

3 Bundesdeutscher Vergleich

Der zweite Teil dieser Untersuchung hat die PISA-Ergänzungsstudie zum Gegenstand.⁹⁵ Im Zentrum steht die Frage, weshalb die Testergebnisse der Schülerinnen und Schüler über die einzelnen Bundesländer stark variieren.

3.1 Testleistung Bundesland

Der nationale Ergänzungstest (PISA-E) misst *mathematische* und *naturwissenschaftliche Kompetenz* sowie *Lesekompetenz* in einer standardisierten nationalen Metrik. Im Vergleich der Bundesländer erzielen Bayern und Baden-Württemberg jeweils die höchsten Werte sowie die Bundesländer Brandenburg und Bremen (insofern berücksichtigt) hingegen niedrigere.

Das PISA-Stichprobendesign sieht so genannte Klumpenstichproben vor. Klumpenstichproben können im Vergleich zu gleich großen Zufallsstichproben bei einfacher Zufallsauswahl zu präziseren Schätzungen von Populationsmerkmalen führen,

wenn die einzelnen Klumpen die Grundgesamtheit möglichst gleichmäßig repräsentieren. (...) Dies ist jedoch bei Schulklassen oder Schulen typischerweise nicht (im Original kursiv, d. Verf.) der Fall. Klassen und Schulen unterscheiden sich leistungsmäßig in Deutschland allein aufgrund der Gliederung des Schulwesens erheblich (...)^{96 97}

Nun besuchen der PISA-E Studie zufolge in Bayern 14,1 Prozent der 15-jährigen Schülerinnen und Schüler bereits berufliche Schulen.⁹⁸ In Baden-Württemberg und Rheinland-Pfalz finden sich Quoten von 9,3 beziehungsweise 8,6 Prozent. In Bundesländern, die eine 10-jährige Vollzeitschulpflicht aufweisen, wie Bremen, Brandenburg, Berlin oder Nordrhein-Westfalen, besuchen praktisch alle 15-Jährigen noch eine allgemein bildende Schule.

⁹⁵ Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich

⁹⁶ Klieme, Baumert, Köller, Bos: Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung von Kompetenzen. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 68

⁹⁷ vgl. auch: Hartung Lehr- und Handbuch der angewandten Statistik, 1999. S. 289 ff

⁹⁸ Baumert, Weiß: Föderalismus und Gleichwertigkeit der Lebensverhältnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 51

Land	Hauptschule	Real- schule	Gymnasium	Integrierte Gesamtschule	Berufliche Schule	Sonder- schule
Baden-Württemberg	26,5	31	28,9	0,5	9,3	3,9
Bayern	28,7	26,6	26,6	0,5	14,1	3,6
Hessen	15,2	28,4	31,4	16,4	5,2	3,4
Niedersachsen	30,4	32,5	24,8	3,8	3,9	4,7
Nordrhein-Westfalen	24,6	26,2	30	14,9		4,4
Rheinland-Pfalz	32,8	26,2	25,6	3,4	8,6	3,4
Saarland	18,1	33,5	28,5	16,5		3,3
Schleswig-Holstein	29,5	29,5	26,2	6,8	4,1	3,9
Alte Länder	26,2	28,3	28,1	7,3	6,1	4
Brandenburg		13,9	28,8	52,2		5
Mecklenburg-Vorpommern	10,4	49,4	25,7	5,7	3,3	5,4
Sachsen	9,2	55	27,7	0,3	3,1	4,8
Sachsen-Anhalt	6,2	54,8	28,3	1	3,6	6,1
Thüringen	16,5	43,5	26,7	2	6	5,4
Neue Länder	8,4	44,3	27,6	11,4	3,1	5,3
Berlin	10,6	21,2	33,8	30,7		3,7
Bremen	24,7	25,1	29,6	15,4		5,2
Hamburg	11,7	18,9	31,8	25,9	6,3	5,5
Stadtstaaten	12,3	21	32,8	27,9	1,7	4,3

Abbildung 44 Bildungsbeteiligung der 15-Jährigen PISA Testteilnehmer (Quelle: PISA)

*Je unterschiedlicher Schulen hinsichtlich der Schulleistungen sind, desto mehr Schulen müssen bei Ziehung der gleichen Anzahl von Schülern in den Schulen in die Untersuchung einbezogen werden, um eine vorgegebene Messgenauigkeit zu erreichen.*⁹⁹

Bei vergleichbaren Populationsanteilen wurden daher in Nordrhein-Westfalen 25 Gesamtschulen und in Hamburg und Berlin jeweils 25 Hauptschulen getestet. In Bayern und Baden-Württemberg hingegen wurden trotz der zu erwartenden Leistungsheterogenität zwischen einzelnen Berufsschulen lediglich drei beziehungsweise zwei Berufsschulen in der Stichprobe gezogen, wobei letztlich nur jeweils zwei Berufsschulen im Datensatz berücksichtigt wurden.¹⁰⁰

Es wurde (...) entschieden, die Gruppe der 15-Jährigen an diesen Schulen in geringerem Umfang in die Stichprobe einzubeziehen (Undersampling) und größere Standardfehler in Kauf zu nehmen. Die vergrößerten Standardfehler gehen in die

⁹⁹ Baumert, Artelt, Carstensen, Sibberns, Stanat: Untersuchungsgegenstand, Fragestellungen und technische Grundlagen der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 21

¹⁰⁰ Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Weiß: Schau genau: Wie begründet sind die Zweifel an den PISA-Ergebnissen aus Bayern? Eine Replik auf den gleichnamigen Artikel von Klaus Klemm in der Süddeutschen Zeitung, vom 3. 9. 2002. S. 1

*Schätzung der Standardfehler der Mittelwerte der Gesamtpopulation, aber auch von Subpopulationen ein. Ein größerer Standardfehler zeigt geringere Schätzgenauigkeit, aber keine systematische Verzerrung in irgend eine Richtung an.*¹⁰¹

Weniger die Schätzgenauigkeit, etwa im Sinne eines besonders geringen Standardfehlers, ist Gegenstand des Interesses, vielmehr steht ein systematischer Fehler (bias) im Vordergrund, der bei großer Heterogenität einer unterschiedlich leistungsselektierten Population der Berufsschüler bei Wertung von nur zwei Schulen in Betracht gezogen werden muss. Wie stark die Leistungsunterschiede im Segment der Berufsschulen ausfallen, bleibt offen.

*Wie sehr sich die 15-Jährigen an beruflichen Schulen in den einzelnen Ländern von Schule zu Schule unterscheiden, ist unbekannt.*¹⁰²

Im Nachhinein ist es dann nicht mehr möglich, über Gewichtungen eventuelle systematische Fehler zu korrigieren.

*Es gibt keine Möglichkeit, dies durch Gewichtung auszugleichen, da keine Daten über Verteilungen in der Grundgesamtheit der 15-Jährigen an beruflichen Schulen verfügbar sind.*¹⁰³

Robitzsch diskutiert das Problem systematischer Fehler in Schulleistungsuntersuchungen (positive Verzerrung) im Falle systematischen Ausfalls von Schulen (not missing at random) unter Verwendung von Modellrechnungen.¹⁰⁴ Nach der von Robitzsch entwickelten erweiterten Korrekturformel reduzieren sich die Testwerte bei systematischem Ausfall von Schulen.

¹⁰¹ Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann: Schau genau: Wie begründet sind die Zweifel an den PISA-Ergebnissen aus Bayern? Eine Replik auf den gleichnamigen Artikel von Klaus Klemm in der Süddeutschen Zeitung vom 3. 9. 2002

¹⁰² Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Schau genau: Wie begründet sind die Zweifel an den PISA-Ergebnissen aus Bayern? Eine Replik auf den gleichnamigen Artikel von Klaus Klemm in der Süddeutschen Zeitung vom 3. 9. 2002. S. 6

¹⁰³ Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Weiß Schau genau: Wie begründet sind die Zweifel an den PISA-Ergebnissen aus Bayern? Eine Replik auf den gleichnamigen Artikel von Klaus Klemm in der Süddeutschen Zeitung vom 3. 9. 2002. S. 6

¹⁰⁴ Robitzsch, A.: Modellierung von Stichprobenausfall in PISA In: Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Schau genau: Wie begründet sind die Zweifel an den PISA-Ergebnissen aus Bayern? Eine Replik auf den gleichnamigen Artikel von Klaus Klemm in der Süddeutschen Zeitung vom 3. 9. 2002. S. 30

Robitzsch spricht bereits zwei Probleme seines Verfahrens an: bimodale und schiefe Verteilungen. Darüber hinaus wären mehrfach geschichtete Klumpen-Stichproben zu berücksichtigen.

An dieser Stelle sollen lediglich Vergleiche unter anderem mit externen Daten angestellt werden. Sollten systematische Verzerrungen der Testergebnisse in denjenigen Bundesländern, die einen erheblichen Teil ihrer 15-jährigen Schülerinnen und Schüler in Berufsschulen unterrichten, vorliegen, könnte sich dies als Wechselwirkung zwischen Anteil der Berufsschüler und weiteren Variablen, insbesondere den Testwerten zeigen.

Ein Nachweis über die Testwerte lässt sich aufgrund der bundesland-spezifischen Standardisierung nicht erbringen. Die bundesland-spezifische Standardisierung im Public Use File wurde aus grundsätzlichen Erwägungen eingeführt¹⁰⁵ und lässt den direkten Vergleich im Rahmen dieser Evaluation nicht zu. Es ließe sich jedoch über die Problematik der so genannten *Nichtleser* eine Analyse führen.

3.2 Nichtleser und Testleistung

Artelt, Stanat, Schneider und Schiefele untersuchen den Zusammenhang zwischen Lesegewohnheit und Testleistung im Lesen. Die nationale Ergänzungsstudie, PISA-E, berichtet einen Zusammenhang zwischen dem Anteil derjenigen, die nicht aus Vergnügen lesen (in der Studie als *Nichtleser* definiert) und mittlerer *Lesekompetenz*,¹⁰⁶ wie folgende Abbildung zeigt.

¹⁰⁵ Information: KMK vom 19. August 2003

¹⁰⁶ Artelt, Schneider und Schiefele: Ländervergleich zur Lesekompetenz. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 79

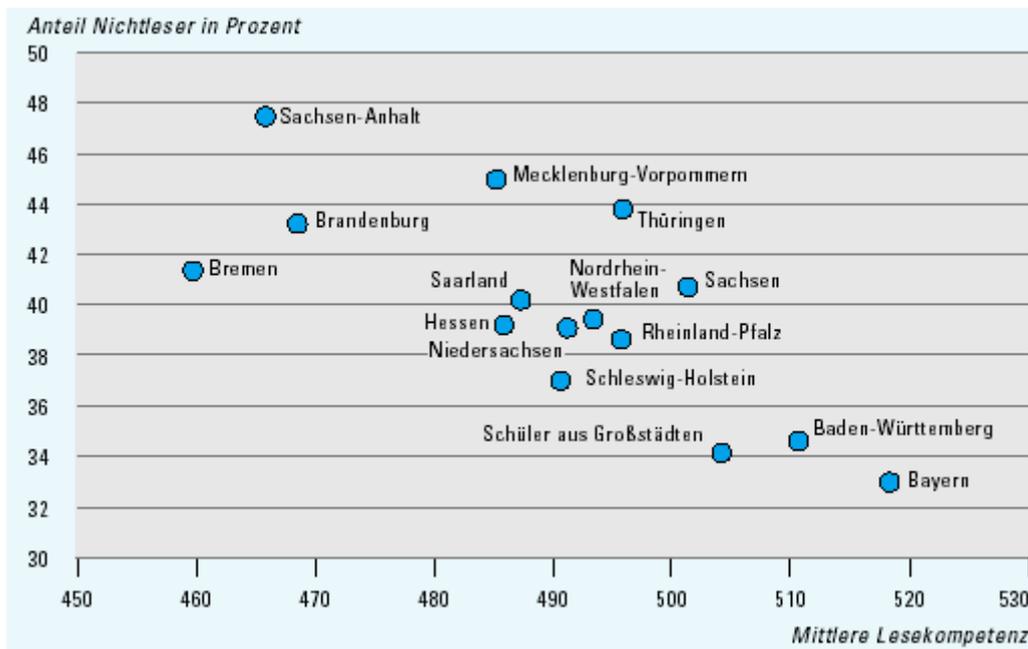


Abbildung 45 Anteil Nichtleser und mittlere *Lesekompetenz* (Quelle: PISA)

Erkennbar ist der negative Zusammenhang zwischen *Lesekompetenz* und Anteil der *Nichtleser*, wobei Bayern, Baden-Württemberg und Schüler aus Großstädten besonders niedrige Anteile an *Nichtlesern* bei gleichzeitig hoher mittlerer *Lesekompetenz* aufweisen, die Bundesländer Bremen, Brandenburg und Sachsen-Anhalt hingegen bei niedriger mittlerer *Lesekompetenz* einen hohen Anteil an *Nichtlesern*. Der Zusammenhang wird mit $r = -0,66$ berichtet.¹⁰⁷

Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Weiß berichten den Anteil der *Nichtleser* auch im internationalen Vergleich.¹⁰⁸

¹⁰⁷ Artelt, Schneider und Schiefele: Ländervergleich zur Lesekompetenz. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 80

