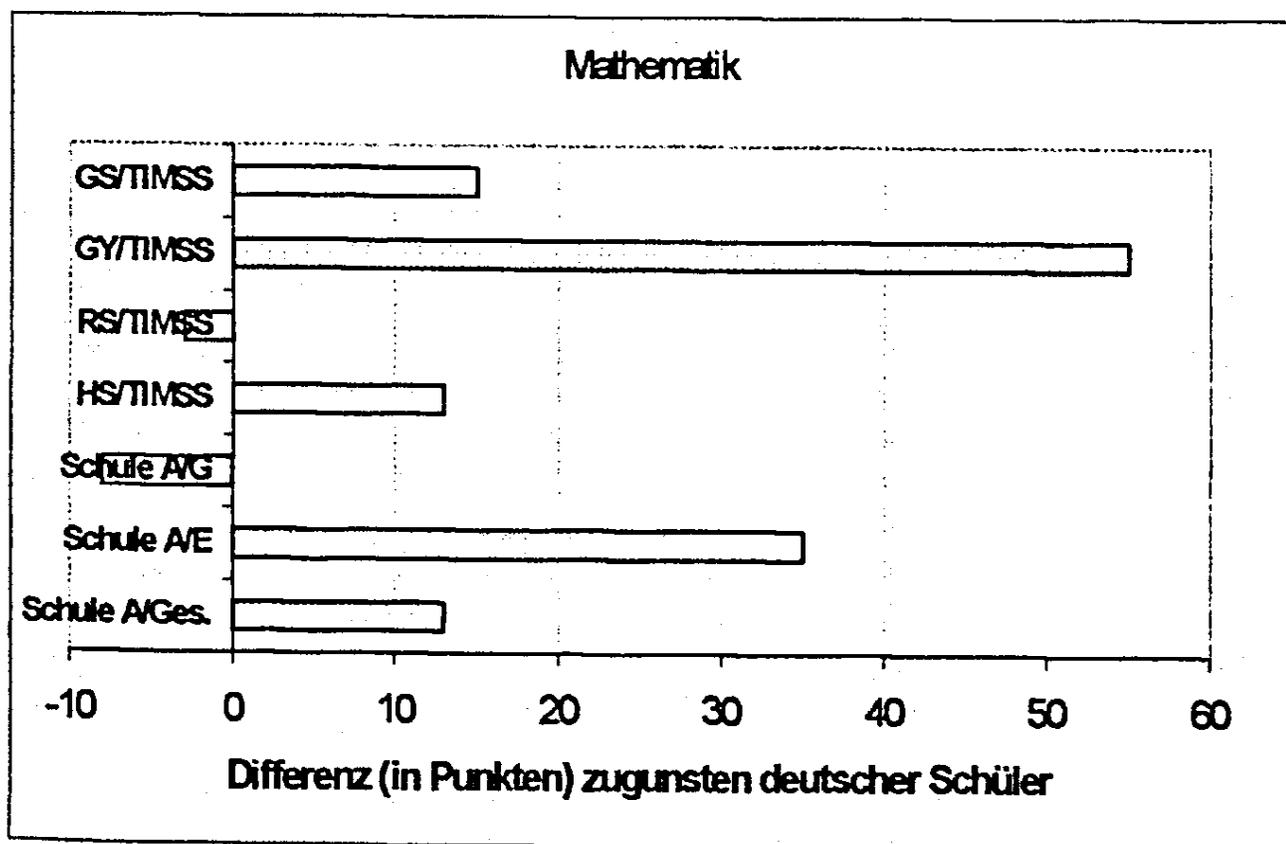
<sup>5</sup> Es hat sich gezeigt (vgl. Baumert, Lehmann u. a., 1997), dass ein zusätzliches Schuljahr bei Schülerinnen und Schülern der Mittelstufe dazu führt, dass diese einen mittleren Wissenszuwachs von etwa 40 Punkten auf der TIMSS-Skala erreichen.

dungsabschluss offen, dort, wo die Gesamtschule fehlt, entfallen überproportional viele Migrantenkinder auf die Hauptschule.

Eine Reihe von Forschungsarbeiten hat wiederholt den Nachweis von insgesamt niedrigeren Leistungen in standardisierten Tests und einer niedrigeren Bildungsbeteiligung von Kindern aus Migrantenfamilien erbracht (vgl. Arbeitsgruppe Bildungsbericht 1994; Lehmann/Peek 1997; Schwippert 1999). Dabei wurde eine Vielzahl denkbarer Ursachen für dieses unterdurchschnittliche Abschneiden im deutschen Schulsystem diskutiert (vgl. z. B. Bender-Szymaniowski/Hesse 1987; Lehmann/Peek 1997). An vorderer Stelle stehen u.a. Hinweise auf die ökonomische Situation der Familien, auf mangelnde elterliche Unterstützung, auf die kulturelle Distanz zwischen Elternhaus und Schule, sowie auf Begabungsunterschiede und Sprachprobleme der Kinder im monolingualen Bildungssystem Deutschlands. Damit zusammenhängend wird die Frage erörtert, inwieweit es im deutschen Bildungssystem zu Benachteiligungen – oder auch Bevorzugungen – von Migrantenkindern kommt (Lehmann/Peek 1997; Nauck/Diefenbach 1997), und wie das Ziel, eine der Begabung des jeweiligen Kindes angemessene Ausbildung zu gewährleisten, erreicht werden kann.

