Studien zur
International und Interkulturell
Vergleichenden Erziehungswissenschaft

herausgegeben von

Band 3

Waxmann Münster / New York
München / Berlin

Hans Döbert, Eckhard Klieme, Wendelin Sroka (eds.)

Conditions of School Performance in Seven Countries

A Quest for Understanding the
International Variation of PISA Results

Waxmann Münster / New York
München / Berlin

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the
Deutsche Nationalbibliografie; detailled bibliographic data
is available in the Internet at http://dnb.ddb.de.

Findings of the Research Project "Comparison of Education Systems in Selected Countries – Understanding the International Variation of PISA Results", Part 1, conducted by a Working Group led by the German Institute for International Educational Research (DIPF), initiated and financed by the German Federal Ministry of Education and Research (BMBF).

Studien zur International und Interkulturell Vergleichenden Erziehungswissenschaft, Band 3

ISSN 1612-2003
ISBN 3-8309-1373-7

Postfach 8603, D-48046 Münster

http://www.waxmann.com
E-Mail: info@waxmann.com

Umschlag: Pleßmann Kommunikationsdesign, Ascheberg
Druck: Zeitdruck GmbH, Münster
Gedruckt auf alterungsbeständigem Papier, DIN 6738

- measures for further developing and securing quality of tuition and school on the basis of binding standards (such as a central *Abitur*) as well as results-oriented evaluation, the individual school having increased responsibility;
- measures aimed at opening up the school to its environment and the development of cooperation with neighbours, sports associations, churches, artists, independent maintaining bodies promoting activities with and by young people, the keywords being partnership, cooperation and dialogue for the nascent educational networks;
- measures aimed at improving the professionalism of the teaching staff, particularly with regard to diagnostic and methodical competence as a component of systematic school development;
- measures aimed at strengthening the change of paradigm in teacher in-service training, from the qualification of the single teacher to their qualification in teams and in the school, as well as to a close-to-reality teacher pre-service training being an important momentum of the necessary educational offensive.

As a whole, a series of recent developments, which could take the form of a long-term and effective reform and innovation strategy, are emerging. Among them are, in particular: the tendency fed mainly from the new Länder to flexible institutional solutions within the school structures, which have been in existence up until now (for example double-tracked secondary level), the step-by-step enforcement of a new control of the school system by limiting central requirements and extending the sphere of responsibility of the individual school, the emphasis on quality of school and instruction as the core of school development, the change of paradigm in the control of the school system towards an output-oriented control as well as the introduction of a system of system monitoring and of quality assurance.

2.7 Assessments/evaluations (*Isabell van Ackeren, Klaus Klemm*)

The following chapter outlines the methods of external examination practised in Germany: central examinations, continuous assessment and System Monitoring. To understand these methods, it is necessary to bear in mind the peculiarities of the German school system, which reflect the political system of cooperative federalism (see section 2.1).