¹⁰⁸ Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Weiß: PISA 2000 Zusammenfassung zentraler Befunde, S.17

Prozentualer Anteil von Schülerinnen und Schülern, die angeben, sie würden nicht zum Vergnügen lesen

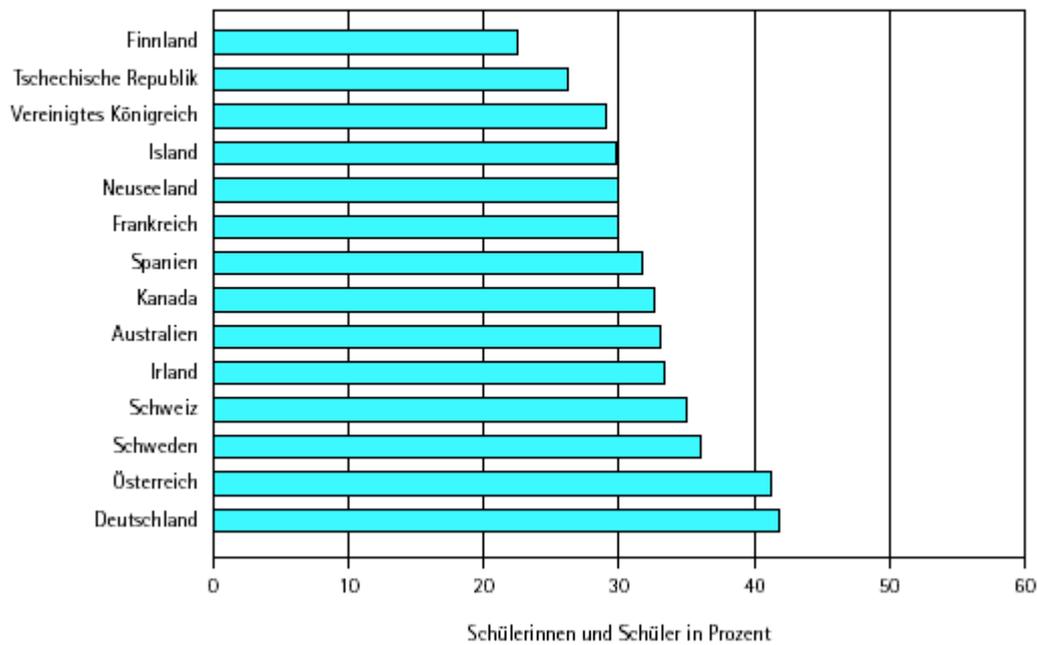


Abbildung 46 Nichtleser im internationalen Vergleich (Quelle: PISA)

Finnland weist in dieser Abbildung den geringsten Anteil *Nichtleser* auf, Deutschland hingegen mit über 40% den größten.

Der Technical Report der internationalen PISA-Studie weist mit der Variable ST34Q01 dasjenige Item auf, mit dem das tägliche Lesen entsprechend Codierung der *Nichtleser* abgefragt wird.¹⁰⁹ Die Stufe 1 (I do not read for enjoyment) entspricht der Codierung der so genannten *Nichtleser*. Die folgende Darstellung berechnet sich aus dem erweiterten Datensatz.

¹⁰⁹ Read each day – Q34 ST34Q01

- 1 I do not read for enjoyment
- 2 30 minutes or less each day
- 3 More than 30 minutes to less than 60 minutes each day
- 4 1 to 2 hours each day
- 5 More than 2 hours each day

Quelle: Adams, Wu (Hrsg.): PISA 2000 Technical Report. S. 300

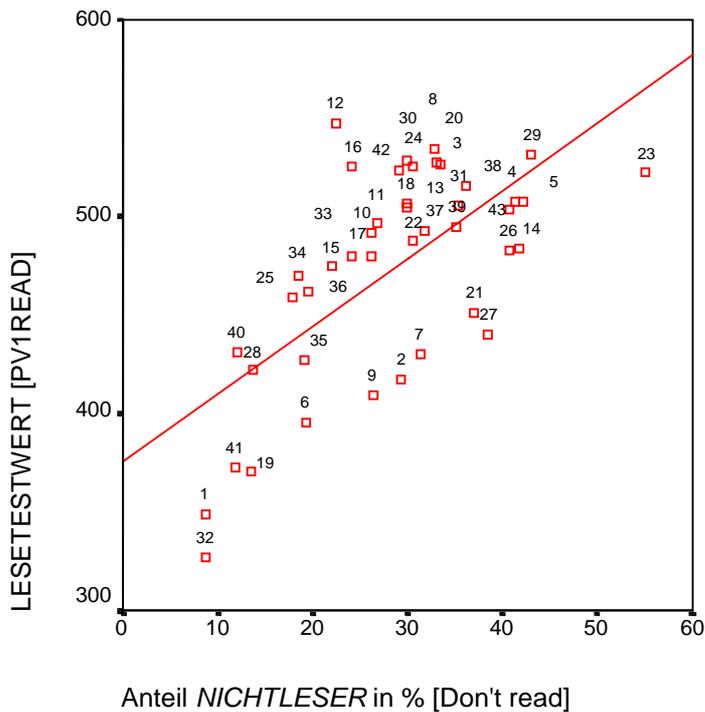
CASE NO.	Read each day - Q34					GRADE	HISEI	PV1READ	
	Don't read	30 min or less	31- 60 min	1-2 hours	More than 2 hours				
	percent								
ALBANIA	1	8,63	19,10	30,63	27,63	14,02	9,71	45,55	349,02
ARGENTINA	2	29,28	30,81	21,50	11,37	7,03	9,59	43,35	417,41
AUSTRALIA	3	33,10	30,47	20,54	11,80	4,09	10,10	52,25	527,79
AUSTRIA	4	41,31	28,60	18,01	8,98	3,10	9,43	49,72	507,53
BELGIUM	5	42,16	24,74	21,42	9,08	2,60	9,61	48,95	507,31
BRAZIL	6	19,27	21,30	31,38	16,81	11,23	8,50	43,93	395,20
BULGARIA	7	31,33	18,95	20,38	19,56	9,76	9,11	49,99	430,54
CANADA	8	32,72	33,73	20,37	9,62	3,56	9,84	52,83	534,49
CHILE	9	26,27	27,84	26,86	12,21	6,83	9,48	39,87	409,64
CZECH REPUBLIC	10	26,23	29,70	25,70	12,87	5,50	9,51	48,31	491,94
DENMARK	11	26,75	36,09	23,32	9,43	4,41	8,97	49,73	496,56
FINLAND	12	22,38	29,09	26,27	18,17	4,09	8,89	50,00	546,99
FRANCE	13	29,97	27,52	28,58	10,57	3,37	9,51	48,27	504,38
GERMANY	14	41,76	26,96	17,90	8,71	4,67	9,06	48,85	483,99
GREECE	15	21,99	26,61	22,68	19,98	8,74	10,05	47,76	474,46
HONG KONG	16	24,10	35,76	23,21	11,52	5,42	9,83	42,25	525,15
HUNGARY	17	26,08	28,19	24,19	13,44	8,10	9,24	49,53	479,83
ICELAND	18	29,82	37,92	22,44	6,93	2,89	10,00	52,73	506,58
INDONESIA	19	13,45	36,04	26,02	16,72	7,76	9,51	36,38	370,76
IRELAND	20	33,37	30,86	20,41	11,57	3,79	9,50	48,43	526,68
ISRAEL	21	37,04	20,19	19,09	16,25	7,42	9,98	55,30	450,96
ITALY	22	30,61	30,18	22,55	13,00	3,66	9,86	47,08	487,36
JAPAN	23	55,03	17,84	15,42	8,18	3,53	10,00	50,54	522,03
KOREA, REPUBLIC OF	24	30,55	29,58	21,91	12,00	5,96	10,00	42,80	525,03
LATVIA	25	17,86	25,61	29,59	19,66	7,28	9,38	50,15	459,14
LIECHTENSTEIN	26	40,63	33,75	16,56	5,00	4,06	8,85	47,46	482,58
LUXEMBOURG	27	38,42	25,56	19,56	11,93	4,54	9,05	44,79	440,40
MEXICO	28	13,60	43,70	27,19	11,51	4,00	9,36	42,48	421,74
NETHERLANDS	29	43,00	31,57	16,72	5,80	2,91	9,43	50,85	531,64
NEW ZEALAND	30	29,92	36,61	19,37	10,44	3,65	10,98	52,20	528,60
NORWAY	31	35,38	34,57	20,16	7,72	2,17	10,00	53,91	505,71
PERU	32	8,71	23,64	32,89	21,34	13,42	9,31	40,42	327,23
POLAND	33	24,17	22,71	28,72	16,45	7,96	9,00	46,03	480,01
PORTUGAL	34	18,41	39,00	26,53	12,35	3,71	9,25	43,85	469,89
ROMANIA	35	19,01	27,53	22,09	19,70	11,67	8,89	45,53	427,19
RUSSIAN FEDERATION	36	19,44	24,62	25,82	17,40	12,72	9,69	49,38	461,91
SPAIN	37	31,77	32,92	24,15	8,75	2,41	9,70	44,99	492,84
SWEDEN	38	36,04	30,78	21,07	8,75	3,36	8,98	50,57	515,39
SWITZERLAND	39	35,12	33,02	20,58	8,33	2,96	8,93	49,21	494,80
THAILAND	40	11,95	39,78	25,87	16,08	6,32	9,38	33,02	431,18
MACEDONIA	41	11,80	19,89	29,61	24,10	14,59	9,25	46,79	327,96
UNITED KINGDOM	42	29,10	35,65	22,91	9,41	2,99	10,68	51,26	523,54
UNITED STATES	43	40,71	31,17	16,21	8,05	3,86	9,53	52,40	503,71

*) DATABASE: INSTUD_READ (WITH CONSID. OF. W_FSTUWT - STUD. POP. WEIGHT -)
download <http://pisaweb.acer.edu.au/oeed>

Abbildung 47 NICHTLESER

Japan und die Niederlande weisen den höchsten Anteil an *Nichtlesern* auf. Brasilien, Portugal, Lettland und Mexiko weisen den geringsten Anteil an *Nichtlesern* in der in 2000/2001 publizierten Studie auf. Hinzu kommen die Länder Indonesien, Thailand, Mazedonien, Peru und Albanien, die noch weniger *Nichtleser* aufweisen.¹¹⁰ Bemerkenswert ist, dass in der internationalen Untersuchung ein zum Datensatz der Ergänzungsstudie (PISA-E) umgekehrter Zusammenhang zwischen Testleistung und Anzahl der *Nichtleser* festzustellen ist, wie folgende Abbildung zeigt.

¹¹⁰ Ergänzung im PUBLIC-USE-FILE download <http://pisaweb.acer.edu.au/oeed>



Korrelationen		PV1READ	HISEI
Don't read	Korrelation nach Pearson	,662**	,572**
	Signifikanz (2-seitig)	,000	,000
	N	43	43

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Korrelationen		PV1READ	HISEI
Spearman-Rho	Don't read	,579**	,540**
	Korrelationskoeffizient		
	Sig. (2-seitig)	,000	,000
	N	43	43

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

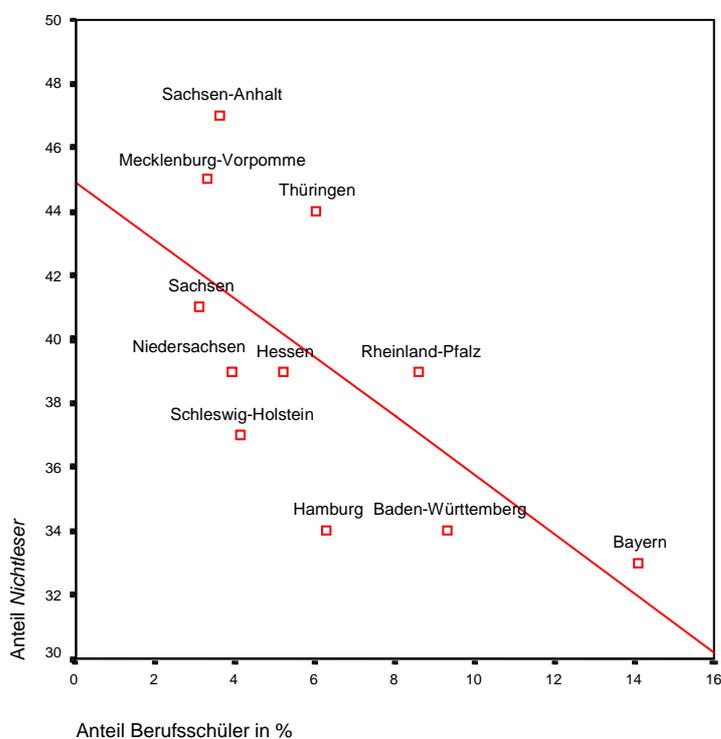
Abbildung 48 *Nichtleser* - Testwert Lesen International/Korrelation

Den größten Anteil *Nichtleser* weisen die japanischen Schüler auf, bei gleichzeitig hohem Lesetestwert. Den geringsten Anteil *Nichtleser* weisen in der in 2000/2001 publizierte Studie die mexikanischen Schüler auf. Sie erreichen den niedrigsten Lesetestwert der damals publizierten Teilnehmerstaaten. Betrachtet man den erweiterten Datensatz, so weisen die Länder Indonesien, Thailand, Mazedonien, Peru und Albanien noch niedrigere Anteile an *Nichtlesern* auf mit Testwerten im Lesen, die zum Teil unterhalb denen der mexikanischen Schülerinnen und Schülern liegen. Mit $r_p = 0,662$ $r_s = 0,579$ besteht ein positiver linearer statistischer Zusammenhang zwischen Anteil der *Nichtleser* und Lesetestleistung. Das bedeutet: Je mehr *Nichtleser*, desto höher die Testleistung im Lesen.