Etwas ein Viertel (23,0 Prozent) der Schülerinnen und Schüler aus Schule A gibt an, dass bei ihnen zu Hause normalerweise nicht Deutsch gesprochen wird. Diese mögliche sprachliche Benachteiligung spiegelt sich interessanterweise nicht in der Kurszugehörigkeit wider: *Von den Jugendlichen mit nicht-deutscher Muttersprache finden sich 59,3 Prozent in Mathematik im E-Kurs*; dieser Prozentsatz liegt sogar leicht über dem der Schüler, in deren Elternhaus üblicherweise Deutsch gesprochen wird (55,3 Prozent). Die Abbildung 3 zeigt die Leistungsunterschiede zwischen beiden Gruppen. Positive Differenzen signalisieren Leis-

Abbildung 3: Leistungen in Mathematik (Mittelwerte) von Schülern mit deutscher bzw. nicht-deutscher Muttersprache; Vergleichsstichproben aus TIMSS; HS: Hauptschule; RS: Realschule; GY: Gymnasium; GS: Gesamtschule; G: G-Kurs; E: E-Kurs; Ges.: Gesamte Schülerschaft der Schule A



tungsvorteile der deutschen Schülerinnen und Schüler. Bei isolierter Betrachtung der TIMSS-Ergebnisse ist zum einen erkennbar, dass am Gymnasium die Leistungsunterschiede zwischen Schülern mit deutscher und nicht-deutscher Muttersprache im Vergleich zu den anderen Schulformen stärker sind.

An der Schule A ergeben sich insgesamt kleine Unterschiede zugunsten der Schüler mit deutscher Muttersprache. Bezogen auf die beiden Kursniveaus zeigt sich, dass die Unterschiede allein auf das E-Kursniveau zurückzuführen sind, auf dem sich bemerkenswerte Differenzen zugunsten von Schülern mit deutscher Muttersprache zeigen, die mit einem Wissenszuwachs von einem Schuljahr korrespondieren.

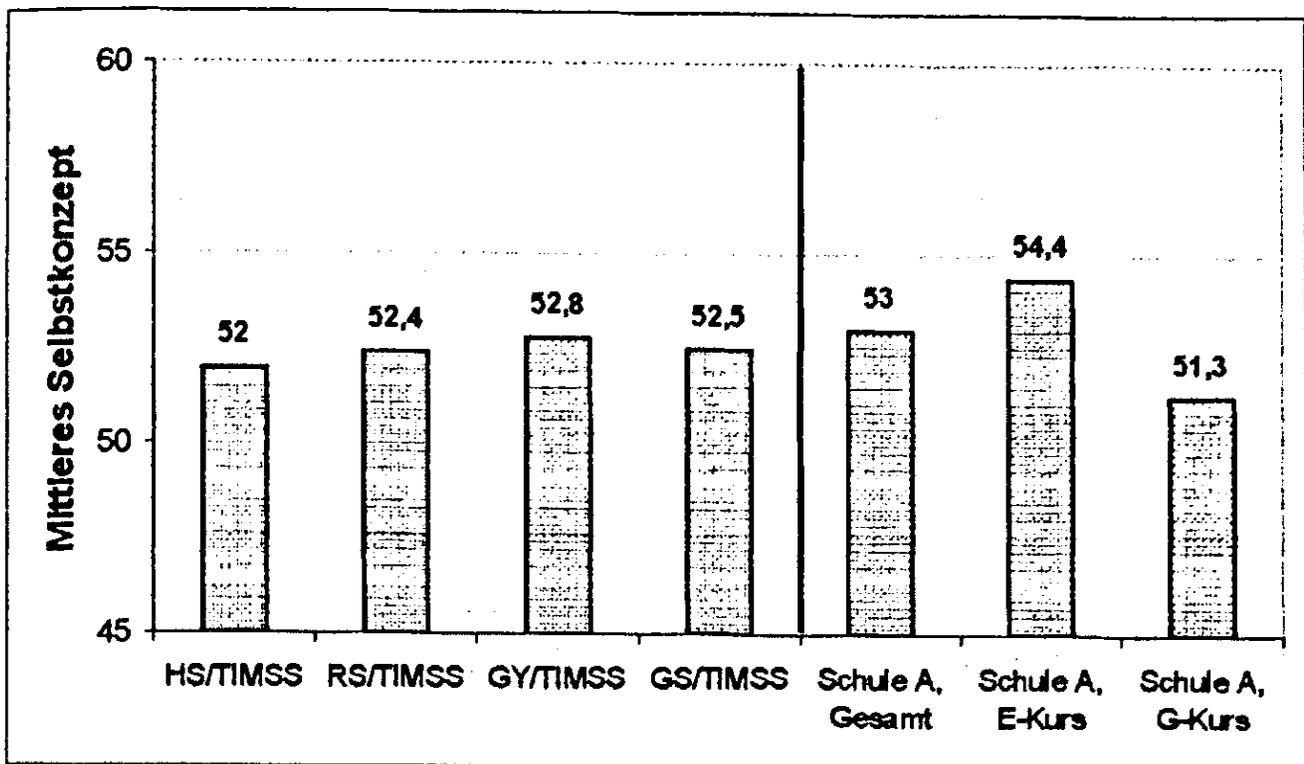
Der Unterschied verschwindet, wenn man den familiären Hintergrund konstant hält, das heißt, durch die Einbeziehung von Hintergrundvariablen sinkt der Unterschied zwischen Schülern mit deutscher bzw. nicht-deutscher Muttersprache von über 30 Punkten auf quasi 0 Punkte. Dieser Befund belegt zum einen die besondere Rolle des Elternhauses für das Leistungsniveau der Schüler. Zum anderen legt er die Interpretation nahe, dass im Rahmen der besonderen Fördermaßnahmen für Migrantenkinder an der Schule A diese trotz etwas schwächerer Leistungen den Erweiterungskursen in Mathematik zugeordnet werden.

