

---

---

## Originalarbeiten

---

---

**Lars Balzer und Reinhold S. Jäger**

### Fachleistung Mathematik in MARKUS

Die aktuelle Bildungspolitik in Deutschland orientiert sich bei der Frage nach der Leistungsfähigkeit des deutschen Schulsystems mittlerweile stark an nationalen sowie internationalen empirischen Schulleistungstudien. Die Ergebnisse internationaler Studien wie TIMSS (vgl. TIMSS/II; Baumert et al., 1997 und TIMSS/III; Baumert, Bos & Lehmann, 2000a, b) und PISA (vgl. Max-Planck-Institut für Bildungsforschung, 2001) und nationaler Studien wie der Hamburger Studie LAU (vgl. Lehmann, Gänsfuß & Peek, 1999, Lehmann, Peek & Gänsfuß, 1997), der Brandenburger Studie QuaSUM (vgl. Lehmann et al., 2000) und jetzt auch der Rheinland-Pfälzischen Studie MARKUS werden zunehmend als Ausgangspunkt gesehen, um Veränderungen im Bildungswesen einzuleiten. Allen genannten Studien ist gemeinsam, dass sie zunächst in sehr detaillierter Art und Weise die Leistungsfähigkeit des Bildungssystems auf verschiedenen Ebenen beschreiben, darüber hinaus aber auch spezifische Fragestellungen bearbeiten, die über die reine Leistungsbeschreibung hinausgehen.

Dieser Artikel steht am Anfang eines Themenheftes zum Projekt MARKUS und widmet sich zunächst der Beschreibung der für die Leistungserhebung verwendeten Instrumente und anschließend der Beschreibung der Mathematikleistung der Schülerinnen und Schüler des Landes Rheinland-Pfalz, bevor in den folgenden Artikeln spezifische Fragestellungen bearbeitet werden.

#### FRAGESTELLUNGEN

Konkret werden in diesem Beitrag nach der kurzen Beschreibung der Instrumente folgende Fragestellungen angegangen:

- a) Wie schneiden rheinland-pfälzische Schülerinnen und Schüler der 8. Klassen im Bildungsgangvergleich ab?
- b) Wie schneiden rheinland-pfälzische Schülerinnen und Schüler der 8. Klassen im Schulartvergleich ab?
- c) Wie schneiden rheinland-pfälzische Schülerinnen und Schüler der 8. Klassen im Vergleich der Bildungsgänge innerhalb der Schularten ab?
- d) Gibt es bildungsgang- und schulartunabhängig Leistungen, die als Spitzenleistungen bezeichnet werden können?

Diesen Fragestellungen liegt die Notwendigkeit zu Grunde, vor komplexen Analysen, wie sie in weiteren Artikeln dieses Themenheftes gezeigt werden, zunächst einmal die vorhandenen Daten zu beschreiben, also eine Bestandsaufnahme des vielen Analysen zu Grunde liegenden Leistungsstandes durchzuführen.

## METHODEN

### Instrumente

Der in MARKUS eingesetzte Mathematiktest besteht aus drei Teilen (für eine ausführliche Beschreibung vgl. Helmke & Jäger (im Druck)):

- Einem aus 15 Items der TIMSS II-Untersuchung bestehenden Testteil (MARKUS-T); eine komplette Liste der veröffentlichten TIMSS II-Items findet sich in Baumert et al., 1998).
- Einem Testteil zur Erfassung von Vorkenntnissen (MARKUS-V), welcher aus 6 Items besteht, die eine Teilmenge von MARKUS-T darstellen.
- Einem neuentwickelten curriculumorientierten Testteil (MARKUS-C), der die Grundlage für diesen Beitrag bildet. Er wird nachfolgend näher beschrieben.

MARKUS-C wurde für jeden der vier rheinland-pfälzischen Bildungsgänge (Gymnasium, Realschule, Hauptschule Aufbaukurs, Hauptschule Grundkurs) als eigener Test entwickelt. Diese Tests wurden jeder Schulart mit all ihren Bildungsgangvarianten zugeordnet. Daraufhin wurde auf der Grundlage des Multi-Matrix-Designs, welches es erlaubt, durch die Aufteilung der Gesamtheit aller Testitems auf mehr als eine Testform einen inhaltlich breiteren Bereich abzudecken, jeder Bildungsgangtest in 2 Testformen A und B unterteilt. Jede Testform

enthielt unterschiedliche Items aus unterschiedlichen Stoffgebieten des jeweiligen Curriculum. Insgesamt wurden also 8 Tests (vier Bildungsgänge x zwei Testformen) entwickelt.