Prinzipiell stellen sich nun zwei Fragen:

1. Wie lassen sich die hohen Testergebnisse mit geringer Bereitschaft zum Lesen vereinbaren ?
2. Zusätzlich stellt sich die Frage nach der Diskrepanz zwischen der PISA Ergänzungsstudie (PISA-E) und der internationalen Studie (PISA-I) hinsichtlich der Richtung des Zusammenhangs.

Die Klärung zunächst der zweiten Frage liegt in der Ermittlung einer so genannten Dritten Variablen, welche mit den übrigen kovariert (Scheinkorrelation). Im Anteil der Berufsschüler je Bundesland ließe sich eine solche finden.



Korrelationen

		LESEN
SCHUL	Korrelation nach Pearson	-,664*
	Signifikanz (2-seitig)	,026
	N	11

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Korrelationen

		LESEN
Spearman-Rho	SCHUL Korrelationskoeffizient	-,754*
	Sig. (2-seitig)	,007
	N	11

**.. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 49 Nichtleser Berufsschüler National/Korrelation

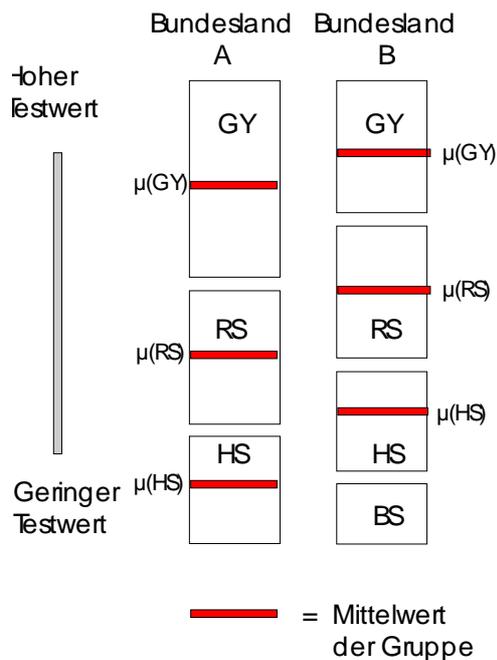
Es zeigt sich ein negativer, linearer Zusammenhang zwischen Anteil der *Nichtleser* und prozentualem Anteil der Berufsschüler des Jahrgangs eines Bundeslandes: Je mehr Berufsschüler ein Bundesland aufweist, desto weniger *Nichtleser* finden sich in den

Daten. Die im Lesetest eher schwächeren Berufsschüler wurden jedoch nicht entsprechend ihres Anteils in der Grundgesamtheit berücksichtigt. Daher liegt das Fehlen dieser Gruppe als Interpretation nahe (Scheinkorrelation).

Japanische und koreanische Schülerinnen und Schüler sind neben ihrem regulären Unterricht und ihrem privaten Zusatzunterricht von erheblichem Umfang kaum noch bereit, freiwillig zu lesen. Mexikanische, brasilianische, indonesische, thailändische, mazedonische, peruanische und albanische Schülerinnen und Schüler hingegen lesen freiwillig, insofern sich für sie eine Gelegenheit bietet. Das Ausmaß des Mangels an freiwilligem Lesen kovariiert mit der Höhe des sozialen Index eines Landes ($r_p = 0,572$ und $r_s = 0,54$) signifikant. Je mehr für die Schule gelesen werden muss, desto geringer das Vergnügen.

3.3 Stage Migration

Der so genannte Scheinzusammenhang beziehungsweise die so genannte *Scheinkorrelation* führt zum Begriff *Stage Migration*. Das Prinzip basiert auf der Gruppierung von Daten. Statistische Kennziffern, wie die des arithmetischen Mittelwertes, lassen sich verändern (wahlweise verbessern oder verschlechtern), indem Daten gruppiert werden. Die folgende Abbildung veranschaulicht diesen Sachverhalt.



Stage Migration:

Angenommen sei eine insgesamt identische Testleistungsverteilung in Bundesland A und B entsprechend nebenstehender Abbildung. Bundesland A setzt die untere (Testleistungs-) Grenze für GY tiefer als Bundesland B. Dadurch verringert sich der mittlere Wert (μ) in GY und jedem folgenden Segment (RS, HS), jedoch ohne dass die Gesamtestleistung von Bundesland A geringer wäre als die von Bundesland B. Bundesland B senkt die untere Grenze von GY nicht und weist dadurch in den Segmenten GY, RS und HS höhere mittlere Werte auf (ohne dass sich – wie eingangs definiert – die Gesamtestleistung von Bundesland A unterscheidet). Ferner weist es im unteren Testleistungssegment eine weitere Kategorie auf (BS). Bleibt diese bei der Mittelwertbildung unberücksichtigt, erhöht sich auch der gesamte Mittelwert von Bundesland B.

Abbildung 50 Stage Migration

Stage Migration ist eine Erscheinung gruppierter Daten, wenn - wie im Beispiel - die Gruppierungskriterien unterschiedlich definiert sind beziehungsweise Inkohärentes zusammengefasst wird. Stage Migration betrifft ebenfalls den Vergleich der Testleistungsspitzen, sofern Quantile gebildet werden, da diese sich an der absoluten

Anzahl der zu messenden Einheiten (z.B. Schüler) orientieren. Das bedeutet, dass auch bei Vergleichen von Lageparametern oberer Testleistungsbereiche einzelner Systeme bei gleicher Gesamtleistung (Testleistungsverteilung) schon aus rein technischen Gründen unterschiedliche mittlere Werte auftreten können.

Folgendes Beispiel soll diesen Sachverhalt noch einmal verdeutlichen:¹¹¹

	2002-03	2003-04
Laura	100	90
James	90	80
Felipe	80	70
Kisha	70	65
Jose	60	55
Raul	20	-
Mean	70	72

Abbildung 51 Stage Migration

Für jede Person einer Gruppe von Probanden wird zu zwei verschiedenen Zeitpunkten t_0 (2002-03) und t_1 (2003-04) je ein personenbezogener Wert gemessen. Sämtliche Meßwerte aus t_0 reduzieren sich in t_1 . Da jedoch diejenige Person mit dem geringsten Meßwert aus t_0 (Raul) in t_1 nicht mehr gemessen wird, erhöht sich der Mittelwert der Gruppe von $\bar{x} = 70$ in t_0 auf $\bar{x} = 72$ in t_1 .

Baumert hatte in diesem Zusammenhang die Bildungsexpansion angesprochen.

Die konsistent über das gesamte Leistungsspektrum hinweg auftretenden Länderunterschiede, die sich in allen untersuchten mathematisch-naturwissenschaftlichen Fächern nachweisen lassen, legen die Frage nahe, ob die Leistungsunterschiede möglicherweise auf unterschiedliche Strukturentwicklungen des Schulsystems in beiden Ländern zurückzuführen seien - etwa auf unterschiedliche Expansionsraten der weiterführenden Schulen. Diese Vermutung wird immer wieder vorgetragen. Allerdings fehlen überzeugende empirische Belege.^{112 113}

¹¹¹ Darling-Hammond: From „Seperate but Equal“ to „No Child Left Behind“: The Collision of New Standards and Old Inequalities. In: Meier, Kohn, Darling-Hammond, Sizer, Wood: Many Children Left Behind. Boston 2004 S.19

¹¹² Baumert, J: TIMSS-Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. In: List: TIMSS Mathematische-naturwissenschaftliche Kenntnisse deutscher Schüler auf dem Prüfstand. S. 57

In folgender Abbildung wird der Anteil Schüler in Gymnasien und *mathematische Kompetenz* im PISA-Ergänzungstest dargestellt. Berichtet wird mit $r=-0,54$ ein negativer, linearer Zusammenhang zwischen beiden Variablen. Je höher der Anteil der Schüler in Gymnasien des Bundeslandes, desto geringer die mittlere *mathematische Kompetenz*.¹¹⁴

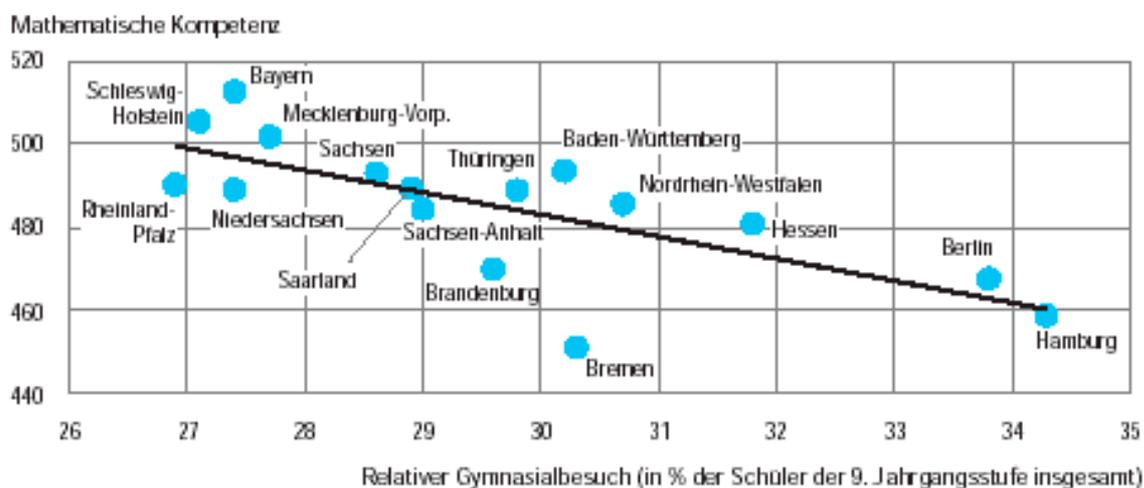


Abbildung 52 Schüler in Gymnasien – *Mathematische Kompetenz* (Quelle: PISA)

In der folgenden Darstellung der so genannten Perzentilbänder zeigt sich, dass das Bundesland Bremen einerseits den geringsten mittleren Wert sämtlicher Bundesländer aufweist, andererseits die höchste Streuungsbandbreite. Die oberen Testleistungswerte hingegen erreichen das Niveau der Testleistungswerte sämtlicher Bundesländer.

¹¹³ Baumert schlägt in diesem Zusammenhang einen Vergleich von TIMSS mit Daten aus einer MPIB Untersuchung von 1968/69 vor.

¹¹⁴ Neubrand, Klieme: Mathematische Grundbildung. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 123ff