*Zusammenfassend lassen sich die Befunde wie folgt interpretieren:* Es besteht im oberen Leistungsbereich eine deutliche Differenz in den Leistungen zwischen Schülern mit deutscher und nicht-deutscher Muttersprache. Berücksichtigt man den unterschiedlichen familiären Hintergrund, werden dadurch die Leistungsunterschiede weitgehend erklärt. Hinter dem relativ hohen Anteil von Migrantenschülern im E-Kurs steht vermutlich die Absicht, benachteiligten Kindern Wege zu besser qualifizierten Schulabschlüssen offen zu halten. Aus einer pädagogischen Perspektive, bei der die optimale Förderung benachteiligter Schüler im Vordergrund steht, kann dies durchaus wünschenswert sein (vgl. Köller et al. 1999).

### *5.3 Selbstkonzept der Begabung in Mathematik*

Schulischer Alltag heißt für Schülerinnen und Schüler auch, dass sie in sämtlichen Fächern regelmäßig Rückmeldungen über ihre eigenen Leistungen erhalten, sei es in Form von Noten nach Klassenarbeiten und in Zeugnissen, oder sei es durch individuelle Rückmeldungen im täglichen Unterrichtsgeschehen. Aufgrund solcher Rückmeldungen und als Folge sozialer Vergleiche mit ihren Mitschülern bilden Kinder und Jugendliche in jedem Fach so genannte Selbstkonzepte eigener Begabungen aus, die Einschätzungen darstellen, für wie begabt sich ein Schüler in dem jeweiligen Fach hält. Zentral für die Ausbildung der Selbstkonzepte ist der Referenzrahmen, in dem Leistungsrückmeldungen im Vergleich zu anderen bewertet werden. Diesen Referenzrahmen bildet üblicherweise die Klasse, d. h. die Höhe des individuellen Selbstkonzepts der Begabung in Mathematik richtet sich daran aus, wie gut man selbst im Vergleich zu den Klassenkameraden ist. Dieser Referenzrahmeneffekt führt in der Regel dazu, dass Schülerinnen und Schüler der Hauptschule ein ähnlich hohes Selbstkonzept der Begabung ausbilden wie Gymnasiasten, obwohl zwischen beiden Schulformen große Leistungsunterschiede bestehen (siehe Abbildung 4).

Abbildung 4: Selbstkonzept der Begabung in Mathematik nach Gruppe; Vergleichsstichproben aus TIMSS; HS: Hauptschule; RS: Realschule; GY: Gymnasium; GS: Gesamtschule; G: G-Kurs; E: E-Kurs



Fachspezifische Selbstkonzepte der Begabung können ihrerseits Leistungsergebnisse beeinflussen. Eine Vielzahl von Studien (siehe z. B. Helmke 1992; Köller, Klemmert, Möller/Baumert 1999) hat mittlerweile belegt, dass Schülerinnen und Schüler mit höheren Selbstkonzepten auch größere Lernerfolge erzielen, indem sie sich mehr anstrengen, ein höheres Interesse haben und aufmerksamer und ausdauernder Aufgaben bearbeiten. In der Untersuchung an den fünf hessischen Gesamtschulen wurden fünf Items zur Erfassung des Selbstkonzepts in Mathematik eingesetzt; Itemformulierungen waren zum Beispiel:

Kein Mensch kann alles. – Für Mathematik habe ich einfach keine Begabung.

Bei manchen Sachen in Mathematik, die ich nicht verstanden habe, weiß ich von vornherein: „Das verstehe ich nie.“

Die Abbildung 4 zeigt die entsprechenden Mittelwerte getrennt nach Gruppen, die Vergleichsstichprobe stammt aus TIMSS. Die Werte sind so transformiert, dass ein Wert von 50 ein mittelhohes Selbstkonzept der Begabung anzeigt. Die Standardabweichung ist 10. Deutlich erkennbar sind in den vier Schulformen der TIMS-Studie die Referenzgruppeneffekte. Die jeweiligen Klassenkameraden stellen die Bezugsgruppe dar, an der sich die Selbsteinschätzung ausrichtet, was dazu führt, dass keine Schulformunterschiede beobachtbar sind.