In jedem dieser 8 Tests gab es einen bildungsgangspezifischen Teil (Additum) und einen bildungsgangübergreifenden Teil (Fundamentum), der Grundlage für alle Bildungsgänge war.

Um auf der einen Seite dem Schulalltag und auf der anderen Seite der Testökonomie gerecht zu werden, wurden Items sowohl im Mehrfachwahlantwortenformat (multiple-choice mit 6 Antwortalternativen) als auch im offenen Antwortformat erstellt.

### Stichprobe

Datengrundlage dieses Artikels ist die im Rahmen der MARKUS-Studie erfolgte Gesamterhebung aller 8. Klassen des Landes Rheinland-Pfalz. Details dieser Population sind bereits im Editorial beschrieben worden.

Die Auswertungen dieses Artikels basieren auf der Gruppe derjenigen Schülerinnen und Schüler, von denen Mathematiktest-Daten vorliegen. Dies waren insgesamt 37520 Schülerinnen und Schüler. Aus dieser Tatsache können kleine Abweichungen gegenüber der Populationsdarstellung im Editorial (in diesem Artikel interessieren nur die Testdaten, nicht die Fragebogendaten) resultieren. Die jeweils zu Grunde gelegten Fallzahlen werden in den entsprechenden Darstellungen deswegen genannt.

### Skalierung

Dem „edumetrischen Skalierungspragmatismus“ (Baumert, Köller, Lehrke & Brockmann, 2000, S. 67) folgend wurde für MARKUS im Rahmen der probabilistischen Testtheorie ein eindimensionales Raschmodell angepasst. Dafür wurden aus skalierungsstatistischen Gründen alle Mathematikitems der MARKUS-Studie für die Berechnung einer MARKUS-Gesamtsskala herangezogen, denn es mussten Items in die Analyse einbezogen werden, die quasi eine Verknüpfung zwischen den verschiedenen Testversionen ermöglichten. Diese Verknüpfung gewährleisten so genannte Ankeritems. In MARKUS waren dies zum einen die von allen Schülerinnen und Schülern zu bearbeitenden TIMSS-Items und zum anderen die in mehreren Testversionen vorkommenden Fundamentumitems.

Für die MARKUS-Gesamtsskala wurde eine Metrik mit Mittelwert 700 und Standardabweichung 100 gewählt. Die Wahl dieser Metrik ist willkürlich, unter-

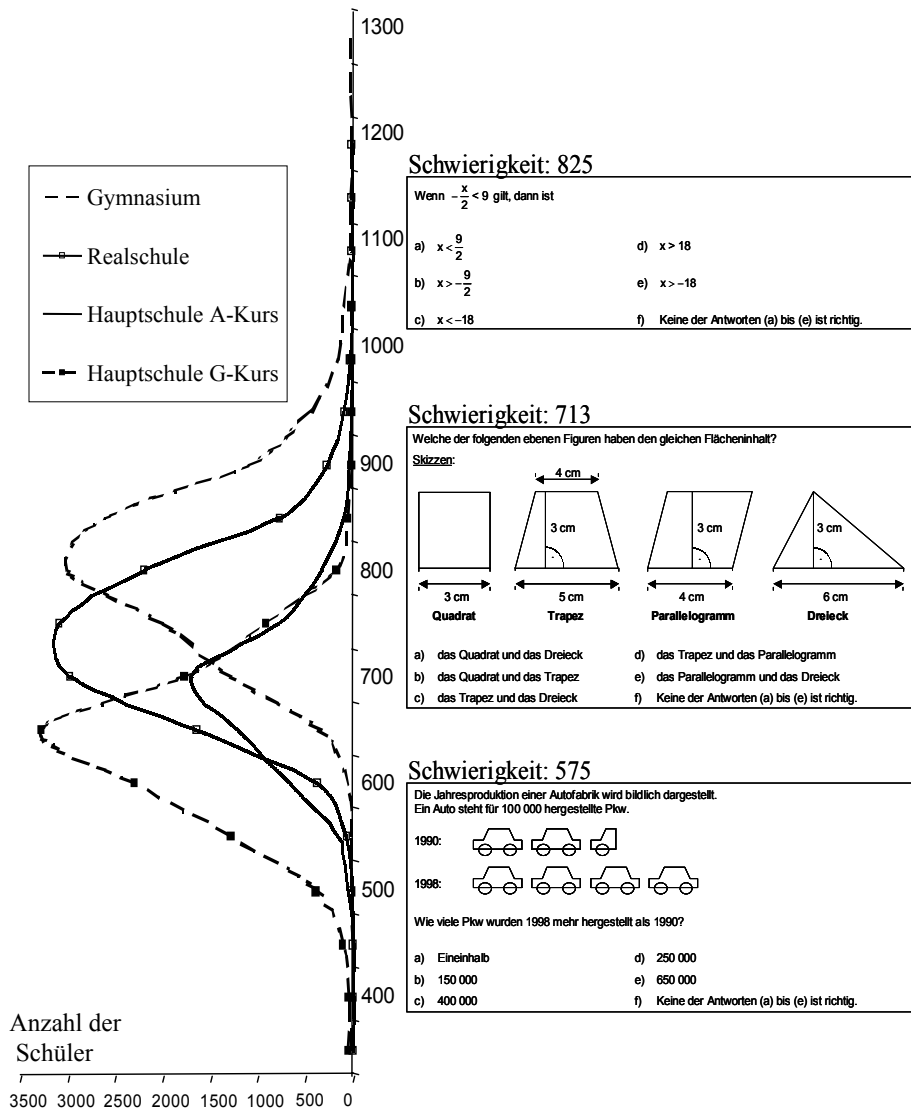
scheidet sich aber von anderen Schulleistungsstudien und verhindert damit methodisch unsaubere Vergleiche.