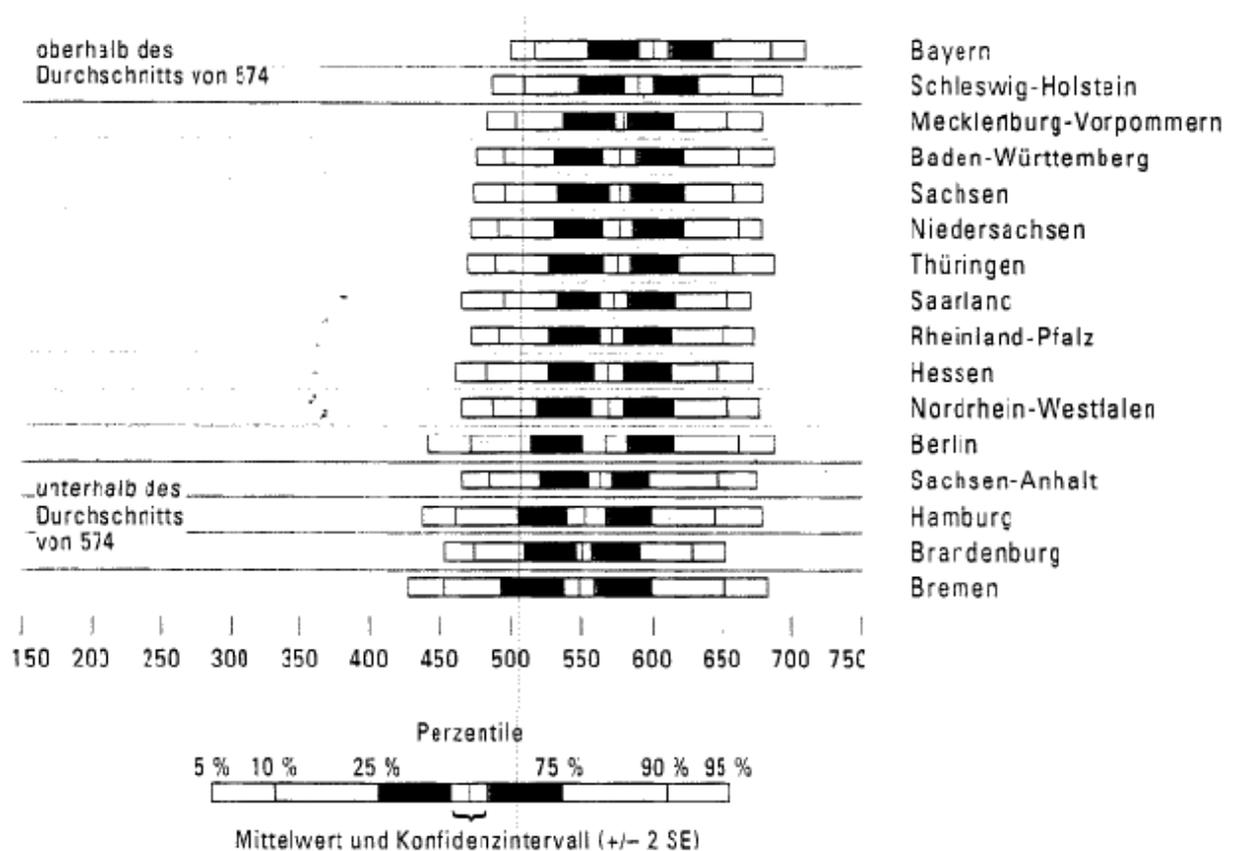
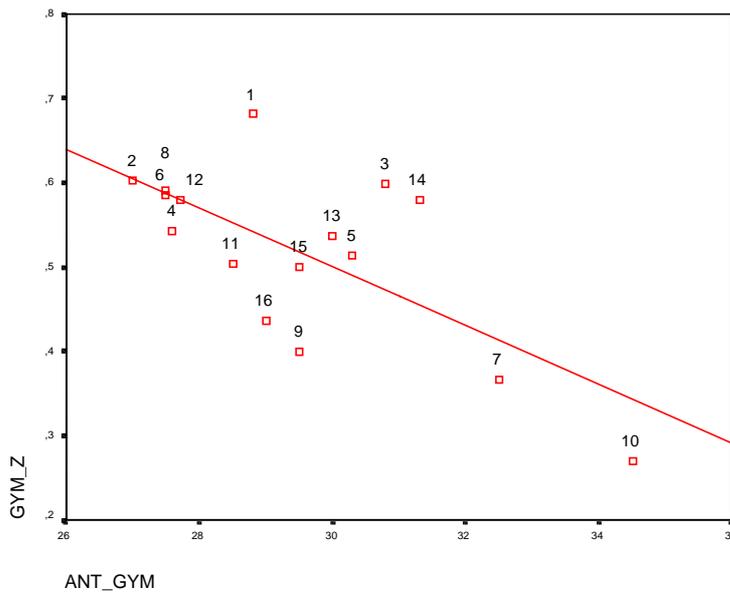


Abbildung 53 Perzentilbänder Testwert Lesen Gymnasien (Quelle: PISA)

Betrachtet man nun in der folgenden Abbildung den durchschnittlichen relativen höchsten sozioökonomischen Index¹¹⁵ der Gymnasiasten der einzelnen Bundesländer (GYM_Z) und trägt diesen gegen den Anteil der Schüler in den Gymnasien (ANT_GYM) ab, zeigt sich ein mit $r_p = -0,679$ $p=0,004$ beziehungsweise $r_s = -0,564$ $p=0,023$ statistisch signifikanter negativer, linearer Zusammenhang. Das bedeutet, je mehr Schülerinnen und Schüler die Gymnasien eines Bundeslandes proportional aufweisen, desto niedriger der relative durchschnittliche sozioökonomische Index. Dieser Zusammenhang wäre als Konsequenz einer Bildungsexpansion zu interpretieren. Berlin weist im Datensatz die höchste Expansion des Gymnasiums auf.

¹¹⁵ Ein Vergleich macht die z-Standardisierung je Bundesland notwendig, da der sozioökonomische Standard über die Bundesländer systematisch variiert.



- 1 Saarland
- 2 Rheinland-Pfalz
- 3 Nordrhein-Westfalen
- 4 Niedersachsen
- 5 Bremen
- 6 Schleswig-Holstein
- 7 Hamburg
- 8 Mecklenburg-Vorpommern
- 9 Brandenburg
- 10 Berlin
- 11 Sachsen
- 12 Bayern
- 13 Baden-Württemberg
- 14 Hessen
- 15 Thüringen
- 16 Sachsen-Anhalt

Correlations

		GYM_Z
ANT_GYM	Pearson Correlation	-,679**
	Sig. (2-tailed)	,004
	N	16

** . Correlation is significant at the 0.01 level

Correlations

			GYM_Z
Spearman's rho	ANT_GYM	Correlation Coefficient	-,564*
		Sig. (2-tailed)	,023
		N	16

* . Correlation is significant at the .05 level (2-tailed).

Abbildung 54 Anteil Gymnasiasten – relativer sozioökon. Index/Korrelation

Die Darstellung der Perzentilbänder legt nun die Interpretation nahe, dass mit der Bildungsexpansion kein Leistungsabfall im oberen Testleistungsbereich zwingend notwendig ist beziehungsweise zumindest anhand der PISA-Daten nicht nachgewiesen werden kann. Lediglich das arithmetische Mittel als Lageparameter sinkt. Letzteres ist das Kennzeichen so genannter Stage Migration. Die PISA 2000 Ergänzungsstudie¹¹⁶ bestätigt diesen Sachverhalt.

¹¹⁶ Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Stanat, Tillmann, Weiß (Hrsg.): PISA 2000 Zusammenfassung zentraler Befunde, 2003

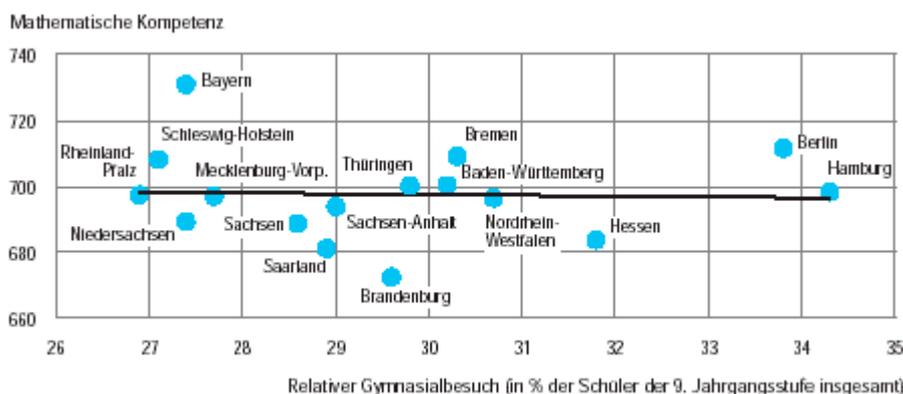


Abbildung 55 Zusammenhang zwischen relativem Gymnasialbesuch und Leistungsniveau der oberen 5 % der Gymnasiasten der 9. Jahrgangsstufe in Mathematik auf Länderebene (Mittelwerte) (Quelle: PISA 2000 Zusammenfassung zentraler Befunde, S.68)

Berlin und Bremen weisen nach Bayern hier die höchsten mittleren Werte auf. Die Mittelwerte der höchsten 5% Testwerte der Bundesländer weisen keinen negativen Zusammenhang zum relativen Gymnasialbesuch mehr auf, lediglich der Mittelwert der gesamten Schüler des Gymnasiums eines Bundeslandes kovariiert negativ mit zunehmendem Gymnasialbesuch.¹¹⁷

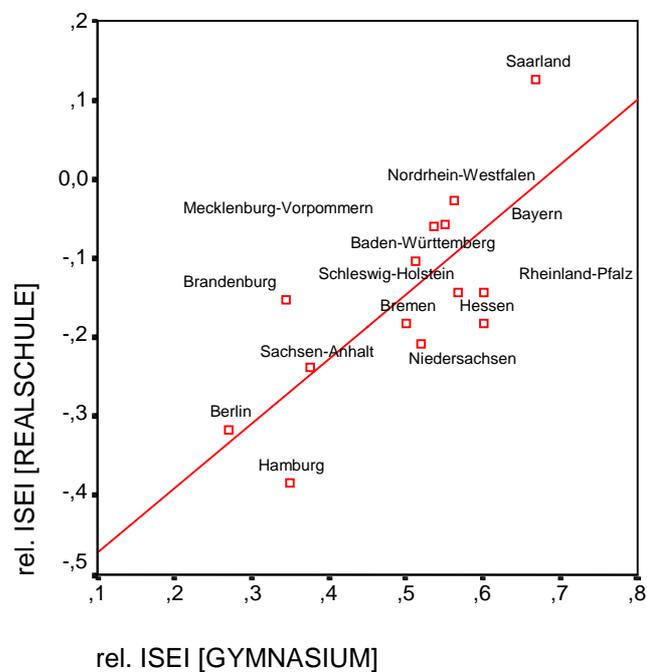
Die folgende Abbildung zeigt den Zusammenhang von durchschnittlichem (je Bundesland) z-standardisierten sozioökonomischen Index der Schülerschaft aus Gymnasien und Realschulen auf der Datenbasis der 31.751 15-Jährigen der 9. Klassenstufe (PISA 2000-E). Der positive lineare Zusammenhang der sozioökonomischen Niveaus der jeweiligen Realschulen und Gymnasien ist statistisch signifikant ($p < 0,01$) und mit $r_p = 0,757$ ($r_s = 0,675$)¹¹⁸ relativ hoch. Dieser Zusammenhang zeigt die unterschiedliche Expansion des Gymnasiums und ihre Auswirkung für die Schulform Realschule. Je geringer die Expansion des Gymnasiums, desto höher ist nach vorliegenden Daten das durchschnittliche sozioökonomische Niveau der Schulform Realschule. Zu berücksichtigen ist jeweils der unterschiedliche Anteil der Schulen mit mehreren Bildungsgängen.

¹¹⁷ (Prinzipiell wären jedoch auch die Mittelwerte dieses Segments Gegenstand der so genannten Stage Migration.)

¹¹⁸ Berechnung über den ungewichteten Datensatz

Bundesland	Realschule (rel. ISEI) *	Gymnasium (rel. ISEI) *	Schule mit mehreren Bildungsgängen		Realschule	Integrierte Gesamtschule	Gymnasium	Gesamt	
			Hauptschule	**					
Saarland	0,13	0,67	321	**	594	307	245	577	2044
Rheinland-Pfalz	-0,14	0,60	565	**	143	631	134	600	2073
Nordrhein-Westfalen	-0,03	0,56	485	**	0	640	564	608	2297
Niedersachsen	-0,21	0,52	492	**	6	588	120	624	1830
Bremen	-0,18	0,50	433	**	0	560	202	564	1759
Schleswig-Holstein	-0,14	0,57	459	**	0	603	375	635	2072
Hamburg	-0,38	0,35	0	**	0	535	0	561	1096
Mecklenburg-Vorpommern	-0,06	0,54	19	**	1144	657	136	678	2634
Brandenburg	-0,15	0,34	0	**	0	675	665	715	2055
Berlin	-0,32	0,27	0	**	0	517	0	589	1106
Sachsen		0,43	0	**	1855	0	0	681	2536
Bayern	-0,06	0,55	639	**	0	547	27	602	1815
Baden-Württemberg	-0,10	0,51	507	**	0	626	17	637	1787
Hessen	-0,18	0,60	477	**	0	576	673	608	2334
Thüringen		0,44	0	**	1969	0	28	717	2714
Sachsen-Anhalt	-0,24	0,38	258	**	0	635	19	687	1599
TOTAL			4655	**	5711	8097	3205	10083	31751

* ISEI standardisiert je Bundesland
 Datensatz: 15-Jährige in Klassenstufe 9 (Datensatz ungewichtet)
 ** Anzahl (N)



Korrelationen

		GYM
REAL	Korrelation nach Pearson	,757**
	Signifikanz (2-seitig)	,002
	N	14

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Korrelationen

		GYM
Spearman-Rho	REAL	,675**
	Korrelationskoeffizient	
	Sig. (2-seitig)	,008
	N	14

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 56 relatives sozioökonomisches Niveau Realschule/Gymnasium

3.4 Migrationshintergrund

Baumert und Weiß machen in der PISA-Ergänzungsstudie auf die Verteilung von Schülern mit Migrationshintergrund über einzelne Bundesländer und Schulformen aufmerksam.

*In Bremen besuchen rund 25 Prozent der 15-Jährigen einen Hauptschulbildungsgang; dies ist eine Größenordnung, die wir auch in Flächenländern antreffen. In Bremen ist allerdings der Anteil an Jugendlichen mit Migrationshintergrund in diesem Bildungsgang besonders hoch.*¹¹⁹

Anteil und *Verteilung* sind nicht gleichzusetzen. Da Bremen ebenso wie Hamburg oder Berlin relativ hohe *Anteile* an ausländischen Schülern (entsprechend Definition des Statistischen Bundesamtes) besitzen, finden sich entsprechend hohe *Anteile* von ausländischen Schülern in sämtlichen Schulformen, u.a. auch in der Hauptschule. Die Frage ist jedoch nicht, ob deren Anteil besonders hoch ist, sondern ob deren Verteilung eine überproportionale Häufung in der Schulform Hauptschule aufweist.

Da die Expansion des Gymnasiums in Bremen stärker ausgeprägt ist als etwa in den Bundesländern Saarland oder Bayern, wäre dieser von Baumert und Weiß benannte Zusammenhang, insofern er sich auf eine überproportionale Häufung beziehen sollte, eher unwahrscheinlich.