Im Vergleich zu TIMSS ist der Mittelwert der Schule A ( $M = 53$ ) unauffällig. Bemerkenswert sind die relativ geringen Unterschiede zwischen E- und G-Kurs-Schülern an der Schule A. Die Forschungsarbeiten zu den Effekten von Leistungsdifferenzierungen innerhalb von Schulen zeigen eher, dass die Differenzierung gerade im unteren Bereich erheblichen Kosten im Begabungselbstkonzept hat (im Überblick Wagner, 1999). Üblicherweise werden größere Unterschiede gefunden. Offenbar gelingt es der untersuchten Schule, leistungsschwächeren Schülern Kompetenzerlebnisse zu vermitteln, die sol-

che unerwünschte Effekte verhindern. Bezogen auf die TIMSS-Stichproben ergeben sich keine signifikanten Unterschiede zwischen beiden Kursniveaus an Schule A und den vier Schulformen aus TIMSS.

Zusammenfassend lässt sich konstatieren, dass es der untersuchten Schule neben der erfolgreichen Vermittlung mathematischen Fachwissens offenbar gelingt, möglichen Problempunkten, die sich aufgrund der unterschiedlichen Herkunft und Leistung ihrer Schülerschaft ergeben können, erfolgreich zu begegnen.

## 6. Die Schule B

Die Schule B wurde 1988 eingerichtet und ist Versuchsschule des Landes Hessen. Sie ist eine Ganztagschule mit flexiblen Unterrichtszeiten von 7.45 bis 16.00 Uhr. Die Jahrgangsstärke beträgt 30 bis 40 Kinder bzw. Jugendliche. *Der Anteil ausländischer Mitschülerinnen und -schüler ist sehr gering.*

An der Schule B wird in *jahrgangsgemischten Kursen* unterrichtet, wobei es insgesamt vier Stufen gibt, bei denen die ersten drei jeweils drei Jahrgänge umfassen (Jg. 0 bis 2, Jg. 3 bis 5, Jg. 6 bis 8) und die letzte Stufe zwei Jahrgänge (Jg. 9 bis 10). In allen Fächern außer Mathematik, Englisch und Französisch, die jahrgangsspezifisch unterrichtet werden, findet der Unterricht für die von uns untersuchten Schülerinnen und Schüler der 8. Jahrgangsstufe gemeinsam mit Schülern aus der 6. und 7. Jahrgangsstufe statt. Somit führt die in der Schule B praktizierte Jahrgangsmischung dazu, dass in vielen Fächern vergleichsweise leistungsheterogene Gruppen zusammen unterrichtet werden. Diese Zusammensetzung soll – so das Ziel der Schule B – u. a. die soziale Kompetenz der Schüler fördern. Die Altersheterogenität und die damit zusammenhängenden Unterschiede im Leistungsvermögen können auch aktiv in den Unterricht eingebracht werden, indem die Schüler einander helfen. Somit kann eine selbstverständliche Kultur des Helfens und der Achtung voreinander etabliert werden, wobei der jahrgangsübergreifende Unterricht verhindert, dass allzu feste Rollen und Rollenerwartungen entstehen.

*Im jahrgangsspezifischen Mathematikunterricht der 8. Klasse* wird intern differenziert. Es werden keine Klassenarbeiten geschrieben. Stattdessen erwerben die Schülerinnen und Schüler in jedem Stoffgebiet so genannte Grund- und Erweiterungsdiplome, deren Schwerpunkt auf der Feststellung des Könnens, nicht des Nichtkönnens, liegt. Der pädagogische Ansatz der Schule baut sehr stark auf Konzepten der Reformpädagogik auf. Konsequenter Weise wird deshalb auch auf die Erteilung von Noten und Ziffernzeugnisse bis zum Ende des 8. Jahrgangs verzichtet.