Diese Gesamtskala ermöglicht es, alle in MARKUS verwendeten Items und alle Schülerinnen und Schüler auf einer gemeinsamen Dimension abzubilden. Ein weiterer großer Vorteil dieser Art der Skalierung ist es, dass die Fähigkeit einer Person auch dann zuverlässig zu schätzen ist, wenn nur eine Teilmenge der Items bearbeitet wurde, wie dies im Multi-Matrix-Design von MARKUS der Fall ist. Ausgehend von den Schwierigkeitsschätzungen dieser Gesamtskala wurden verschiedene weitere Skalen gebildet. Die wichtigste ist die Skala MARKUS-C, die nur curriculare Items beinhaltet. Um Verwechslungen mit der alle Items umfassenden Gesamtskala auszuschließen, wurde diese Skala auf einen Mittelwert von 300 mit einer Standardabweichung von 50 normiert. Ebenso wurde mit bereichsspezifischen Skalen (bereichsbezogene Skalen für die einzelnen Lehrplanbereiche wie z. B. „Beschreibende Statistik“) verfahren.

Abbildung 1 veranschaulicht die gemeinsame Dimension von Itemschwierigkeit und Personenfähigkeit auf Grundlage der Gesamtskala. Auf der linken Seite der Abbildung sind die Fähigkeitswerte der Schülerinnen und Schüler abgetragen. Die inhaltliche Interpretation solcher Fähigkeitswerte erfolgt im Ergebnisteil. Auf der gleichen Dimension werden die Items mit ihren Schwierigkeiten abgebildet. In Abbildung 1 erfolgt dies exemplarisch für drei Items.

Abbildung 1: Fähigkeiten und Schwierigkeiten in MARKUS

**Fähigkeit / Schwierigkeit (M = 700, SD = 100)**



## ERGEBNISSE

Im Folgenden werden zunächst die Leistungen aller Schülerinnen und Schüler im Bildungsgangvergleich, dann im Schulartvergleich und daraufhin im Vergleich der Bildungsgänge innerhalb der Schularten dargestellt. Im Anschluss daran wird untersucht, ob es bildungsgang- und schulartunabhängig Leistungen gibt, die als Spitzenleistungen bezeichnet werden können.

Von besonderem Interesse für die Beschreibung der Leistungsfähigkeit der rheinland-pfälzischen Schülerinnen und Schüler ist dabei derjenige Teil des Tests, der curriculare Validität beansprucht, da er sich in seinen Inhalten auf das komplette Curriculum im Fach Mathematik der 8. Klasse bezieht. Dieser wird wie schon beschrieben als MARKUS-C bezeichnet. Deswegen werden die zentralen Berechnungen zur Beschreibung der mathematischen Fachleistungen auf der Grundlage von MARKUS-C erfolgen.

### Mathematische Fachleistung im Bildungsgangvergleich

Das mittlere Niveau der mathematischen Fachleistung der Schülerinnen und Schüler, bezogen auf die einzelnen Bildungsgänge, unterscheidet sich erwartungsgemäß (vgl. Tabelle 1): Schülerinnen und Schüler des Bildungsgangs Gymnasium schneiden besser ab als diejenigen im Bildungsgang Realschule, diese wiederum sind besser als diejenigen im Bildungsgang Hauptschule A-Kurs, welche wiederum besser sind als diejenigen im Bildungsgang Hauptschule G-Kurs.