2.7.1 External evaluation in Germany: Central examinations, continuous assessment and System Monitoring

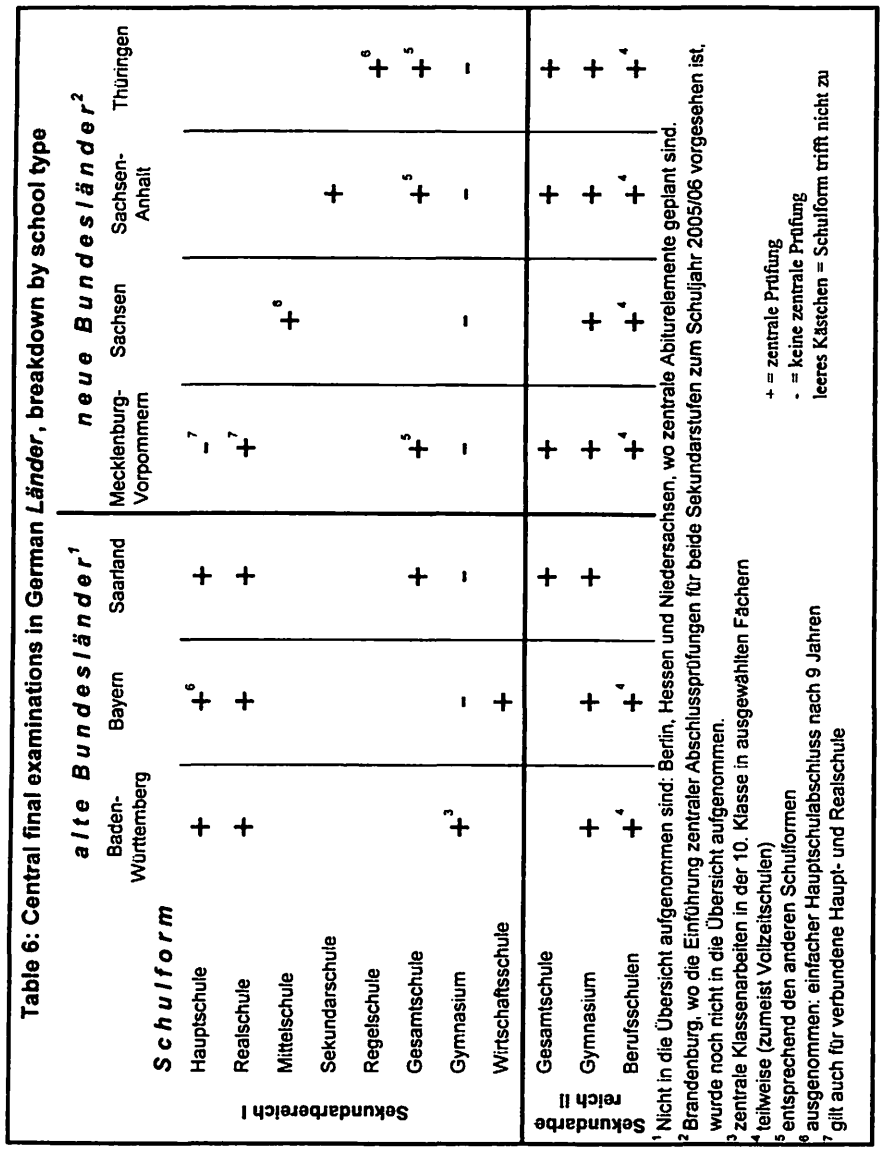
The differences in school systems between the Länder, the evident inadequacies of the German school system, and the increasing autonomy of individual schools throughout Germany have all contributed to the increasing emphasis placed on external evaluation since the mid 90s: central final examinations for all school-leavers, standardised tests during the school career and international Large Scale Assessments.

Two contrasting models: centralised and decentralised final examinations

Almost half of the Länder use a traditional model of external assessment in the form of central final examinations. In seven Länder, there are not only detailed syllabuses, course-book tests, detailed examination prescriptions, teacher training in assessment, requirements for putting students into the next class, and the school inspection, as control measures. Even the KMK resolutions for "standards for the middle school final examination" and for "standardised examination requirements" are not deemed sufficient by these Länder to ensure quality of school education and the equivalence of certificates. They therefore also have central examinations (see Klemm, 1998). These central examinations are taken in Baden-Württemberg, Bavaria, Saarland and in all the Länder in the former East Germany, apart from Brandenburg. They differ from Land to Land and exist for the end of the lower secondary level as well as for the Abitur (see table 6).

The tradition of central examinations in the three old Länder goes back to the influence of the occupying allied forces at the end of World War II, particularly of France. The French missionary attitude to education was noticeable in Saarland and what was later called Baden-Württemberg. In Rhineland-Palatinate, a central Abitur was introduced during the French occupation, but abolished a few years later. The central final examinations in Bavaria reflect the traditionally more centralised structure of the Land as a whole. The new Länder created at reunification in 1990 also had a long tradition of central examinations in the centrally managed Unitary Education System in the GDR. This tradition was carried on in all new Länder, except Brandenburg. But even there, the tendency towards increasing independence of individual schools, together with the need to test and compare school results in connection with the debate about quality have led to the decision to introduce central examinations for the school year 2005/2006 at the latest. Reacting to the PISA results, Lower Saxony also announced that it would introduce standardised examinations allowing for some freedom for individual schools.