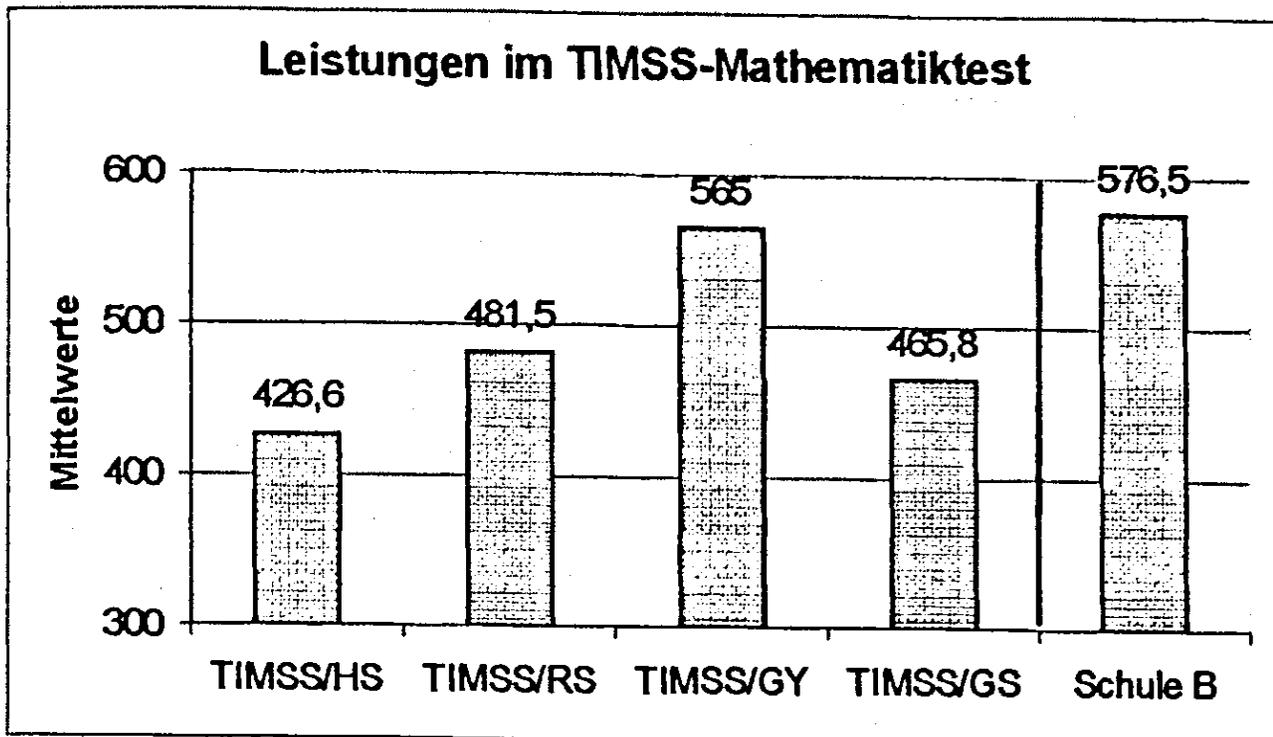
*Die soziale Zusammensetzung der Schule B* verdient eine besondere Erwähnung. Zwar wurde diese Schule als Gesamtschule konzipiert, sie zieht jedoch aufgrund ihres pädagogischen Profils insbesondere Kinder aus Familien mit großer Bildungsnähe und einem hohen sozio-ökonomischen Status an, wie man ihn üblicherweise nur am Gymnasium antrifft.

Im Folgenden berichten wir in aller Knappheit die Ergebnisse in dem TIMSS-Mathematiktest, bevor wir ausführlicher darauf eingehen, wie der Mathematikunterricht an der Schule B von den Schülerinnen und Schülern beurteilt wird.

## 6.1. Die Ergebnisse in den TIMSS-Leistungstests

Die Abbildung 5 zeigt die Ergebnisse der Schule B im Mathematiktest.

Abbildung 5: Mathematikleistungen der Schule B im Vergleich zu den in TIMSS untersuchten 8. Jahrgängen an verschiedenen Schulformen; HS: Hauptschule; RS: Realschule; GY: Gymnasium; GS: Gesamtschule



Insgesamt liegt der Mittelwert der untersuchten Schüler aus der Schule B knapp über dem von Gymnasien in der TIMS-Untersuchung. Dies weist darauf hin, dass die pädagogische Schwerpunktsetzung an der Schule B nicht zu Lasten der Leistungsentwicklung in Mathematik geht. Das relativ hohe Leistungs-niveau ist auch vor dem Hintergrund, dass in dieser Schule keine Klassenarbeiten geschrieben werden, bemerkenswert.

## 6.2. Wie beurteilen Schülerinnen und Schüler den Unterricht?

Inwieweit ist das gute Ergebnis der Schule B im TIMSS-Mathematiktest auf einen besonderen Mathematik-Unterricht zurückzuführen? Lassen sich womöglich besondere Stärken oder Schwächen identifizieren?

Die Suche nach einem „guten“ Unterricht gehörte und gehört zu den zentralen Themen in der Lehr-Lernforschung. Mit zunehmendem Erkenntnisgewinn hat sich dabei gleichzeitig die Hoffnung zerstoben, dass einfache Ursache-Wirkungs-Modelle mit Anspruch auf Allgemeingültigkeit die Komplexität des Unterrichtsgeschehens adäquat erklären können. Vielmehr wurde betont, dass (1) Lehrer sowohl guten als auch schlechten Unterricht auf sehr verschiedene Weise halten können und dass (2) gleich erscheinende Verhaltensweisen des Lehrers unter verschiedenen Bedingungskonstellationen und für verschiedene Schülertypen unterschiedliche Wirkungen haben können (Helmke/Weinert 1997). Zudem zeigte sich eine wechselseitige Kompensierbarkeit von Unterrichtsaspekten: Schwächen in einigen Bereichen können durch Stärken in anderen Bereichen ausgeglichen werden.