Tabelle 1: Bildungsgänge im Leistungsvergleich: MARKUS-C

<b>Bildungsgang</b>	<b>N</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>
Gymnasium	11069	339	42,8
Realschule	11355	306	36,9
Hauptschule A-Kurs	4952	278	33,9
Hauptschule G-Kurs	10144	262	41,2

Die Unterschiede zwischen den Bildungsgängen sind statistisch signifikant ( $p < .001$  sowohl für die einfaktorielle Varianzanalyse wie auch für alle paarweisen post-hoc-Vergleiche, vgl. Tabelle 2).

Besonders bei sehr großen Untersuchungsgruppen sollte bei der Frage nach der statistischen Signifikanz allerdings nicht Halt gemacht werden. Ob ein gefundener statistisch signifikanter Effekt auch tatsächlich praktisch bedeutsam ist, wird deswegen mit dem Effektstärkemaß  $\omega^2$  untersucht (vgl. Wolf, 2001).

Ein  $\omega^2$  von .373 bescheinigt der Varianzanalyse einen sehr großen Effekt und weist damit auf praktisch bedeutsame Gruppenunterschiede innerhalb dieser Analyse hin (die z. B. von Cohen (1977) und Bortz & Döring (1995) formulierten Konventionen zur Interpretation von Effektstärken sollten zwar nicht dogmatisch für alle Themenbereiche und Analysemethoden exakt übernommen werden (vgl. Wolf, 2001), können aber als Anhaltspunkt zur Interpretation herangezogen werden). Die Signifikanzwerte und Effektstärken der paarweisen post-hoc-Vergleiche sind der folgenden Tabelle 2 zu entnehmen.

Tabelle 2: Effektstärken der Leistungsunterschiede (MARKUS-C) im Bildungsgangvergleich

Bildungsgang	Vergleichsbildungsgang	p	$\omega^2$
Gymnasium	Realschule	< .001	.148
	Hauptschule A-Kurs	< .001	.373
	Hauptschule G-Kurs	< .001	.460
Realschule	Hauptschule A-Kurs	< .001	.121
	Hauptschule G-Kurs	< .001	.242
Hauptschule A-Kurs	Hauptschule G-Kurs	< .001	.042

Einen eher kleinen Effekt und damit einen praktisch eher wenig bedeutsamen Leistungsunterschied findet man beim Vergleich der Bildungsgänge Hauptschule A-Kurs mit Hauptschule G-Kurs. Eher mittlere Effekte liegen den Leistungsunterschieden zwischen den Bildungsgängen Gymnasium und Realschule und zwischen den Bildungsgängen Realschule und Hauptschule A-Kurs zu Grunde. Die restlichen Effektstärken weisen auf praktisch sehr bedeutsame Gruppenunterschiede hin.

Betrachtet man in einem weiteren Schritt die Standardabweichungen der einzelnen Gruppen, so ist zu erkennen, dass erhebliche Überschneidungen im Fähigkeitsspektrum der Schülerinnen und Schüler zwischen den verschiedenen Bildungsgängen bestehen. Auf diesen Aspekt wird im Ergebnisteil zu Spitzenleistungen noch näher eingegangen.

### Mathematische Fachleistung im Schulartvergleich

Der Vergleich des mittleren Niveaus der mathematischen Fachleistung, bezogen auf die einzelnen Schularten, ist inhaltlich etwas komplexer. In Rheinland-Pfalz gibt es außer den Schularten Gymnasium und Realschule auch Schularten, die mehrere Bildungsgänge umfassen können. Es sind dies die Hauptschule, die

Integrierte Gesamtschule, die Regionale Schule und die Duale Oberschule. Bei diesen werden Schülerinnen und Schüler verschiedener Bildungsgang- bzw. Leistungszuordnung im Klassenverband (sogenannte binnendifferenzierte Klassen bzw. Kurse), oder nach Leistungen differenziert in verschiedenen Kursen unterrichtet. Nimmt man einen Vergleich auf der Ebene dieser Schularten vor, so zeigt sich, dass Schülerinnen und Schüler der Schulart Gymnasium am besten abschneiden, gefolgt von Schülerinnen und Schülern der Realschule, der Integrierten Gesamtschule, der Regionalen Schule, der Dualen Oberschule und der Hauptschule (vgl. Tabelle 3).