Die nachfolgende Tabelle zeigt die *Verteilung* ausländischer Schüler über die einzelnen Schulen je Bundesland für das Jahr 2000.¹²⁰ Die Zellbesetzungen werden durch standardisierte Residuen (u_{ij}) auf Signifikanz geprüft. Erweitert man die Perspektive und betrachtet nicht nur die 15-jährigen Schülerinnen und Schüler der 9. Klassenstufe, so fällt auf, dass ausländische Schüler in Bremens Hauptschulen im Bundesvergleich *unterproportional* vertreten sind mit 1200 (16,4%),¹²¹ während Bayern hier einen überproportionalen Anteil von 36900 (40,2%) aufweist. Baden-Württemberg weist mit 49100 (34,6%) ebenfalls einen überproportionalen Anteil auf. Ein weiterer erheblicher Teil bayerischer und baden-württembergischer ausländischer Schüler findet sich überproportional im Segment der Berufsschulen (im dualen System). In den bayerischen und baden-württembergischen Gymnasien sind ausländische Schüler statistisch signifikant unterproportional vertreten (Bayern 10700 (11,7%), Baden-Württemberg

¹¹⁹ Baumert, Weiß: Föderalismus und Gleichwertigkeit der Lebensverhältnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 50

¹²⁰ Quelle: Statistisches Bundesamt „Ausländische Schüler an allgemein bildenden Schulen/Ausländische Schüler an beruflichen Schulen“ Länder 2000

¹²¹ Die Prozentuierung basiert auf insgesamt N = 710.750 Schülerinnen und Schülern einer Auswahl von Schultypen.

12500 (8,5%)). Bremen weist zu Bayern und Baden-Württemberg eine gegenläufige Tendenz auf. Ausländische Schülerinnen und Schüler finden sich hier signifikant überproportional in den Gymnasien (1200 (16,4%)), hingegen signifikant unterproportional in Haupt- und Berufsschulen (Hauptschule 1200 (16,4%), Berufsschule im dualen System 1200 (16,4%)). Diese Daten sind konform zu den nachgewiesenen unterschiedlichen Bildungsexpansionsraten der Bundesländer. Da ausländische Schülerinnen und Schüler in der Bundesrepublik generell niedrigere sozioökonomische Standards aufweisen als die deutsche Vergleichsgruppe, finden sie sich in denjenigen Bundesländern mit geringerer Bildungsexpansion weniger häufig in höheren Bildungseinrichtungen.

Bundesland * Schultyp Kreuztabelle

		Schultyp															Fachs- Berufs- akade- mien	Gesamt		
		Hauptschule	Schularten mit mehreren Bildungsgängen	Realschulen	Gymnasien	Integrierte Gesamt- schulen	Freie Waldorf- schulen	Abend- schulen und sonstige Kollegs	Sonder- schulen	Berufs- schulen im dualen System	Berufs- grund- bildungsjahr	Berufsaufbau schulen	Berufsfach- schulen	Berufs- Technische Oberschulen	Fachober- schulen	Fach- gymnasien			Fachschulen	
Bundesland	BW Baden-Württemberg	Anzahl	49100	0	16100	12500	700	600	1000	13100	28300	4000	100	12600	100	0	2900	900	0	142000
		% von Bundesland	34,6%	,0%	11,3%	8,8%	,5%	,4%	,7%	9,2%	19,9%	2,8%	,1%	8,9%	,1%	,0%	2,0%	,6%	,0%	100,0%
		Korrigierte Residuen	73,9	-43,8	-1,0	-45,9	-127,5	20,2	-20,7	5,7	25,1	1,9	20,0	57,4	-2,0	-38,5	58,6	-9,1	-11,2	
BY Bayern		Anzahl	36900	0	7300	10700	400	100	300	8100	21400	1200	0	1700	500	1700	0	900	500	91700
		% von Bundesland	40,2%	,0%	8,0%	11,7%	,4%	,1%	,3%	8,8%	23,3%	1,3%	,0%	1,9%	,5%	1,9%	,0%	1,0%	,5%	100,0%
		Korrigierte Residuen	98,4	-33,7	-35,2	-7,2	-98,9	-6,9	-26,9	-1	48,3	-28,5	-3,8	-53,9	51,5	36,6	-29,2	5,4	58,1	
BE Berlin-Ost Berlin-West		Anzahl	4600	0	3400	6100	6300	100	600	2100	3600	1300	0	2000	0	700	200	300	0	31300
		% von Bundesland	14,7%	,0%	10,9%	19,5%	20,1%	,3%	1,9%	6,7%	11,5%	4,2%	,0%	6,4%	,0%	2,2%	,6%	1,0%	,0%	100,0%
		Korrigierte Residuen	-49,5	-18,8	-3,1	39,0	67,7	4,6	10,9	-13,6	-29,2	15,6	-2,1	5,3	-5,3	28,0	-3,3	2,6	-4,8	
BB Brandenburg		Anzahl	0	0	0	400	800	30	10	100	100	0	0	0	0	0	0	0	0	1440
		% von Bundesland	,0%	,0%	,0%	27,8%	55,6%	2,1%	,7%	6,9%	6,9%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-23,0	-3,9	-13,6	17,7	60,6	15,8	-1,9	-2,5	-10,7	-6,4	-5	-9,3	-1,1	-3,5	-3,4	-3,5	-1,0	
HB Bremen		Anzahl	1200	0	1000	1200	600	10	300	600	300	0	0	800	0	100	0	0	0	7310
		% von Bundesland	16,4%	,0%	13,7%	16,4%	8,2%	,1%	4,1%	8,2%	16,4%	4,1%	,0%	10,9%	,0%	1,4%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-20,2	-8,9	6,1	10,5	-3,2	-1,3	22,1	-1,9	-2,8	7,2	-1,0	19,4	-2,5	5,1	-7,7	-7,9	-2,3	
HH Hamburg		Anzahl	4300	900	1800	5200	7300	100	200	2300	3000	2300	0	2400	0	200	400	300	0	30700
		% von Bundesland	14,0%	2,9%	5,9%	16,9%	23,8%	,3%	,7%	7,5%	9,8%	,0%	,0%	7,8%	,7%	1,3%	1,0%	1,0%	,0%	100,0%
		Korrigierte Residuen	-51,8	32,4	-31,3	24,7	89,5	4,8	-9,6	-8,5	-37,0	52,1	-2,1	16,2	-5,2	-3,5	10,1	2,9	-4,8	
HE Hessen		Anzahl	11200	0	12000	12600	10300	100	1400	6000	14900	2600	0	3400	0	1500	1000	700	0	77700
		% von Bundesland	14,4%	,0%	15,4%	16,2%	13,3%	,1%	1,8%	7,7%	19,2%	3,3%	,0%	4,4%	,0%	1,9%	1,3%	,9%	,0%	100,0%
		Korrigierte Residuen	-82,6	-30,7	37,5	34,2	40,5	-5,0	14,7	-11,6	11,8	10,9	-3,5	-17,0	-8,6	35,8	16,1	2,3	-7,8	
MV Mecklenburg-Vorpom- mern		Anzahl	100	100	300	300	100	0	0	0	0	0	0	0	0	0	0	0	0	900
		% von Bundesland	11,1%	11,1%	33,3%	33,3%	11,1%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-10,6	29,3	20,7	19,1	1,9	-1,4	-3,4	-9,3	-13,9	-5,0	-4	-7,4	-9	-2,7	-2,7	-2,7	-8	
NI Niedersachsen		Anzahl	9300	200	5300	4500	1900	100	100	6100	6400	2000	0	2800	0	400	500	400	0	40000
		% von Bundesland	23,3%	,5%	13,3%	11,3%	4,8%	,3%	,3%	15,3%	16,0%	5,0%	,0%	7,0%	,0%	1,0%	1,3%	1,0%	,0%	100,0%
		Korrigierte Residuen	-16,5	-11,4	11,9	-7,2	-32,1	2,0	-18,5	46,5	-8,9	28,4	-2,4	11,4	-6,0	3,9	10,3	3,9	-5,5	
NW Nordrhein-Westfalen		Anzahl	61100	0	28400	27500	35300	200	20300	38700	4300	0	12000	0	1000	0	2000	0	0	235400
		% von Bundesland	26,0%	,0%	12,1%	11,7%	15,0%	,1%	2,0%	8,6%	16,4%	1,8%	,0%	5,1%	,0%	,4%	,0%	,8%	,0%	100,0%
		Korrigierte Residuen	-11,4	-61,7	12,2	-12,8	117,0	-15,8	37,6	-4,6	-18,9	-33,3	-7,0	-15,7	-17,2	-26,5	-53,3	1,3	-15,7	
RP Rheinland-Pfalz		Anzahl	8600	1700	2900	3000	700	30	20	2200	4300	1100	0	1600	0	100	300	300	0	26850
		% von Bundesland	32,0%	6,3%	10,8%	11,2%	2,6%	,1%	,1%	8,2%	16,0%	4,1%	,0%	6,0%	,0%	,4%	1,1%	1,1%	,0%	100,0%
		Korrigierte Residuen	19,7	85,5	-3,2	-6,2	-38,4	-3,5	-17,7	-3,8	-7,2	13,8	-2,0	1,8	-4,9	-8,4	5,9	5,3	-4,4	
SL Saarland		Anzahl	200	2500	200	900	900	30	40	600	1400	300	0	400	0	200	0	0	0	7670
		% von Bundesland	2,6%	32,6%	2,6%	11,7%	11,7%	,4%	,5%	7,8%	18,3%	3,9%	,0%	5,2%	,0%	2,6%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-48,1	269,9	-24,4	-1,8	7,5	3,6	-5,8	-3,2	1,4	6,3	-1,0	-1,9	-2,6	17,3	-7,9	-8,1	-2,3	
SN Sachsen		Anzahl	0	1300	0	700	0	30	100	100	200	0	0	100	0	100	0	0	0	2630
		% von Bundesland	,0%	49,4%	,0%	26,6%	,0%	1,1%	3,8%	3,8%	7,6%	,0%	,0%	3,8%	,0%	3,8%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-31,1	241,6	-18,4	22,2	-16,4	10,6	11,8	-9,1	-13,5	-8,6	-6	-4,2	-1,5	-4,7	17,3	-4,7	-1,4	
ST Sachsen-Anhalt		Anzahl	40	300	200	400	10	0	10	100	0	0	0	100	0	0	0	0	0	1160
		% von Bundesland	3,4%	25,9%	17,2%	34,5%	,9%	,0%	,9%	8,6%	,0%	,0%	,0%	8,6%	,9%	,0%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-18,0	82,2	6,3	22,8	-9,9	-1,5	-1,2	-3	-15,8	-5,7	-4	4,3	-1,0	-3,1	-3,1	-3,1	-9	
SH Schleswig-Holstein		Anzahl	3900	0	2200	1800	600	30	200	1100	2000	100	0	700	0	300	100	0	0	13030
		% von Bundesland	29,9%	,0%	16,9%	13,8%	4,6%	,2%	1,5%	8,4%	15,3%	,8%	,0%	5,4%	,0%	,0%	2,3%	,8%	,0%	100,0%
		Korrigierte Residuen	8,1	-12,0	19,8	5,0	-18,6	,6	3,0	-1,6	-7,0	-13,9	-1,4	-1,7	-3,3	-10,5	19,4	-8	-3,1	
TH Thüringen		Anzahl	0	600	0	300	20	0	0	40	0	0	0	0	0	0	0	0	0	960
		% von Bundesland	,0%	62,5%	,0%	31,3%	2,1%	,0%	,0%	4,2%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	,0%	100,0%
		Korrigierte Residuen	-18,8	185,2	-11,1	17,7	-7,7	-1,4	-3,5	-5,1	-14,4	-5,2	-4	-7,6	-9	-2,8	-2,8	-2,8	-8	
Gesamt		Anzahl	190540	7600	81100	88100	65930	1460	8880	62840	125500	19500	100	40600	600	5900	5700	5900	500	710750
		% von Bundesland	26,8%	1,1%	11,4%	12,4%	9,3%	,2%	1,2%	8,8%	17,7%	2,7%	,0%	5,7%	,1%	,8%	,8%	,8%	,1%	100,0%

Abbildung 57 Migration – Schulform Quelle: Statistisches Bundesamt (Angaben gerundet)

3.5 Risikoschüler

Ein weiteres Interesse des deutschen PISA-Konsortiums galt der Erkennung leseschwacher Schüler durch Lehrkräfte. In Schulen mit Hauptschulbildungsgang wurden die für die Durchführung der PISA-Tests zuständigen Schulkoordinatoren gebeten, bei den Klassen- beziehungsweise Deutschlehrkräften Auskunft darüber einzuholen, welche der Schüler in der PISA-Stichprobe über eine so geringe Lesefähigkeit verfügen, dass sie erhebliche Schwierigkeiten beim Übergang in das Berufsleben haben dürften.