In der vorliegenden Untersuchung wurde auf einige *Variablen* zurückgegriffen, die sich in der Vergangenheit als bedeutsam für Unterrichtserfolg erwiesen haben. Wichtig im Hinblick auf die Interpretation der Befunde ist, dass lediglich von Schülerinnen und Schülern Beschreibungen des Unterrichts vorliegen. Beobachtungsdaten über Unterrichtsmerkmale können auch anderen Quellen entstammen – so der Einschätzung von Lehrkräften oder Fremdbeobachtungen beispielsweise durch Kolleginnen und Kollegen. Die TIMSS-Video-Studie (vgl. Baumert, Lehmann u.a. 1997) hat weiterhin gezeigt, wie hilfreich gerade das Videographieren von Schulstunden für Fremdbeobachtungen sein kann. Schülerangaben, die sich insgesamt als nützliche Beschreibungen des Unterrichtsgeschehens erwiesen haben (Gruehn 2000), unterliegen bestimmten Beschränkungen. So bildet sich natürlich im individuellen Schülerurteil auch dessen Persönlichkeit ab, man spricht üblicherweise von der psychologischen Umwelt eines Schülers. Erst ein Aggregieren auf Klassenebene bereinigt die Schülerangaben um eventuelle Idiosynkrasien, und es entsteht ein Bild der von allen Schülerinnen und Schülern geteilten Umwelt.

Für den vorliegenden Bericht wurden diejenigen Unterrichtsmerkmale ausgewählt, bei denen sich besonders große Unterschiede zwischen den Vergleichsstichproben und der Schule B gezeigt haben. Dabei handelt es sich um die *Klarheit und Strukturiertheit des Unterrichts*, die vorhandene *Regelklarheit*, das Ausmaß an *Unterrichtsstörungen* und um die so genannte *Individuelle Bezugsnormorientierung* der Lehrkräfte. Itembeispiele sind:

**Klarheit und Strukturiertheit des Unterrichts**

„Unser Mathematiklehrer erklärt besonders an schwierigen Stellen ganz langsam und sorgfältig.“

**Regelklarheit hinsichtlich unerwünschten Verhaltens**

„Im Mathematikunterricht hat unser Lehrer klargemacht, was passiert, wenn man Regeln verletzt.“

**Unterrichtsstörungen**

„Im Mathematikunterricht wird der Unterricht oft sehr gestört.“

**Individuelle Bezugsnormorientierung der Lehrkraft \***

„Wenn sich ein schwacher Schüler verbessert, ist das für unseren Lehrer eine gute Leistung, auch wenn der Schüler immer noch unter dem Klassendurchschnitt liegt.“

Während die ersten drei Aspekte keiner Erklärung bedürfen, sind einige Worte zur *individuellen Bezugsnormorientierung* angebracht. Aus motivationspsychologischer Perspektive ist wiederholt argumentiert worden, dass es im Hinblick auf die Optimierung der Lernmotivation oder des Selbstkonzepts der eigenen Begabung wünschenswert ist, dass Schüler ihre aktuellen mit ihren früheren Kenntnissen vergleichen (Rheinberg 1998; Rheinberg/Krug 1999), da bei dieser Vergleichsform sofort Wissensfortschritte deutlich werden. Dieser Vergleich mit früheren Leistungen wird in der Lehr-Lernforschung üblicherweise „individuelle Bezugsnorm“ genannt (z.B. Rheinberg 1998) und der „sozialen Bezugsnorm“ (als Vergleich mit Mitschülern) und der „kriterialen Bezugsnorm“ (als einer inhaltlich begründeten Mindestleistung für eine bestimmte Note) gegenübergestellt (vgl. hierzu auch Klauer 1989). Tatsächlich zeigen verschiedene Studien, dass die Bezugsnormorientierung der Lehrer einen bedeutenden Einfluss auf die Motivation der Schülerinnen und Schüler hat und eine individuelle Bezugsnormorientierung des Lehrers sogar Leistungen verbessern kann (Krampen 1985; Mischo/Rheinberg 1995; Rheinberg/Krug 1999).

Tabelle 2: Aspekte der Instruktion im Mathematikunterricht. Vergleichsmittelwerte aus der BIJU-Befragung in Nordrhein-Westfalen sowie aus TIMSS

	Regelklarheit	Klarheit und Strukturiertheit	Unterrichtsstörungen	Individuelle Bezugsnormorientierung der Lehrkraft
Hauptschule/BIJU	51.3	52.3	50.4	51.8
Realschule/BIJU	52.4	53.0	50.4	52.5
Gymnasium/BIJU	51.8	53.9	48.6	53.7
Gesamtschule/BIJU	52.2	53.0	50.1	53.2
Hauptschule/TIMSS	53.1	54.6	50.1	53.9
Realschule/TIMSS	52.8	52.6	50.7	51.3
Gymnasium/TIMSS	52.0	51.4	49.1	52.1
Gesamtschule/TIMSS	52.5	52.9	51.5	52.1
Schule B	57.8	61.8	44.8	60.8

Anmerkungen: Vergleichsstichproben aus TIMSS und BIJU

Die Tabelle 2 zeigt die Befunde (Mittelwerte) zu den vier Skalen in der Schule B im Vergleich zu den Stichproben aus der TIMS- und BIJU-Studie. Die Daten sind wiederum so transformiert, dass eine Standardabweichung von 10 in der Vergleichsstichprobe vorliegt. In der Schule B liegt die Streuung etwas niedriger.