Table 6: Central final examinations in German *Länder*, breakdown by school type



[1] Nicht in die Übersicht aufgenommen sind: Berlin, Hessen und Niedersachsen, wo zentrale Abiturelemente geplant sind.
[2] Brandenburg, wo die Einführung zentraler Abschlussprüfungen für beide Sekundarstufen zum Schuljahr 2005/06 vorgesehen ist, wurde noch nicht in die Übersicht aufgenommen.
[3] zentrale Klassenarbeiten in der 10. Klasse in ausgewählten Fächern
[4] teilweise (zumeist Vollzeitschulen)
[5] entsprechend den anderen Schulformen
[6] ausgenommen: einfacher Hauptschulabschluss nach 9 Jahren
[7] gilt auch für verbundene Haupt- und Realschule

In Hesse, landwide standardised examination tasks are planned for parts of the written Abitur. For these tasks, the conditions still have to be developed. Thus Hesse is also heading towards an Abitur with some centralised elements. This is coupled with further restrictions to the 'gymnasiale Oberstufe' (upper level of the grammar school), aimed at a better control over the outcome: *ie,* greater emphasis on core subjects, a higher number of examination subjects and a more restricted choice of subjects for Abitur. In Berlin, a central Abitur is planned for the year 2006. In North Rhine-Westphalia, Rhineland-Palatinate, Schleswig-Holstein, Bremen and Hamburg, a central Abitur is not currently being discussed.

This overview of the use of central examinations and the efforts to standardise the final examination shows that central final examinations have a secure place in Germany, which will be strengthened even more in future. The discussions have been spurred on by the results of the large-scale international achievement studies of the last few years. In CDU-governed and, now, also in SPD-governed Länder, the equality of decentralised and centralised examinations has been called into question in favour of centralised examinations. There has, however, not yet been any academic debate about the structure and effects of centralised versus decentralised examinations.

There are two notable deviations between the Länder concerning the examination set-up. First, not all Länder submit precise topic areas for the central examinations at the beginning of the last two years of the upper level of the grammar school. Bavaria, for example, chooses the examination topics from the syllabus without further specification. Secondly, the marking procedure differs from land to land: Bavaria and Saarland use internal examiners, while other Länder, such as Baden-Württemberg and Saxony, use external examiners and ensure student anonymity. In spite of these Länder-specific variations, the examination procedures are comparable (see Klemm, 1998). There have not yet been any publications of individual schools' results, apart from one attempt in Saxony, which has since been abandoned. It must be borne in mind that only 20% of the full qualification achieved by each student at the end of the 'gymnasiale Oberstufe' (upper level of the grammar school) is directly influenced by centralised examinations. By far the greatest part of the total marks is made up from the internal results achieved during the upper-level career. Even so, the normative effect of examinations, with reference to the assessment requirements in all subjects, is a topic of public and academic debate. The critics of centralised examinations repeatedly claim that centralised examinations have a normative effect: the announcement of examination topics affect instructions. They claim that centralised examinations encourage reproductive learning and suppress creativity, which cannot be tested adequately since the actual classroom work cannot be fitted into examination structures. For some subjects in particular, it is difficult to formulate uniform assessment criteria. Finally, they argue that this form of examination does not take into account the different conditions of achievement within schools, since it is not possible to check whether a particular topic has been thoroughly discussed in class. Conversely, the advocates of centralised assessment argue that decentralised examinations could be systematically below the required level, could fail to cover the topic areas sufficiently and that examinations could be too intensively and therefore unfairly prepared for. Moreover, the equivalence of the assessment criteria could be undermined.

Looking at these arguments in the light of the TIMSS and PISA results, neither the international comparisons nor the results within Germany give convincing evidence for an explicit general connection between types of examination and achievement (Baumert, 2001; OECD 2001). In the PISA study, the group of the highest achieving Federal Länder at grammar school level includes both Länder with a centralised Abitur (*eg,* Bavaria and Baden-Württemberg) and Länder with decentralised examinations (*eg,* Schleswig-Holstein, Rhineland-Palatinate). The comparison also shows that there are Länder with central examinations which do less well than average (*eg,* Saxony-Anhalt), which does not indicate that central examinations secure a minimum standard in Germany.

Nonetheless, centralised final examinations do enable comparisons to be made between the results of different schools and discrepancies to be shown, which provides a starting point for

political educational reform. The examinations -whether centralised or decentralised- have, up until now, primarily been used to provide school-leavers with a recognised qualification. There has not yet been any systematic feedback or use of previous results in schools, although the examination results in all Länder are statistically drawn up and evaluated. Access to the data -except via short press statements- has, up until now, been reduced or prohibited by the ministries, with reference to data protection laws.

A new option: standardised tests during the course of the school career

The international comparative studies of the last few years have shown that results fall short of a minimum standard, results are not equivalent in the whole nation and results are not evaluated adequately. Thus, there are inadequacies both in the German education system and in the individual schools. Appropriate educational policies are being worked on to redeem this situation. Alongside the intensified discussion of centralised final examinations, control methods are being developed, which will directly affect daily school life and so support the individual school. These are intended, in turn, to help achieve equivalence on a Länder, a national and an international level. These measures are known as parallel and comparative work.