Tabelle 3: Schularten im Leistungsvergleich: MARKUS-C

<b>Schulart</b>	<b>N</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>
Gymnasium	10598	340	43,0
Realschule	9651	307	36,9
Integrierte Gesamtschule	1876	289	45,2
Regionale Schule	2014	286	41,8
Duale Oberschule	86	268	42,3
Hauptschule	13295	267	39,6

Die einfaktorielle Varianzanalyse erweist sich bei einem  $p < .001$  als statistisch signifikant und ist bei einem  $\omega^2$  von .347 mit einer sehr großen Effektstärke versehen. Das bedeutet, dass es innerhalb dieser Analyse praktisch bedeutsame Gruppenunterschiede gibt.

Die Effektstärken und Signifikanzwerte der paarweisen post-hoc-Vergleiche sind der folgenden Tabelle 4 zu entnehmen.



Tabelle 4: Effektstärken der Leistungsunterschiede (MARKUS-C) im Schulartvergleich

Schulart	Vergleichsschulart	p	$\omega^2$
Gymnasium	Realschule	< .001	.144
	Integrierte Gesamtschule	< .001	.143
	Regionale Schule	< .001	.177
	Duale Oberschule	< .001	.022
	Hauptschule	< .001	.431
Realschule	Integrierte Gesamtschule	< .001	.024
	Regionale Schule	< .001	.038
	Duale Oberschule	< .001	.010
	Hauptschule	< .001	.212
Integrierte Gesamtschule	Regionale Schule	.388	--
	Duale Oberschule	< .001	.009
	Hauptschule	< .001	.025
Regionale Schule	Duale Oberschule	.005	.007
	Hauptschule	< .001	.022
Duale Oberschule	Hauptschule	1.000	--

Tabelle 4 zeigt, dass beinahe alle Vergleiche zwischen den Schularten statistisch signifikant ausfallen. Eine Ausnahme stellen die Vergleiche zwischen Integrierter Gesamtschule mit der Regionalen Schule und Dualer Oberschule mit der Hauptschule dar.

Betrachtet man die praktische Bedeutsamkeit mit Hilfe von  $\omega^2$ , so fällt der Vergleich mit der Schulart Gymnasium durchweg bedeutsam aus. Mit Ausnahme des Vergleichs mit der Dualen Oberschule, der bei ähnlicher Mittelwertsdifferenz auch auf Grund der geringeren Gruppengröße der Dualen Oberschule einen geringeren t-Wert und in der Folge davon eine geringere Effektstärke aufweist, resultieren durchweg mittlere bis hohe Effektstärken. Die Leistungsunterschiede der anderen Schularten sind hingegen (mit Ausnahme des Vergleiches Realschule mit Hauptschule) praktisch weniger relevant.

Ungeachtet der mittleren Leistungsunterschiede zeigt aber auch diese Darstellung anhand der Standardabweichungen große Überlappungen in den mathematischen Fachleistungen zwischen den Schularten.

### Mathematische Fachleistung im Vergleich der Bildungsgänge innerhalb der Schularten

Den besonderen Gegebenheiten der Schularten mit mehreren Bildungsgängen wird man am besten dadurch gerecht, dass man die einzelnen Bildungsgänge separat aufschlüsselt und daraufhin auf der Basis „Bildungsgänge innerhalb von Schularten“ Vergleiche anstellt, denn ansonsten würde die Einschätzung der Leistungsfähigkeit einer Schulart in hohem Maße durch die Größe einer Bildungsganggruppe innerhalb dieser Schulart bestimmt. Deswegen werden in Tabelle 5 die Leistungen der Schülerinnen und Schüler nach den Bildungsgängen innerhalb von Schularten aufgeschlüsselt.

Auch bei dieser Aufschlüsselung fällt auf, dass zum einen die Reihenfolge der Bildungsgänge auf der Basis der Leistungsdaten erhalten bleibt und dass sich zum anderen auch bei dieser Art der Gruppierung nicht zu vernachlässigende Überschneidungen in der Leistungsfähigkeit zwischen den Analysegruppen zeigen.