*Die Einschätzungen in der Schule B fallen für alle hier berücksichtigten Indikatoren signifikant günstiger aus als die der Vergleichsgruppen. Paarweise Mittelwertvergleiche (t-Tests) der Schule B mit allen anderen Subgruppen werden ohne Ausnahme signifikant ( $p < .05$ ).*

An der Schule B wird der Mathematikunterricht insgesamt als klarer und strukturierter wahrgenommen, es werden weniger Störungen berichtet und Konsequenzen von regelwidrigem Verhalten werden als klar definiert betrachtet. *Außerdem weisen die Angaben der Schülerinnen und Schüler aus der Schule B darauf hin, dass in ihrem Mathematikunterricht eine sehr starke Orientierung an einer individuellen Bezugsnorm stattfindet.* Dieser Befund mag v.a. auf eine Besonderheit zurückführbar sein: An der Schule B wird bis zur 8. Jahrgangsstufe auf den Einsatz von Noten und Zensuren verzichtet. Dies reduziert möglicherweise die Auftretenshäufigkeit interindividueller (sozialer) Vergleiche, die eher negative Konsequenzen für individuelle Selbstkonzepte und Interessen haben (Köller et al., 1998).

Die erhobenen Schülerurteile deuten insgesamt darauf hin, dass die guten Mathematikleistungen auch ein Produkt des Unterrichts sind und nicht lediglich ein Ergebnis des günstigen familiären Hintergrunds darstellen.

## 7. Fazit

Die Untersuchung mit TIMSS-Instrumenten an den fünf hessischen Gesamtschulen hat mehrere wichtige Erkenntnisse im Hinblick auf Schulentwicklungsmaßnahmen gebracht. Zum einen haben die Schulen wichtige Hinweise darüber erhalten, wie sie hinsichtlich einer großen Anzahl von Aspekten ihre Schülerinnen und Schüler im Vergleich mit anderen Schulen stehen. Hier hat sich die Einsetzbarkeit von TIMSS jenseits einer Momentaufnahme bundesdeutscher Fachleistungen gezeigt. *Zum anderen wurde empirisch die von uns*

*vertretene Ansicht gestützt, dass es nicht die Schulform (dreigliedriges Schulsystem vs. Gesamtschule) per se ist, die bestimmt, wie die Lernleistung der Schüler ausfällt, sondern dass es innerhalb verschiedener Schularten die Möglichkeit gibt, mit überzeugenden pädagogischen Konzepten eine erfolgreiche Schule zu gestalten.*

Was geschah und geschieht nun mit den Daten der einzelnen Schulen? Innerhalb der teilnehmenden Schulen und zusammen mit Vertretern des Max-Planck-Instituts wurden die Untersuchungsergebnisse ausführlich diskutiert. An den Schulen wird derzeit überlegt, wie die Stärken der Schulen auch in der Öffentlichkeit noch besser zur Geltung gebracht werden können und wie mit den Bereichen umgegangen werden kann, bei denen die Untersuchung noch Optimierungsbedarf offenbarte und die je nach Schule ganz unterschiedlich waren. Dabei tauschen sich die Schulen auch intensiv miteinander über ihre jeweiligen Wege aus.

## Literatur

- Arbeitsgruppe Bildungsbericht am Max-Planck-Institut für Bildungsforschung (Hg.) 1994: Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick. Reinbek: Rowohlt
- Baumert, Jürgen; Wilfried Bos und Rainer H. Lehmann (Hg.) 2000a.: Dritte Internationale Mathematik- und Naturwissenschaftsstudie: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen: Leske + Budrich
- Baumert, Jürgen; Wilfried Bos, und Rainer H. Lehmann (Hg.) 2000b: Dritte Internationale Mathematik- und Naturwissenschaftsstudie: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske + Budrich
- Baumert, Jürgen und Olaf Köller 2000: Unterrichtsgestaltung, verständnisvolles Lernen und multiple Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe. In: Jürgen Baumert, Wilfried Bos/Rainer H. Lehmann (Hg.): Dritte Internationale Mathematik- und Naturwissenschaftsstudie: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen Kompetenzen am Ende der gymnasialen Oberstufe Opladen: Leske + Budrich, S. 271-316
- Baumert, Jürgen, Rainer H. Lehmann u.a. 1997: TIMSS: Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen: Leske + Budrich
- Baumert, Jürgen u.a. 1996: Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). In: Klaus-Peter Treumann, Günther Neubauer, Renate Möller und Jürgen Abel (Hg.): Methoden und Anwendungen empirisch pädagogischer Forschung: Münster: Waxmann, S. 170-180
- Baumert, Jürgen und Olaf Köller 1998a: Nationale und internationale Schulleistungstudien: Was können sie leisten, wo sind ihre Grenzen? In: Pädagogik. 50, 1998, 6, S. 12-18
- Baumert, Jürgen; Olaf Köller und Kai Schnabel 2000: Schulformen als differentielle Entwicklungsmilieus – eine ungehörige Fragestellung? In Gewerkschaft Erziehung und Wissenschaft GEW (Hg.): Messung sozialer Motivation. Eine Kontroverse. Frankfurt/Main: Bildungs- und Förderungswerk der GEW, S. 28-68
- Bender-Szymanski, Dorothea und Hermann-Günter Hesse 1987: Migrantenforschung. Köln: Böhlau
- Gruehn, Sabine 1995: Vereinbarkeit kognitiver und nicht-kognitiver Ziele im Unterricht. In: Zeitschrift für Pädagogik. 41, 1995, S. 531-553