- Parallel work' involves comparing students of parallel classes within one school;
- 'Comparative work' or 'orientation work' involves comparing different schools within a defined region.

Parallel work involves joint work being done by several parallel classes in core subjects (particularly German, English and Maths) in various school years at primary and secondary level (this varies from Land to Land). The tasks are conceived jointly by the teachers of the parallel classes and accomplished at particular time intervals. The ministry provides example tasks as guidelines for difficulty level and assessment criteria. The discussion of parallel work, designed to enable comparison of individual class results and assessments, reflects a basic debate: the effort to achieve objectivity and equivalence, especially given the tendency towards public accountability and transparency. This has been started in many Länder through assessment within the school.

Parallel work is being established as a norm-orientated motor for quality evaluation and development. It is well known that staff have difficulty establishing equivalent results and assessment across the classes. The function of parallel work is therefore (*eg,* Forthaus/Röhr 2001):

- to establish and fix what is to be learned;
- to encourage discussion between teachers of parallel classes about teaching methods and their effectiveness;
- to compare not only student achievement, but also staff evaluation methods and
- to encourage discussion within the school, but also between neighbouring schools, for example about didactics.

Example tasks (also known as 'sample tasks' or 'task pools') have been devised for parallel work; example tasks for comparative work will follow. Their function is:

- to clarify requirements set out in guidelines and syllabuses and
- to aid classroom diagnosis other than assessment.

Comparative work has similar aims to parallel work. It has come to the centre of interest in the current national debate following the PISA study. The Standing Conference of Ministers of Education and Cultural Affairs of the Länder of the FRG (KMK) has spoken in favour of implementing "orientation and comparative work", cross-nationally and within individual Länder, to control the adherence to standards still to be devised to take place during the course of the school career, similar to those which exist for the final examinations (Resolutions of May and October 2002; see www.kmk.org). So comparative work involves the comparison of several or all schools in the Land. Comparative work and parallel work are carried out during a school career, and the diagnosis made is not a basis for negative selection mechanisms, but rather, a basis for better student support. As with parallel work, the comparisons are intended to be implemented in the whole Land, both in primary and secondary schools. There are no plans as yet to gather background information about the individual schools to help evaluate discrepancies in achievement. It is also unclear whether the results will be centrally gathered and perhaps published and whether they will contribute to the students' grades. Here, further developments and decisions are awaited.

There has been very little research in Germany so far regarding parallel and comparative work. The available literature discusses the organisational structures still largely at the planning stage (*eg,* Kempfert/Rolff, 2000). It links these to the educational policies behind them and addresses the first results to have appeared. The discussions are predominantly normative and do not have an academic basis. Thus, little is known about the effects of these control measures. Parallel and comparative tasks are a sign of the change from an emphasis on context factors to an intensified and systematised emphasis on process and effect of schooling, based on empirical comparative research (see Helmke, 2000 on this paradigm shift). This is also apparent in the introduction of national comparative studies in the Länder, which complement the System Monitoring of international large-scale assessments. This is the subject of the next section.

System Monitoring: International and national Large Scale Assessments

A type of comparative assessment of chosen year groups, either the whole population or a representative sample, carried out irregularly in the course of their educational path, came to prominence in Germany in the 1990s. These studies are intended to provide general knowledge about educational processes and it is also hoped that they will be useful for the individual school. The second point particularly applies to the studies carried out within Germany.

This type of external evaluation has three forms:

- international-scale studies with German participation;
- cross-national studies within Germany;
- comparative assessment within one Land.

The two tabular overviews shown here refer to these three forms. Table 7 gives a chronological overview of Germany's participation in international comparative studies since the 1960s. Table 8 refers to cross-national and single *Länder* studies within Germany. The following section will deal firstly with the important international Large Scale Assessments with German participation and their effects for the development of the German education system. Secondly, the comparative assessments introduced within Germany will be outlined with examples (for further information, see van Ackeren/Klemm, 2000, Bos/Schwippert, 2002, Weinert, 2001).