*Als deutlicher Befund zeigt sich hierbei, dass die meisten der schwachen Leserinnen und Leser von den Lehrkräften unerkannt bleiben. Insgesamt wurden von den Lehrkräften nur 55 dieser Risikoschülerinnen und –schüler im hier definierten Sinne als schwache Leser identifiziert. Von diesen Jugendlichen erzielten 28 Leistungen im PISA-Test, die unterhalb der Kompetenzstufe I liegen, 13 erreichten die Kompetenzstufe I, bei weiteren 13 Fällen lagen die Leistungen dagegen innerhalb der Kompetenzstufe II, und einer der als schwacher Leser identifizierten Jugendlichen erreichte Kompetenzstufe III. Wechselt man nun die Perspektive und fragt, welcher Anteil der Schülerinnen und Schüler, die aufgrund ihrer PISA-Testergebnisse im Hinblick auf ihre Lesekompetenz als Risikoschülerin beziehungsweise –schüler einzustufen sind, von ihren Lehrerinnen und Lehrern als schwache Leser angegeben wurde, ergibt sich ein relativ kleiner Überlappungsbereich (...). Da PISA keine umfassende Erhebung der diagnostischen Kompetenz von Lehrkräften beinhaltet, kann die Ursache für den relativ hohen Anteil unerkannt gebliebener Risikoschülerinnen und –schüler nicht genauer bestimmt werden. Die Ergebnisse sprechen jedoch dafür, dass es lohnenswert sein könnte, diesem Aspekt weiter nachzugehen und die Diagnose der Lesefähigkeiten von Schülerinnen und Schülern durch Lehrkräfte systematisch zu untersuchen*¹²²

¹²² Artelt, Stanat, Schneider und Schiefele: Lesekompetenz: Testkonstruktion und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 120

3.5.1 Risikoschüler – Leseschwäche (I)

Der Datensatz der PISA-Studie bietet die Möglichkeit der Analyse über die Variable *lese* im Datensatz der Ergänzungsstudie.¹²³ Im Haupttest kennzeichnet diese Variable die Schülerinnen und Schüler, die von den Lehrkräften als schlechte Leser klassifiziert wurden.¹²⁴

Vorgehensweise: Es werden zwei Gruppen gebildet, 1. die von Lehrkräften als schwache Leser identifizierten Schüler, 2. sonstige Schüler als Vergleichsgruppe. Wenn Lehrkräfte nicht in der Lage sein sollten, schwache Leser zu identifizieren, dann ließe sich kein systematischer Unterschied zwischen identifizierten Schülern und der Vergleichsgruppe hinsichtlich ihrer Lesefähigkeit nachweisen. Die Ausgangshypothese lautet entsprechend, dass der mittlere Lesetestwert derjenigen Schülerinnen und Schüler, die von den Lehrkräften als Risikoschüler diagnostiziert wurden, gleich dem Lesetestwert der Vergleichsgruppe ist. Die Gegenhypothese, die von der Diagnosefähigkeit der Lehrkräfte ausgeht, behauptet eine statistisch signifikant geringere mittlere Lesefähigkeit der identifizierten Schüler.

Modell: $X_i \sim N(\mu_x, \sigma_x^2)$, $Y_j \sim N(\mu_y, \sigma_y^2)$. Wegen $\sigma_x^2 \neq \sigma_y^2$ wird der Welch-Test¹²⁵ eingesetzt. $H_0: \mu_x = \mu_y$, wobei die eine Gruppe aus identifizierten, die Vergleichsgruppe aus den nicht identifizierten Schülerinnen und Schülern besteht.¹²⁶ Die Ausgangshypothese beinhaltet die Gleichheit der Mittelwerte der identifizierten Schüler mit der Vergleichsgruppe entsprechend der vorangegangenen Untersuchung. Die Darstellung berücksichtigt Haupt- und Sonderschule.

Gruppenstatistiken^a

	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Internat. Lesen PV 1				
Schwache Leser rec				
nicht ident	5051	386,1710	77,44783	1,08973
identifiziert	255	323,8776	64,41659	4,03392

a. Schulform = Hauptschule

¹²³ PISA-OVE-9kl

¹²⁴ Kunter, Schümer, Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Stanat, Tillmann, Weiß
PISA 2000 Dokumentation der Erhebungsinstrumente. S. 24

¹²⁵ Test bei unabhängigen Stichproben und ungleichen Varianzen (Welch-Test)

¹²⁶ (Ipvread als ZV im KS-Test auf Normalverteilung $p=0,524$) Berechnungen über den ungewichteten Datensatz

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
Internat. Lesen PV 1	Varianzen sind gleich	17,695	,000	12,625	5304	,000	62,2934	4,93407	52,62060	71,96621
	Varianzen sind nicht gleich			14,908	292,347	,000	62,2934	4,17852	54,06960	70,51720

a. Schulform = Hauptschule

Abbildung 58 Leseschwäche Diagnose (I) Hauptschule

Gruppenstatistiker^a

		N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Internat. Lesen PV 1	Schwache Leser rec nicht ident	672	244,0936	53,20198	2,05231
	identifiziert	80	203,1785	76,06480	8,50430

a. Schulform = Sonderschule

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
Internat. Lesen PV 1	Varianzen sind gleich	36,512	,000	6,172	750	,000	40,9150	6,62926	27,90089	53,92913
	Varianzen sind nicht gleich			4,677	88,434	,000	40,9150	8,74844	23,53052	58,29950

a. Schulform = Sonderschule

Abbildung 59 Leseschwäche Diagnose (I) Sonderschule

In der Hauptschule erreicht die Vergleichsgruppe mit $\bar{x}_{\text{nicht_ident}} = 386$ zu $\bar{x}_{\text{identifiziert}} = 324$ der identifizierten Schülerinnen und Schüler ein statistisch signifikant ($p < 0,01$) höheres mittleres Testergebnis. Unter den Schülern der Sonderschule erreicht die Vergleichsgruppe mit einem Mittelwert von $\bar{x}_{\text{nicht_ident}} = 244$ bei $N=672$ zu einem Mittelwert von $\bar{x}_{\text{identifiziert}} = 203$ bei $N=80$ ebenfalls einen statistisch signifikant höheren Wert. Der Vergleich zeigt, dass die von Lehrkräften als leseschwach eingestuften Schüler statistisch signifikant niedrigere Testwerte auf der internationalen Skala erzielen als die Vergleichsgruppe.

3.5.2 Risikoschüler – Leseschwäche (II)

In einer Zusatzbefragung wurden die Eltern der Schüler nach der Grundschulempfehlung gefragt. Der Beantwortung dieser Frage kamen rund 72% der Eltern nach. Ein großer Teil der Eltern orientiert sich nicht an der Grundschulempfehlung beim Übergang ihrer Kinder in die Sekundarstufe. Es bestehen systematische Unterschiede zwischen den Bundesländern. Wenn nun die Diagnosefähigkeit der Lehrkräfte an Grundschulen ähnlich ausgeprägt sein sollte, wie die von Artelt, Stanat, Schneider und Schiefele in der PISA-Studie berichtete der Lehrkräfte an Hauptschulen, dann dürfte kein systematischer Unterschied in den Testwerten zwischen denjenigen Schülern eines Bildungsgangs mit niedrigerer, höherer oder dem Bildungsgang adäquater Grundschulempfehlung nachweisbar sein. Vorausgesetzt wird die Bildung einer Variablen für jeden Schüler, die angibt, ob die Grundschulempfehlung den von den Eltern tatsächlich gewählten Bildungsgang unter- oder überschätzt beziehungsweise ihm adäquat ist. Im Sinne von Artelt, Stanat, Schneider und Schiefele wäre nun die Gültigkeit der Annahme zu zeigen, dass sich in Bezug auf die Testergebnisse zwischen den Gruppen kein statistisch signifikanter Unterschied nachweisen lässt. Wenn also zum Beispiel Realschüler mit Hauptschulempfehlung die gleichen Testresultate bei PISA erzielt haben sollten, wie Realschüler mit Realschulempfehlung oder sogar Gymnasialempfehlung, dann wäre dies ein Beleg für mangelnde Diagnosefähigkeit von Lehrkräften.

Das hier gewählte Verfahren ist die Varianzanalyse. Die Mittelwertdiagramme werden berichtet. Als abhängige Variable kommen der Lesetestwert, der Mathematiktestwert sowie der Biologietestwert in Frage.