Tabelle 5: Bildungsgänge in Schularten im Leistungsvergleich: MARKUS-C

<b>Bildungsgang</b>	<b>in Schulart</b>	<b>N</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>
Gymnasium	am Gymnasium	10598	340	43,0
	an der Integrierten Gesamtschule	471	325	35,9
Realschule	an der Realschule	9651	307	36,9
	an der Integrierten Gesamtschule	723	293	34,7
	an der Regionalen Schule	949	304	36,8
	an der Dualen Oberschule	32	291	33,8
Hauptschule A-Kurs	an der Hauptschule	4952	278	33,9
Hauptschule G-Kurs	an der Hauptschule	8343	261	41,4
	an der Integrierten Gesamtschule	682	259	40,6
	an der Regionalen Schule	1065	270	39,3
	an der Dualen Oberschule	54	254	41,0

Eine einfaktorielle Varianzanalyse auf der Basis der unabhängigen Variable „Bildungsgänge innerhalb der Schularten“ erweist sich als statistisch signifikant ( $p < .001$ ). Die zugehörige Effektstärke  $\omega^2$  von .376 ist als sehr großer Effekt zu interpretieren. Das bedeutet, dass es innerhalb dieser Analyse praktisch bedeutende Gruppenunterschiede gibt.

Ein wichtiger Aspekt dieser Darstellung besteht nun darin, nicht die Bildungsgänge oder die Schularten miteinander zu vergleichen, denn dies ist schon in vorangegangenen Analysen erfolgt. Viel interessanter ist es nun zu überprüfen, ob die Schulartunterschiede innerhalb eines Bildungsgangs statistisch signifikant und praktisch bedeutsam sind. Aus diesem Grund werden die Signifikanzwerte und Effektstärken der paarweisen post-hoc-Vergleiche in der nachfolgenden Tabelle 6 ausschliesslich innerhalb eines Bildungsgangs dargestellt.

Tabelle 6: Effektstärken der Leistungsunterschiede (MARKUS-C) im Vergleich der Bildungsgänge innerhalb der Schularten

Bildungsgang innerhalb Schulart	Vergleichsgruppe	p	$\omega^2$
Gymnasium am Gymnasium	Gymnasium an der Integrierten Gesamtschule	<.001	.007
Realschule an der Realschule	Realschule an der Integrierten Gesamtschule	<.001	.011
	Realschule an der Regionalen Schule	.818	--
	Realschule an der Dualen Oberschule	.852	--
Realschule an der Integrierten Gesamtschule	Realschule an der Regionalen Schule	<.001	.022
	Realschule an der Dualen Oberschule	1.000	--
Realschule an der Regionalen Schule	Realschule an der Dualen Oberschule	.969	--
Hauptschule G-Kurs an der Hauptschule	Hauptschule G-Kurs an der Integrierten Gesamtschule	1.000	--
	Hauptschule G-Kurs an der Regionalen Schule	<.001	.004
	Hauptschule G-Kurs an der Dualen Oberschule	.998	--
Hauptschule G-Kurs an der Integrierten Gesamtschule	Hauptschule G-Kurs an der Regionalen Schule	<.001	.015
	Hauptschule G-Kurs an der Dualen Oberschule	1.000	--
Hauptschule G-Kurs an der Regionalen Schule	Hauptschule G-Kurs an der Dualen Oberschule	.609	--

Tabelle 6 zeigt, dass es zwar einige statistisch signifikante Unterschiede innerhalb eines Bildungsgangs gibt, diese sind aber praktisch nicht bedeutsam. Das bedeutet, dass Schülerinnen und Schüler aller gleich benannten Bildungsgänge auch vergleichbare Leistungen erbringen, unabhängig davon, in welcher Schulart sie unterrichtet werden.

### Spitzenleistungen in der mathematischen Fachleistung

Aus der Darstellung in den vorangegangenen Abschnitten geht hervor, dass zwischen der mathematischen Fachleistung der Schülerinnen und Schüler unterschiedlicher Bildungsgänge und Schularten zum Teil erhebliche Überlappungen existieren.

Es ist daher die Frage interessant, ob es in einer Gruppe der Jahrgangsbesten auch Schülerinnen und Schüler der Bildungsgänge und Schularten gibt, die zu formal geringer qualifizierten Abschlüssen führen. Zur Gruppe der Jahrgangsbesten werden in diesem Zusammenhang diejenigen Schülerinnen und Schüler gerechnet, deren mathematische Fachleistung mindestens eine Standardabweichung über dem Mittelwert der Fähigkeitsskala von MARKUS-C bezogen auf die gesamte Jahrgangsstufe liegt. Auf der Grundlage dieser Festlegung werden alle diejenigen Schülerinnen und Schüler dieser Gruppe zugerechnet, die bei MARKUS-C einen Fähigkeitswert von mindestens 350 erreichen; dies sind 5717 Personen.

Betrachtet man nun, wie viel Prozent eines jeden Bildungsgangs zu dieser Spitzengruppe gehören, so stellt man fest, dass 38,7% aller Schülerinnen und Schüler des gesamten Bildungsgangs Gymnasium und 10,5% des Bildungsgangs Realschule, aber auch 1,7% des Bildungsgangs Hauptschule A-Kurs und 1,5% des Bildungsgangs Hauptschule G-Kurs in dieser Gruppe vertreten sind (vgl. Tabelle 7).