Participation in the TIMS study marked an important turning point in the development of the German education system. Germany's results, in all types of school, all subject areas and all Länder, were so mediocre in comparison with other countries, that they incited a debate, still ongoing and now greatly intensified after PISA, about the competence of the German education system. One consequence of this debate was that the hitherto widespread scepticism in Germany regarding such large-scale comparative studies was quickly abandoned and even practically reversed. Although there had already been an international comparative study with German participation- the "International Study of Reading Literacy" in the 90s- the debate about TIMSS and its effects on Germany shows that TIMSS marked a turning-point in Germany's involvement with Large Scale Assessments. It was only after the TIMS studies had awakened interest in the achievement of German students that belated attention was paid to the International Study of Reading Literacy. This delay in awareness is also surely connected with the fact that public interest in the results of comparative studies was inseparable from the debate about Germany's economic situation. In this context, mathematical and scientific literacy evidently seemed to the public to be more important than reading literacy and reading habits. This has markedly changed since the PISA study has shown the fundamental importance of reading literacy for other subject areas. The trend towards mediocrity which was confirmed by the international reading study and by TIMSS, has been re-confirmed by PISA. Indeed, Germany's

place on the ranking lists shows that the downward trend is continuing. Large-scale, international empirically based assessment has only been the subject of intensive research by educationalists and brought to the attention of education policy makers in Germany since the 1990s, even though Germany had taken part in the first international comparative studies in the 1960s. Although the results in all competency areas and age groups tested were mediocre, and although individual Länder (see diagram 2) did take part in the first maths study FIMS and the first science study FISS, Germany was nonetheless absent from most of the international comparisons until the early 1990s (van Ackeren, 2002). Bos and Postlethwaite see a possible explanation for this virtual abstinence in the 1970s and 1980s in the idea that "German pedagogy is traditionally humanities-based" (Bos/Postlethwaite, 2001, p. 253).

Looking back at the development of the German education system in the last few decades, it is clear that Germany has missed out on the structural and didactic modernisation process which other important countries have carried out and financed.

As a consequence of the so-called TIMSS shock and with the aim of overcoming the "black box" image which had been criticised by TIMSS, individual Länder carried out national achievement studies from the mid 90s, which took account of various factors and their complex interaction. Thus, the aim was not just to find differences through comparison, but also to explain these differences using information about context and process variables. The range of these studies was extended to include variables within and outside the school which influence achievement. These are ascertained through questionnaires, which complement the achievement tests and are evaluated, giving more depth to the results (van Ackeren/Rolff 2002). It is no longer merely a question of drawing up general reports of results; rather, the data are intended to give schools an orientation to help improve their teaching.

Three of these studies are particularly important:

- The study "Aspects of learning contexts (and of learning development)" (LAU), conducted in Hamburg for the 5th, 7th, 9th and 11th Years as a comprehensive longitudinal section. So far, it has been done in 1996, 1998 and 2000 for German, Maths and First Foreign Language. It includes the study of attitudes towards learning and teaching, the study of family backgrounds, and it uses student and parents questionnaires (Lehmann/Peek/Gänsefuß, 1997 and 1999; Lehmann *et al.* 2002).
- The "Quality Study of School Mathematics Lessons" (QuaSUM), Brandenburg, 1999, which investigated the achievement level of grades five and nine in Maths. This study used a representative sample and student/parents questionnaires about school- and learning-related attitudes, taking into account their life and educational influences outside school (Lehmann *et al.*, 2000).
- The most recent of the three studies, the "General Maths Enquiry in Rhineland-Palatinate: Competencies, Lesson characteristics, School context" (MARKUS), the first total study of a complete school year (8th Year) in a non-city Land. This was carried out in May 2000 and dealt with student achievement in Maths, lesson and learning structures and contextual influences within and outside school. It used questionnaires for students, teachers and heads of school (Helmke/Jäger, 2002).

The feedback from these comparative studies in the Federal Länder has not yet been uniformly presented; it leaves many questions open and is in need of extensive research. For the projects QuaSUM and MARKUS, plans have been initiated to make the feedback and data obtained applicable for schools. These projects have not yet led to any results being published. The preliminary considerations can, for example, be found in Peek 2001.