- Helmke, Andreas 1992: Selbstvertrauen und schulische Leistungen. Göttingen: Hogrefe
- Helmke, Andreas und Friedrich Wilhelm Schrader 1990: Zur Kompatibilität kognitiver, affektiver und motivationaler Zielkriterien des Schulunterrichts – Clusteranalytische Studien. In: Monika Knopf und Wolfgang Schneider (Hg.): Entwicklung: Festschrift zum 60. Geburtstag von Franz Emanuel Weinert. Göttingen: Hogrefe, S. 180-200
- Köller, Olaf 1998: Zielorientierungen und schulisches Lernen. Münster: Waxmann
- Köller, Olaf; Jürgen Baumert, und Kai Schnabel 1999: Wege zur Hochschulreife: Offenheit des Systems und Sicherung vergleichbarer Standards. Analysen am Beispiel der Mathematikleistungen von Oberstufenschülern an integrierten Gesamtschulen und Gymnasien in Nordrhein-Westfalen. In: Zeitschrift für Erziehungswissenschaft. 2, 1999, S. 370-405
- Köller, Olaf; Hella Klemmert, Jens Möller und Jürgen Baumert 1999: Leistungsbeurteilungen und Fähigkeitsselbstkonzepte: Eine längsschnittliche Überprüfung des Modells des Internal/External Frame of Reference. In: Zeitschrift für Pädagogische Psychologie. 13, 1999, S. 128-134
- Krampen, Günter 1985: Differentielle Effekte von Lehrerkommentaren zu Noten bei Schülern. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie. 17, 1985, S. 99-123
- Lehmann, Rainer H. und Rainer Peek 1997: Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Hamburger Schulbehörde
- Mischo, C. und Falko Rheinberg 1995: Erziehungsziele von Lehrern und individuelle Bezugsnormender Leistungsbewertung. In: Zeitschrift für Pädagogische Psychologie. 9, 1995, S. 139-151
- Nauck, Bernhard und Heike Diefenbach 1997: Bildungsbeteiligung von Kindern aus Familien ausländischer Herkunft. Eine methodenkritische Diskussion des Forschungsstandes und eine empirische Bestandsaufnahme. In Folker Schmidt (Hg.): Methodische Probleme der empirischen Erziehungswissenschaft. Hohengehren: Schneider, S. 289-307
- Rheinberg, Falko 1998: Bezugsnormorientierung. In: Detlef H. Rost (Hg.): Handwörterbuch Pädagogische Psychologie Weinheim: Beltz/PVU, S. 39-43
- Rheinberg, Falko und Siegbert Krug (Hg.) 1999: Motivationsförderung im Schulalltag (2. Auflage). Göttingen: Hogrefe
- Schnabel, Kai 1998: Prüfungsangst und Lernen. Münster: Waxmann
- Schwippert, Knut 1999: Einfluss schulischer und unterrichtlicher Faktoren auf das Leseverständnis Vierzehnjähriger unter Berücksichtigung ihrer sozialen Lage. Unveröffentlichte Dissertation an der Humboldt-Universität zu Berlin
- Wagner, Jürgen W. L. 1999: Soziale Vergleiche und Selbsteinschätzungen. Münster: Waxmann

*Olaf Köller*, geb. 1963, Dr. phil.; Diplom in Psychologie 1991 an der Universität Kiel, ab 1992 wissenschaftlicher Mitarbeiter im Institut für die Pädagogik der Naturwissenschaften in Kiel; seit 1996 wissenschaftlicher Mitarbeiter am Max-Planck-Institut für Bildungsforschung; Promotion 1997 in Psychologie;

Anschrift: Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, D-14195 Berlin

E-Mail: koeller@mpib-berlin.mpg.de

*Ulrich Trautwein*, geb. 1972; Diplom in Psychologie 1999 an der Universität Göttingen, seit 1999 wissenschaftlicher Mitarbeiter am Max-Planck-Institut für Bildungsforschung;

Anschrift: Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, D-14195 Berlin

E-Mail: trautwein@mpib-berlin.mpg.de