Tabelle 7: Bildungsgangspezifische Anteile an der Spitzengruppe

<b>Bildungsgang</b>	<b>N-Gesamt</b>	<b>N-Spitzengruppe</b>	<b>Anteil in %</b>
Gymnasium	11069	4288	38,7%
Realschule	11355	1189	10,5%
Hauptschule A-Kurs	4952	84	1,7%
Hauptschule G-Kurs	10144	156	1,5%

Analog zur Darstellung der vorangegangenen Abschnitte wird die Spitzengruppe auch hinsichtlich der Schulart und hinsichtlich des Bildungsgangs mit der Zuordnung zur Schulart näher beschrieben. Die Ergebnisse sind Tabelle 8 und Tabelle 9 zu entnehmen.

Tabelle 8: Schulartspezifische Anteile an der Spitzengruppe

Schulart	N-Gesamt	N-Spitzengruppe	Anteil in %
Gymnasium	10598	4169	39,3%
Realschule	9651	1061	11,0%
Integrierte Gesamtschule	1876	158	8,4%
Regionale Schule	2014	122	6,1%
Duale Oberschule	86	1	1,2%
Hauptschule	13295	206	1,5%

Tabelle 9: Anteile an der Spitzengruppe der Gruppe Bildungsgänge in Schularten

Bildungsgang	in Schulart	N Gesamt	N Spitzen- gruppe	Anteil in %
Gymnasium	am Gymnasium	10598	4169	39,3%
	an der Integrierten Gesamtschule	471	119	25,3%
Realschule	an der Realschule	9651	1061	11,0%
	an der Integrierten Gesamtschule	723	33	4,6%
	an der Regionalen Schule	949	94	9,9%
	an der Dualen Oberschule	32	1	3,1%
Hauptschule A-Kurs	an der Hauptschule	4952	84	1,7%
Hauptschule G-Kurs	an der Hauptschule	8343	122	1,5%
	an der Integrierten Gesamtschule	682	6	8,8%
	an der Regionalen Schule	1065	28	2,6%
	an der Dualen Oberschule	54	0	0,0%

Man kann Tabelle 8 und Tabelle 9 entnehmen, dass es in jeder Schulart und beinahe auch in jedem Bildungsgang innerhalb einer Schulart Schülerinnen und Schüler gibt, die mit ihren Leistungen in die Spitzengruppe des gesamten Jahrganges vorzudringen vermögen.

## ZUSAMMENFASSUNG UND AUSBLICK

Die vorliegende Arbeit hat in einem ersten Schritt die mathematische Fachleistung, wie sie im Projekt MARKUS in allen 8. Klassen des Landes Rheinland-Pfalz erhoben wurde, dargestellt. Dabei ergaben Vergleiche der Bildungsgänge statistisch signifikante und zugleich bedeutsame Unterschiede in der erwarteten Reihenfolge (Gymnasium, Realschule, Hauptschule A-Kurs und Hauptschule G-Kurs).

Die wegen der schulartspezifischen inneren Leistungsdifferenzierung inhaltlich etwas komplexeren Vergleiche zwischen den in Rheinland-Pfalz existierenden Schularten Gymnasium, Realschule, Integrierte Gesamtschule, Regionale Schule, Duale Oberschule und Hauptschule stellten diese Unterschiede unter einem etwas anderen Blickwinkel dar. Hier ergaben sich statistisch signifikante und bedeutende Unterschiede zwischen der Schulart Gymnasium und allen anderen Schularten sowie der Schulart Realschule mit der Schulart Hauptschule. Alle anderen Schulartvergleiche sind weniger relevant.

Der Vergleich der Bildungsgänge innerhalb der Schularten ergab keine bedeutenden Unterschiede. Die Bildungsgänge innerhalb der verschiedenen Schularten erbringen also weitgehend vergleichbare Leistungen.

Inhaltlich interessanter als die reine Darstellung der Leistungsunterschiede erwies sich allerdings eine Analyse der Spitzenleistungen der mathematischen Fachleistungen. Es zeigte sich, dass aus beinahe allen Schularten und Bildungsgängen und den daraus resultierenden Kombinationen Schülerinnen und Schüler stammen, die zu den Jahrgangsbesten gezählt werden können. Sollten sich diese Ergebnisse in der weiteren Forschung über Mathematikleistung und über Leistung in anderen Fächern bestätigen, könnte für die Schulorganisation Klärungs- und Handlungsbedarf entstehen. So müsste z. B. die Frage gestellt und beantwortet werden, ob das derzeitige Schularten- und Bildungsgangsystem optimal geeignet ist, die Leistungsfähigkeit der Schülerinnen und Schüler bedarfsgerecht zu fördern.