Table 7: Participation of Germany in Internationalen Large Scale Assessments (until 2002)

Bezeichnung der Studie *(alphabetische Ordnung)*		Koordination	Zeitraum: Planung	Zeitraum: Erhebung	Zeitraum: Ergebnisse	Inhalte: fachliche Leistungen	Inhalte: Hintergrundmerkmale *(Bsp.)*	Alters-gruppe	Schul-form	Anzahl Länder	beteiligte Bundesländer
Civic Education Study/ Citizenship Education	CIVICS	IEA	1993	1995-97; 1999	2001	politische Bildung	politische Einstellungen, politisches Engagement	14-Jährige	alle ohne Sonder-schulen	28	alle
Computers in Education Study	ComPed	IEA	1985	*Phase I:* 1989/90; *Phase II:* 1992	1993	Computerwissen/ und -umgang (Textverarbeitung, Programmierung)	Lehrerqualifikation, schulische Ausstattung, Lehrpläne, Lehreraus- und -fortbildung	5., 8., 11. Klassen	alle ohne Sonder-schulen	21	9 Bundesländer Pop II und III
First International Mathematics Study	FIMS	IEA	1961	1963-64	1967	Mathematik	Unterrichtspraxis, Lernumgebung, häuslicher und schulischer Einfluss auf Leistung	13-Jährige; Ende S II	Volks-, Realschulen, Gymnasien	12	nur Hessen und Schleswig-Holstein
First International Science Study *(Teil der Six-Subject-Study)*	FISS	IEA	1966	1970	1973	Naturwissenschaften (Erdkunde, Biologie, Physik, naturwiss. Methoden)	Verhaltensmerkmale, fachliches Interesse, Einstellungen, Unterrichtspraxis	10-Jährige, 14-Jährige, Ende S II	Volks-, Realschulen, Gymnasien	mgg. 19	alte Bundesländer
International Study of Reading Literacy	IRLS	IEA	1988	1990-91	1992 *(int.)*, 1995 *(nat.)*	Leseverständnis, Lesegewohnheiten	Lesegewohnheiten, familiäre Merkmale, Einstellungen zum Lesen, Unterricht	3. und 8. Klassen	alle ohne Sonder-schulen	26 bzw. 31	alle
Programme for International Student Assessment	PISA	OECD	1998	2000 2003 2006	Ende 2001 *(erste Phase)*	Leseverständnis, Mathematik, Naturwissenschaften	fächerübergreifende Kompetenzen, vielfältige Indikatoren für Lernergebnisse und ihre Bedingungen	15-Jährige	alle	32	alle
Progress in International Reading Literacy Study *deutsches Äquivalent:* Internation. Grundschul-Lese-Untersuchung	PIRLS/ IGLU	IEA	1999	2001	2002/ 2003	Leseverständnis	Leseinteresse und -gewohnheiten, Freizeitaktivitäten, professionelle Standards, Unterrichtsprozesse	9- bis 10-Jährige	Grund-schulen	38	alle
Third International Mathematics and Science Study	TIMSS	IEA	1989	*SI:* 1994/95 *SII:* 1995/96	1996-1998	Mathematik, Naturwissenschaften	Unterricht, Lehrer, Schulen, außerschulische Lebensumwelt, psychosoziale / individuelle Schülermerkmale	7./ 8. Kl., Ende S II	alle ohne Sonder-schulen	40	alle (ohne Baden-Württemberg)
Written Composition Study bzw. Hamburger Aufsatzstudie		IEA	1981	1984-85	1988	Aufsatzleistungen	Aufsatzcurriculum, Lehrervariablen, schulbedingte und persönliche Faktoren	11. Klassen	allgemein/ berufsbild. Schulen	14	Hamburg

Source: Data compiled by authors

2.7.2 Results of the Monitoring: Weak points of German schools

While the results of the literacy study in the early 1990s (Lehmann, 1995) were initially widely ignored in Germany with the exception a small group of experts, a fierce debate about the quality of the German education system was launched when the TIMS-II and TIMS-III studies were published. The debate was carried out publicly, by education policy makers, school professionals, education researchers and didactics experts. It gained new momentum from the results of the international PISA study, presented in 2001, and the evaluation of PISA within Germany, presented the following year. The conclusions for Germany derived from these studies were shocking, both for the general public and for the experts. The conclusions, whether new or a confirmation of previous conclusions, carried particular weight through the comparison with other industrial nations. The main conclusions were:

- The competency of German students in maths, science and reading literacy at the end of their compulsory school career and at the end of upper secondary level are far behind those of the internationally leading nations, and generally below OECD average.
- This weakness in student achievement is apparent both among the stronger and among the weaker students. This finding attracted particular attention because it contradicted the assumption behind the German tracked school system, namely that students have more support if taught in homogenous learning groups. The finding did not, however, trigger off a broad debate about the advantages and disadvantages of the German school structure.
- Not only is across-the-board school achievement below average, but the difference in achievement between stronger and weaker students is greater in Germany than in any other country.
- The connection between the attainment of competencies and social background is stronger in Germany than in any other country. The German PISA publication (Baumert *et al.*, 2001) regards this as an effect of assigning students at an early age to educational paths of different difficulty levels.
- The evaluation of the school achievement of young people from migrant backgrounds indicates that, compared with other countries with a migrant population comparable to Germany's, German schools encourage students from migrant backgrounds less successfully than schools in other OECD countries.
- Finally, extended samples taken in the different Federal Länder enabled an internal analysis of PISA within Germany, which showed that there is a large difference between cognitive achievements of individual *Länder* in the areas tested. The most successful *Land* is, in international terms, only at the lower end of the top third of the participating countries; the gap between the Länder is alarming, given the commitment made in the Basic Law (*Grundgesetz)* to maintaining equivalent living standards within the Federal Republic.

The main findings from the international Large Scale Assessments, briefly outlined here, were partly confirmed and partly extended by the achievement studies carried out in individual Federal Länder (see 2.3.3). Three results are particularly important:

- Contrary to the philosophy of the tracked school system, the decisions about transition from primary school to the different types of secondary school only partly reflect student ability, if at all.
- The teachers' diagnostic competency is insufficient for a reliable recommendation to a particular type of school. This inadequacy is connected with the fact that primary school recommendations are dealt with according to social class, which means that there are notable overlaps of ability level between students in different types of secondary school.
- The LAU study from Hamburg shows that there is only a slight learning curve in schools at lower secondary level. The first school years of the lower secondary level evidently serve to achieve a homogenous level within the class, appropriate for the type of school, and not so much for learning as such.

The German school system has no evident advantages to counteract the weak points revealed by the Large Scale Assessments of the last ten years. In the course of the education policy debate, which brought this problem to the public's attention, this has had considerable consequences:

- The debate incited the introduction of national quality evaluation programmes for the school system, in nearly all Länder.
- The debate, together with the Federation-Länder-Commission for Educational Planning and Research Funding (BLK), and the Standing Conference of Ministers of Education and Cultural Affairs (KMK), caused the committees responsible for the coordination of educational policy across Länder boundaries to initiate national development programmes to improve school quality. The most important reforms in these programmes were and are aimed at changing instructions.
- The debate gave an unprecedented boost to empirically orientated school research, which most notably led to the support programmes of the German Research Community (DFG).
- The debate led to the KMK's resolution that Germany must participate in future international Large Scale Assessments. Another consequence was that several Länder carried out *land*-wide achievement tests. The tables here show how broad the spectrum of such tests has become within a short time (see also section 2.7.2).
- Finally, the difference between the PISA results of different Länder caused the KMK to initiate the development of national standards of achievement; to be regularly controlled (see chapter 2.3.2). The shock caused by the discrepancy in achievement between Länder caused the revival of an earlier discussion of the advantages and disadvantages of cultural federalism -at least in the run-up to the parliamentary election in the summer and autumn of 2002.

The initial attempts to improve quality in German schools, which have been briefly described here, do not yet add up to a coherent whole. This is partly because the education policy debate is still ongoing, but also partly because the complicated consensus mechanisms of the federal education system make unified responses to universally recognised problems difficult, if not impossible.

Table 8: Large Scale Assessments within Germany (until 2002)