3.5.2.1 Hauptschule

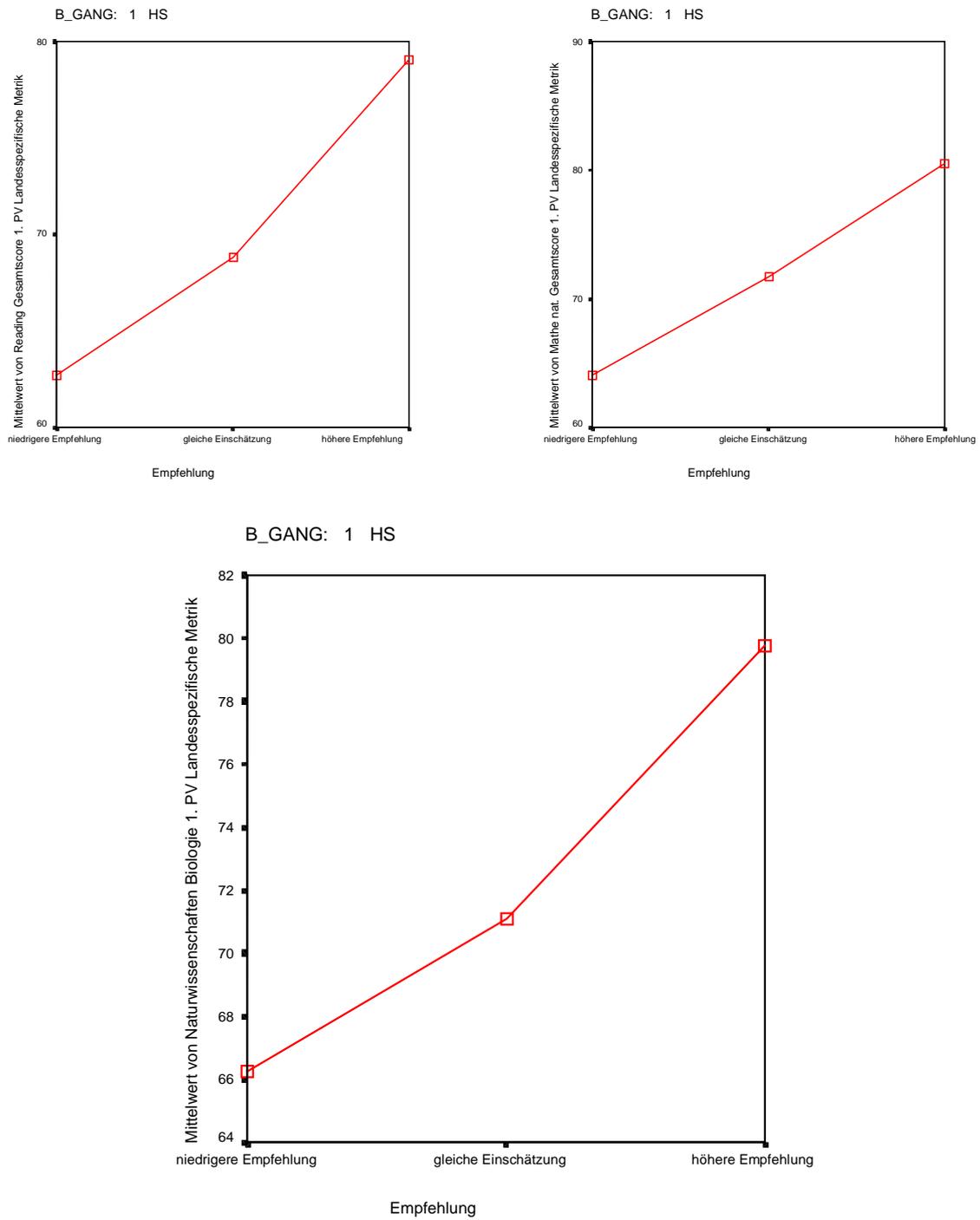


Abbildung 60 Leseschwäche Diagnose(III) Hauptschule

3.5.2.2 Realschule

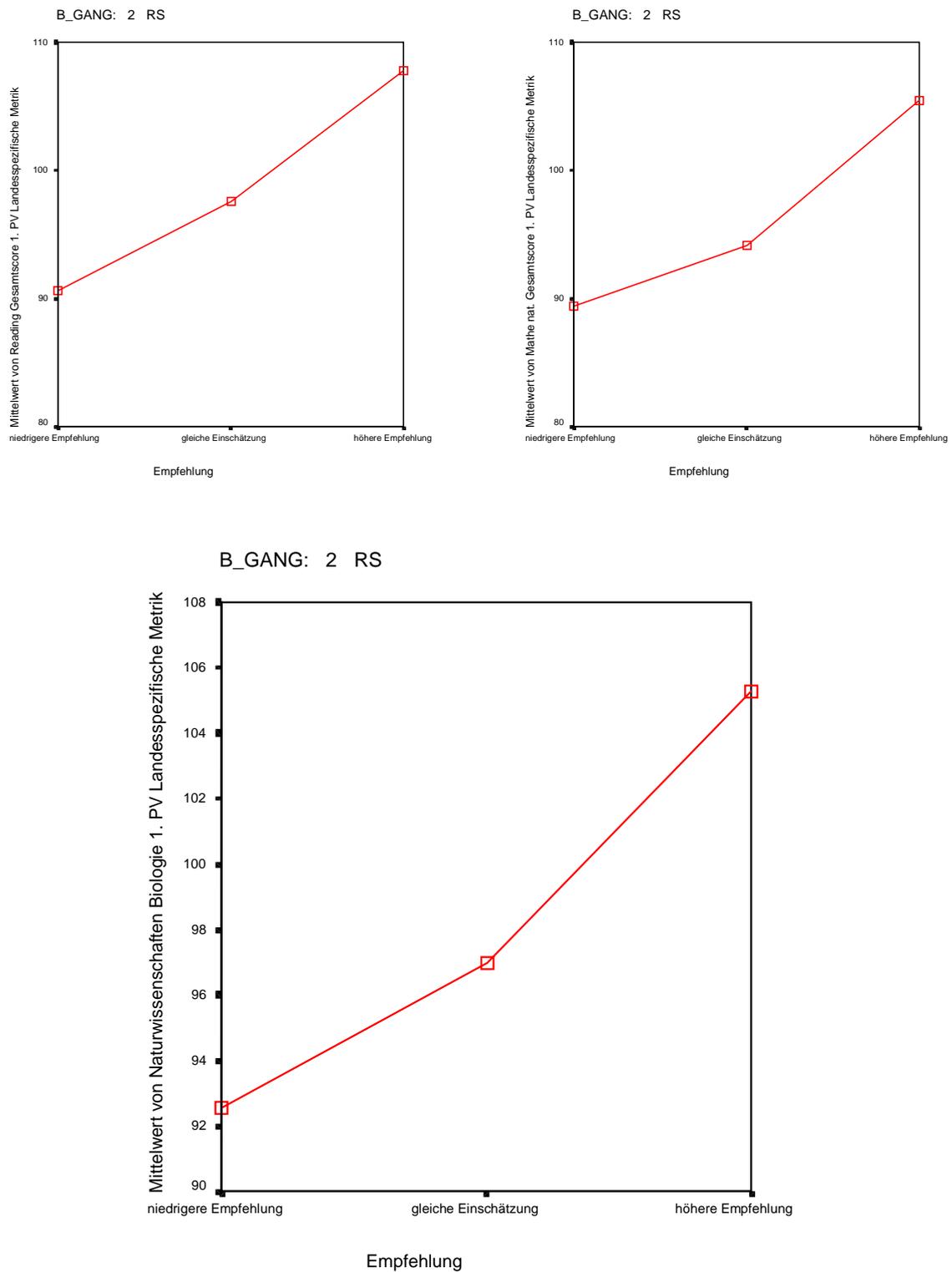


Abbildung 61 Leseschwäche Diagnose(III) Realschule

3.5.2.3 Gymnasium

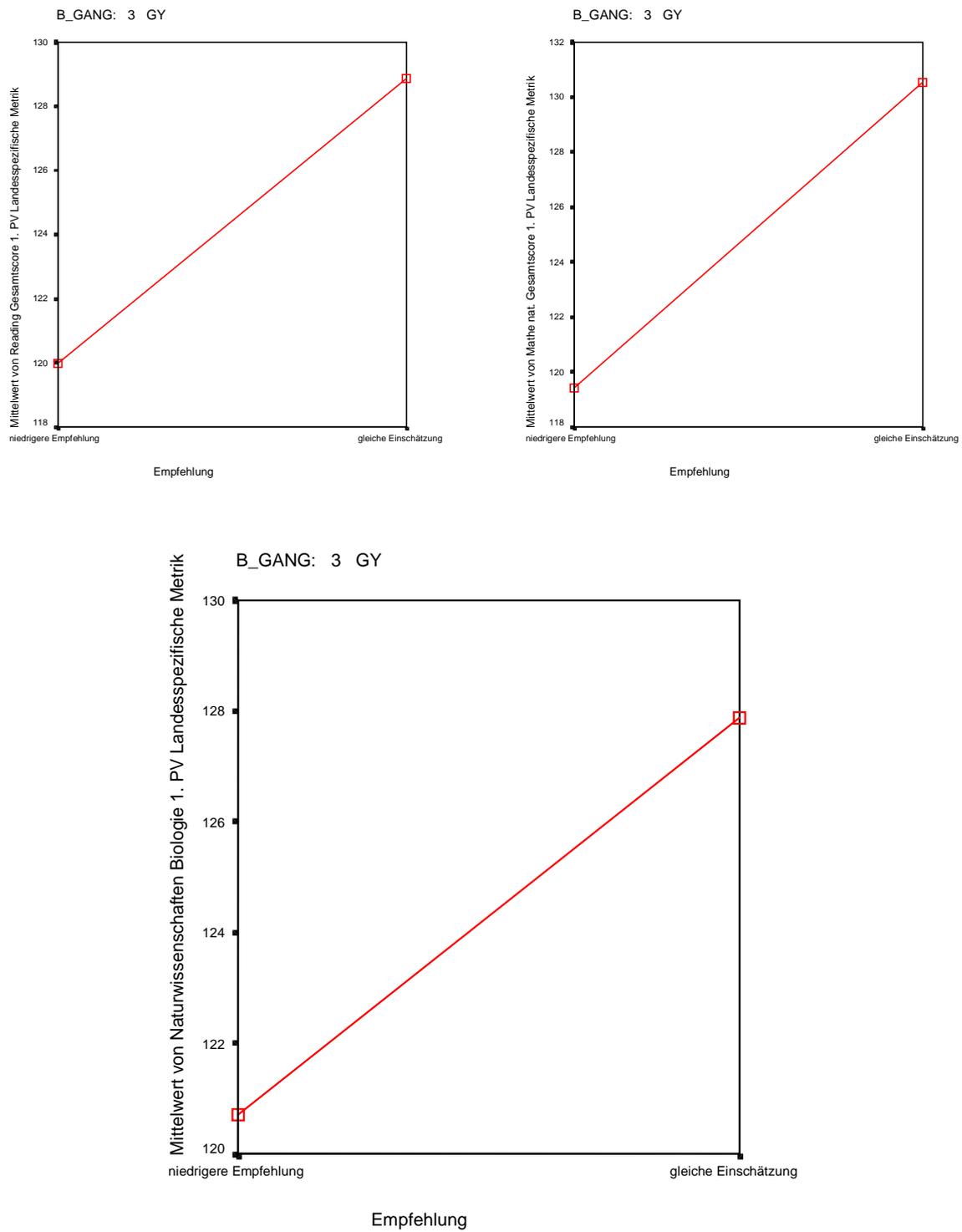


Abbildung 62 Leseschwäche Diagnose(III) Gymnasium

Das Ergebnis lässt sich wie folgt zusammenfassen: Die Ausgangshypothese der Gleichheit der Mittelwerte kann in sämtlichen Fällen verworfen werden. Diejenigen Schülerinnen und Schüler mit in Relation zu der tatsächlichen Wahl der Eltern niedrigeren Grundschulempfehlungen erzielen deutlich geringere Testwerte als diejenigen Schüler mit in Relation höheren Grundschulempfehlungen.

3.5.3 Risikoschüler - Familienstruktur

Der Zusammenhang von Testleistungsergebnissen und Familienstruktur war Gegenstand des nationalen Ergänzungstests. Dieser Aspekt steht ebenfalls im Zusammenhang mit der vorangegangenen Fragestellung.¹²⁷ Die internationalen Berichte weisen für Schülerinnen und Schüler aus alleinerziehenden Haushalten (Students from single-parent families) niedrigere Testwerte aus.¹²⁸ In der deutschen PISA-Studie hingegen gilt dieser Zusammenhang als nicht bestätigt.¹²⁹

Der folgende Abschnitt soll dieser Frage nachgehen. Die Codierung der Variablen wurde in der PISA-Studie wie folgt vorgenommen:

*Students were asked to indicate who usually lives at home with them. The variable FAMSTRUC is based on the first four items of this question (ST04Q01, ST04Q02, ST04Q03 and ST04Q04). It takes four values and indicates the nature of the student's family: (i) single-parent family (students who reported living with one of the following: mother, father, female guardian or male guardian); (ii) nuclear family (students who reported living with a mother and a father); (iii) mixed family (students who reported living with a mother and a male guardian, a father and a female guardian, or two guardians); and (iv) other response combinations.*¹³⁰

Die folgende Abbildung trägt die Variablen *höchster sozioökonomischer Index* (der Familie) als Mittelwert je Teilnehmerstaat und den Anteil der Kernfamilien (*nuclear family*) ab.

¹²⁷ Schümer, Weiß, Steinert, Baumert, Tillmann, Meier: Lebens- und Lernbedingungen von Jugendlichen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 478

¹²⁸ vgl. z.B.: OECD: Learning for Tomorrow's World - First Results from PISA 2003, S.392

¹²⁹ Tillmann, Meier & Schümer, Weiß, Steinert, Baumert, Tillmann, Meier: Lebens- und Lernbedingungen von Jugendlichen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 479

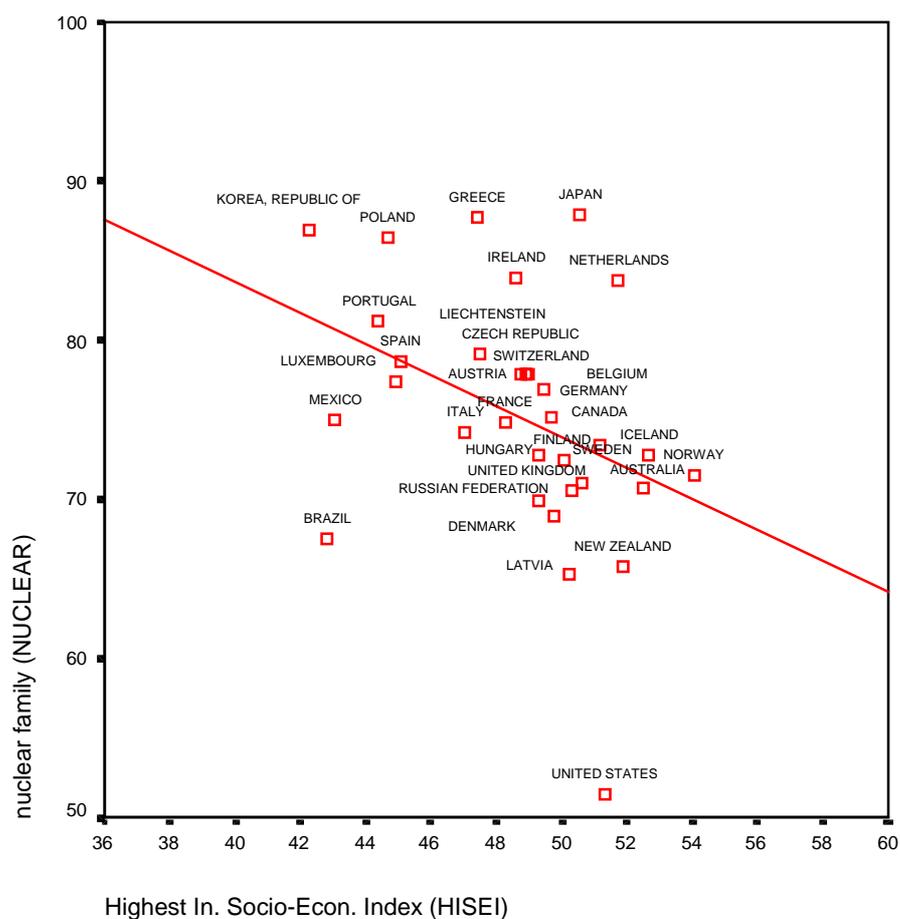


Abbildung 63 Anteil Kernfamilie in Abhängigkeit zum sozioökonomischen Index

Erkennbar ist der negative Zusammenhang von höchstem sozioökonomischem Index der Familie als Mittel je Teilnehmerstaat und Anteil der Kernfamilien. Ein linearer Zusammenhang lässt sich mit $r_p = -0,389$ ($p < 0,05$) beziehungsweise von $r_s = -0,474$ ($p < 0,01$) feststellen.

Korrelationen			HISEI_M
NUCLEAR	Korrelation nach Pearson		-,387*
	Signifikanz (2-seitig)		,028
	N		32

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Korrelationen			HISEI_M
Spearman-Rho	NUCLEAR	Korrelationskoeffizient	-,474**
		Sig. (2-seitig)	,006
		N	32

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 64 Anteil Kernfamilie in Abhängigkeit zum sozioökonomischen Index

Die Querschnittsanalyse zeigt bei steigendem höchstem familiärem sozioökonomischem Index als Mittelwert der Stichprobe eines Landes eine Verringerung des Anteils der Kernfamilie, während der Anteil der Alleinerziehenden (single parent) und geringfügig stärker der Anteil der neuen Partnerschaften (mixed family) zunimmt.