## LITERATUR

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

- Baumert, J., Köller, O., Lehrke, M. & Brockmann, J. (2000). Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS/III) – Technische Grundlagen. In J. Baumert, W. Bos, & R. Lehmann (Hrsg.), TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Bd. 1) (S. 31-84). Opladen: Leske + Budrich.
- Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J., Patjens, S., Jungclaus, H. & Günther, W. (Hrsg.) (1998). Testaufgaben Mathematik TIMSS 7./8. Klasse (Population 2) – (Materialien aus der Bildungsforschung, Band 60). Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller O. & Neubrand, J. (1997). TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Borz, J. & Döring, N. (1995). Forschungsmethoden und Evaluation (2. Aufl.). Berlin: Springer.
- Cohen, J. (1977). Statistical power analysis for the behavioral sciences. New York: Academic Press.
- Helmke, A. & Jäger, R. S. (Hrsg.). (im Druck). MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik.
- Lehmann, R. H., Gänsfuß, R., & Peek, R. (1999). Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern an Hamburger Schulen – Klassenstufe 7. Hamburg: Behörde für Schule, Jugend und Berufsbildung, Amt für Schule.
- Lehmann, R. H., Peek, R., & Gänsfuß, R. (1997). Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Hamburg: Behörde für Schule, Jugend und Berufsbildung, Amt für Schule.
- Lehmann, R. H., Peek, R., Gänsfuß, R., Lutkat, S., Mücke, S. & Barth, I. (2000). QuaSUM. Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik. Ergebnisse einer repräsentativen Untersuchung im Land Brandenburg. Potsdam: Ministerium für Bildung, Jugend und Sport des Landes Brandenburg.
- Max-Planck-Institut für Bildungsforschung (2001). PISA Deutschland – Programme for International Student Assessment. Online im Internet: URL: <http://www.mpib-berlin.mpg.de/pisa/> [Stand: 01.11.2001].
- Wolf, B. (2001). Effektstärkemaße. In D. H. Rost (Hrsg.). Handwörterbuch Pädagogische Psychologie (S. 96-102). Weinheim: Psychologische Verlags Union.

*Anschrift der Autoren:*

Lars Balzer, Dipl.-Psych., Zentrum für empirische pädagogische Forschung (ZepF) der Universität Koblenz-Landau, Campus Landau, Friedrich-Ebert-Str. 12, D-76829 Landau. Email: [balzer@zepf.uni-landau.de](mailto:balzer@zepf.uni-landau.de)

Prof. Dr. Reinhold S. Jäger, Zentrum für empirische pädagogische Forschung (ZepF) der Universität Koblenz-Landau, Campus Landau, Friedrich-Ebert-Str. 12, D-76829 Landau. Email: [jaeger@zepf.uni-landau.de](mailto:jaeger@zepf.uni-landau.de)



Balzer, L. & Jäger, R. S. (2001). Fachleistung Mathematik in MARKUS. *Empirische Pädagogik*, 15, 535-551.

Im Rahmen der Studie MARKUS (Mathematik-Gesamterhebung Rheinland-Pfalz, Kompetenzen, Unterrichtsmerkmale, Schulkontext) werden Daten zur mathematischen Fachleistung auf Grundlage einer Testung der gesamten 8. Jahrgangstufe in Rheinland-Pfalz ausgewertet. Dies sind 37520 Schülerinnen und Schüler. Nach einer Darstellung der erwarteten Unterschiede zwischen Bildungsgängen und Schularten und nach dem Beleg, dass sich gleiche Bildungsgänge in verschiedenen Schularten nicht bedeutsam voneinander unterscheiden, wird gezeigt, dass es bildungsgang- und schulartunabhängig Leistungen gibt, die als Spitzenleistungen innerhalb des Jahrgangs angesehen werden können.

Balzer, L. & Jäger, R. S. (2001). Mathematical achievement in the MARKUS-study. *Empirische Pädagogik*, 15, 535-551.

Based on the results of a test of 37520 pupils (all 8th graders in Rheinland-Pfalz, Germany) in the MARKUS-study mathematical achievement is described. After showing the expected differences between different groups of the German school system mathematical top performance is analysed. It can be pointed out that mathematical top performance happens in all different groups of the German school system.