	Bezeichnung der Studie	(alphabetische Ordnung)	Koordination/ wiss. Begleitung	Zeitraum: Planung	Zeitraum: Erhebung	Zeitraum: Ergebnisse	Inhalte: fachliche Leistungen	Inhalte: Hintergrundmerkmale	Altersgruppe	Schulform	Anzahl Schüler	beteil. Länder	Erhebungsart: Stichprobe	Erhebungsart: flächendeckend
länderübergreifende Studien	Bildungsverläufe und psychosoz. Entwicklung im Jugendalter	BIJU	MPIB Berlin	1989	1991, 93, 95, 97, 2000	laufend	Deutsch, Engl., Math., Biol., Physik,	psychosoziale Aspekte, Motivation, Unterricht	1991 Beginn mit 7. Klassen	alle	*z.B. 1. Kohorte:* ca. 10.000	Berlin, MV, NRW, SA	● *(vor allem Längsschnitt)*	
	Deutsch-Englisch-Schülerleistungen-International	DESI	DIPF	1998	2003 *und* 2004	Herbst 2005	Englisch, Deutsch	system., schul. klassenspez. u. individ. Merkmale + Videostudie	9. Klassen	alle	*jeweils ca.* 11.000	alle 16; für Englisch int. Vgl.	● *(Längs- u. Querschn.-Elemente)*	
	Internationale Grundschul-Lese-Untersuchung, nationale Erweiterung *(für 12 Bundesländer)*	IGLU-E	Universität Hamburg	1999	2001	2003	Leseverständnis, Mathem., Naturwiss.	Leseinteresse/ -gewohnheiten, Freizeitaktivit., Unterricht...	Ende 4. Klassen	Grundschule	8.600	alle außer Brandenb., MV, Nieders., SA	●	
	Programme for International Student Assessment - Erweiterung	PISA-E	MPIPF München	1998	2000 2003 2006	2002 *(erste Phase)*	Leseverständnis, Mathe., Naturwiss.	nat. Erweiterung: z.B. zu (außer)-schul. Einflüssen...	Ende S I	alle	50.000	alle	●	
länderinterne Studien	Bayerischer Mathematiktest und Bayerischer Jahrgangsstufentest Deutsch	BMT	Staatsinstitut für Schulpädagogik und Bildungsforschung	1998	*seit* 1998 *bzw.* 2000 *regelmäßig*	jedes Schuljahr	Mathematik, Deutsch	werden nicht erhoben	7. Kl. *(Mathe., Hauptsch.)*, 8. Kl. *(Deutsch, alle Schul.,* 9. Kl. *(alle außer Hauptsch.)*		*bis 2000: Gym. Math.* 33.000	Bayern		●
	Gesamtschule im Vergleich	Fend-Untersuchungen	Zentrum für Bildungsforschung Univ. Konstanz	1969 und 1962	1973, 1976 sowie 1977, 1979	1982	Deutsch, Mathematik, Engl., Physik	Schullaufbahnen, Chancengleichheit, soziale Umwelt	6., 9., 10. Klassen	alle	20.000	vor allem Hessen, NRW, Niedersachsen	●	
	Aspekte der Lernausgangslage (und der Lernentwicklung)	LAU	Amt für Schule/ Humboldt-Univ. Berlin	1995	1996, 98, 2000, 2002	1997, 1999	Deutsch, Mathem., 1. Fremdspr. E/L	fächerübergr. Kompet., schulbez. Einstellungen, soz. Herkunft,...	5., 7., 9., 11. Klassen	alle ohne Sonderschulen	*jeweils ca.* 12.000	Hamburg		● *(Längsschnitt)*
	Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen Unterrichtsmerkmale, Schulkontext	MARKUS	Bildungsminist., Univiversität Koblenz-Landau	1998	2000	Ende 2000	Mathematik	Unterricht, Schulqualität, Lernvoraussetzungen ...	8. Klassen	alle ohne Sonderschulen	44.000	Rheinland-Pfalz		●
	Münchner Hauptschulstudie *(dt. Erweiterung der Classroom Environment Study)*	Projektverbund Münchner Studien	MPIPF München	Beginn 80er Jahre	1983-1985	1989	Mathematik	leistungsrelevante Motive und Einstellungen, Elternverhalten..	Beginn der 5. bis Ende der 6. Klassen	Hauptschulen	1.000	Bayern *(München)*	● *(Längsschnitt)*	
	Longitudinalstudie zur Genese individueller Kompetenzen (LOGIK)			1982	1984-1993 *1997/98 Nachuntersuchung*	laufend	Lesen, Rechtschreibung, Mathe., Naturwiss.	Intelligenz, Gedächtnis, Motivation, soz. Fähigkeiten ...	4-12 Jahre	Schwerpunkt: Grundschulzeit	200	Bayern *(München)*	● *(Längsschnitt)*	
	Schulorganisierte Lernangebote und Sozialisation von Talenten, Interessen und Kompetenzen (SCHOLASTIK)			1988	1988-1991	1993	Mathematik, Deutsch	kogn. Fertigkeiten, motivationale Schülermerkmale..	1 - 4. Klasse	Grundschulen	1.300	Bayern	● *(Längsschnitt)*	
	Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik	QuaSUM	Bildungsminist., Humboldt-Universität	1997	1999	2000	Mathematik	Unterrichtsmerkmale, Schulalltag, Lernumwelt	5. und 9. Klassen	alle ohne Sonderschulen	13.000	Brandenburg	●	
	Vergleichsuntersuchung im Saarland		Schulbehörde	1998	1998/1999	Schuljahr 1998/1999	Mathematik, Deutsch	werden nicht erhoben	Klassen 7 und 9	Gymn., Real-/ Gesamts	ca. 9.000	Saarland		●