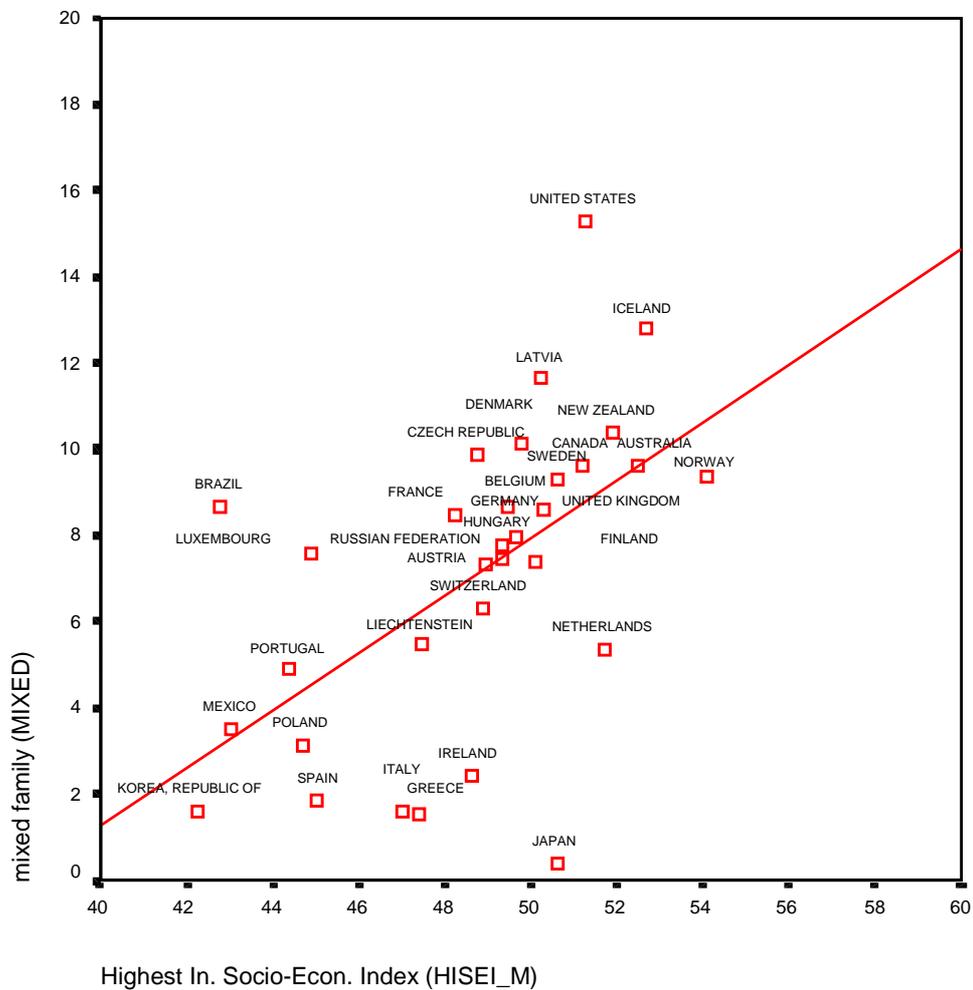
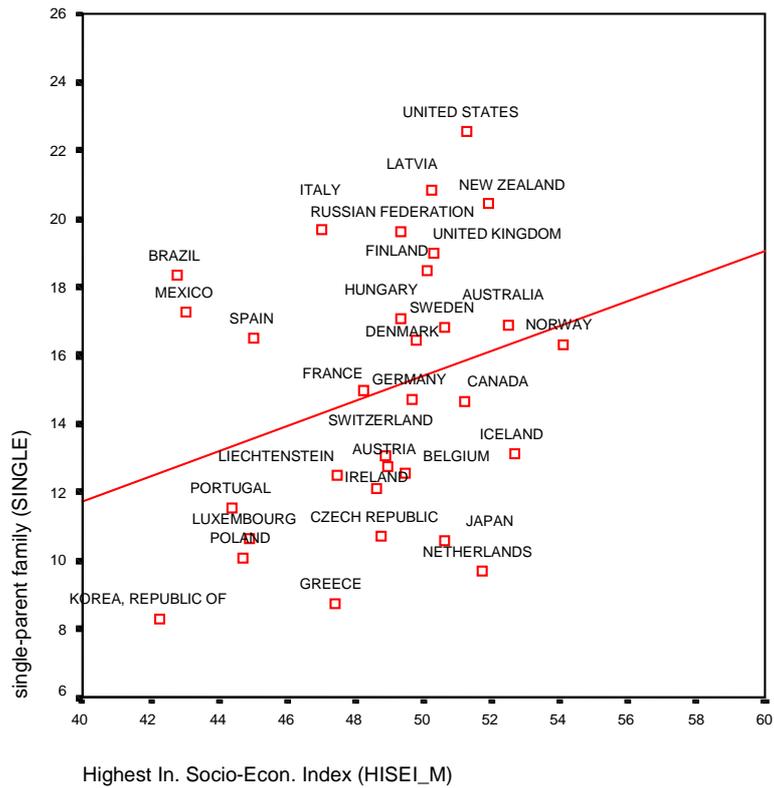


Abbildung 65 Anteil Neuer Partnerschaften (mixed family) in Abhängigkeit zum sozioökonomischen Index



Korrelationen

		HISEI_M
SINGLE	Korrelation nach Pearson	,288
	Signifikanz (2-seitig)	,110
	N	32
NUCLEAR	Korrelation nach Pearson	-,387*
	Signifikanz (2-seitig)	,028
	N	32
MIXED	Korrelation nach Pearson	,563*
	Signifikanz (2-seitig)	,001
	N	32
OTHER	Korrelation nach Pearson	-,106
	Signifikanz (2-seitig)	,562
	N	32

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Korrelationen

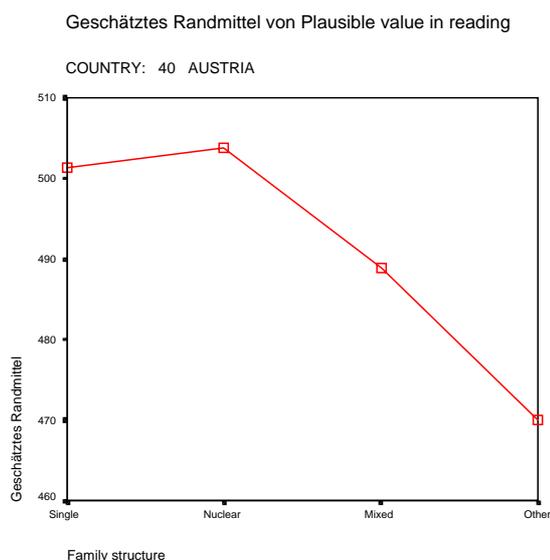
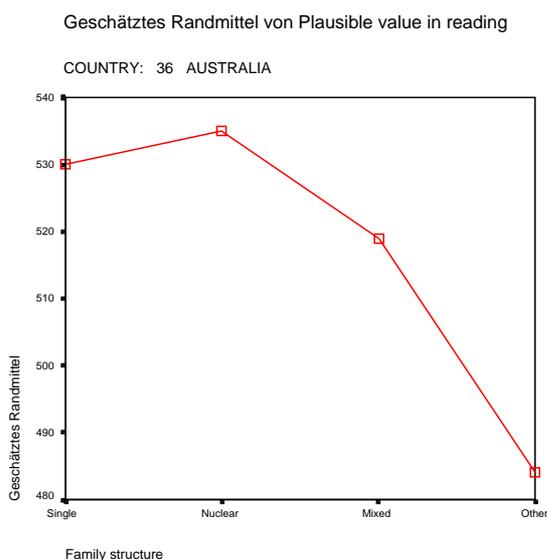
		HISEI_M
Spearman-Rho SINGLE	Korrelationskoeffizient	,304
	Sig. (2-seitig)	,090
	N	32
NUCLEAR	Korrelationskoeffizient	-,474*
	Sig. (2-seitig)	,006
	N	32
MIXED	Korrelationskoeffizient	,622*
	Sig. (2-seitig)	,000
	N	32
OTHER	Korrelationskoeffizient	-,235
	Sig. (2-seitig)	,195
	N	32

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 66 Haushaltstyp [Single, Nuclear, Mixed, Other] in Abhängigkeit zum sozioökonomischen Index

Wie wirkt sich dieser Zusammenhang auf die Testleistungen in PISA aus? Tillmann und Meier hatten eine multivariate, zweifaktorielle Kovarianzanalyse berechnet, in der als Faktoren Bildungsgang und Familientyp und als Kovariate Sozialschicht berücksichtigt wurden.¹³¹

Am internationalen Datensatz soll vergleichbar vorgegangen werden. Als Kovariaten finden *Geschlecht*, höchster soziökonomischer Index (*HISEI*), zusätzlich aber auch die Klassenstufe Berücksichtigung.¹³²

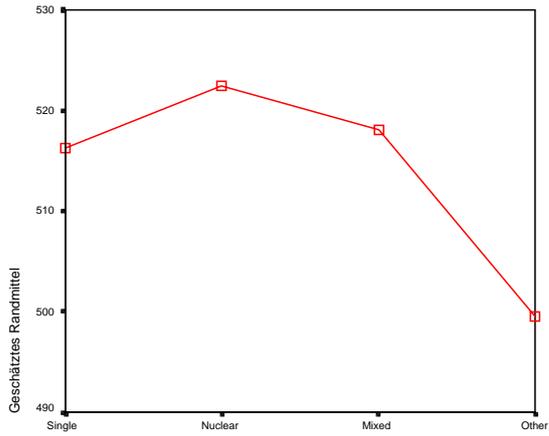


¹³¹ Schümer, Weiß, Steinert, Baumert, Tillmann, Meier: Lebens- und Lernbedingungen von Jugendlichen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 480

¹³² Kovariaten: höchster soziökonomischer Index, Klassenstufe, Geschlecht. Berichtet werden die geschätzten Randmittel derjenigen Teilnehmerstaaten, deren Variablen single, nuclear, mixed in Hinblick auf die Effektgröße mindestens ein N=238 aufweisen. Brasilien und Mexiko sind deshalb nicht berücksichtigt.

Geschätztes Randmittel von Plausible value in reading

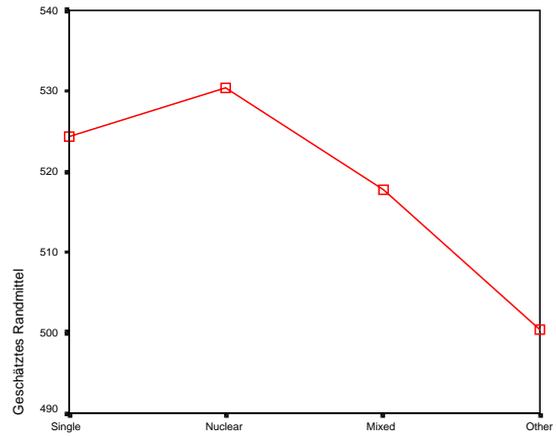
COUNTRY: 56 BELGIUM



Family structure

Geschätztes Randmittel von Plausible value in reading

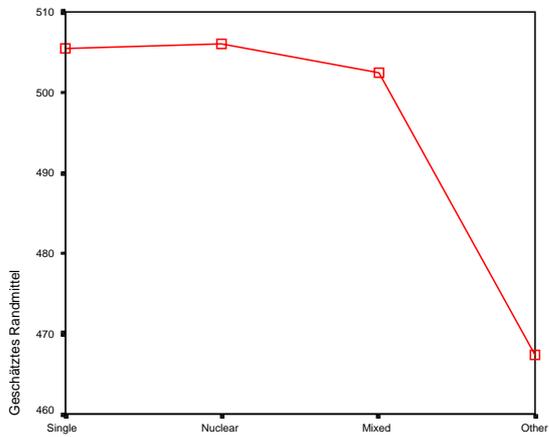
COUNTRY: 124 CANADA



Family structure

Geschätztes Randmittel von Plausible value in reading

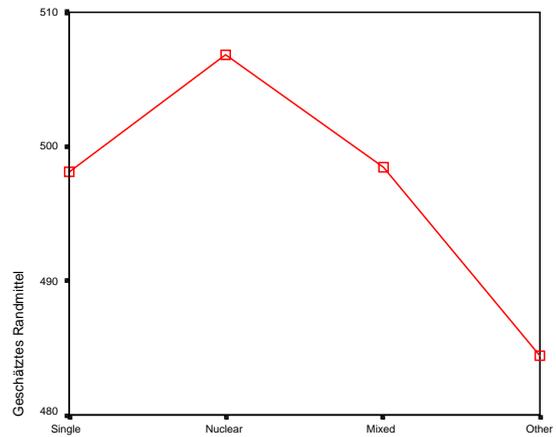
COUNTRY: 203 CZECH REPUBLIC



Family structure

Geschätztes Randmittel von Plausible value in reading

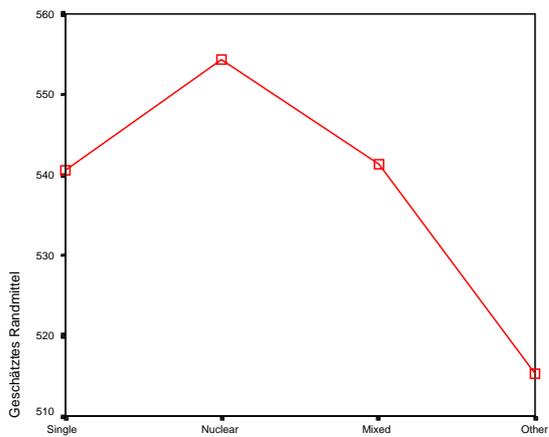
COUNTRY: 208 DENMARK



Family structure

Geschätztes Randmittel von Plausible value in reading

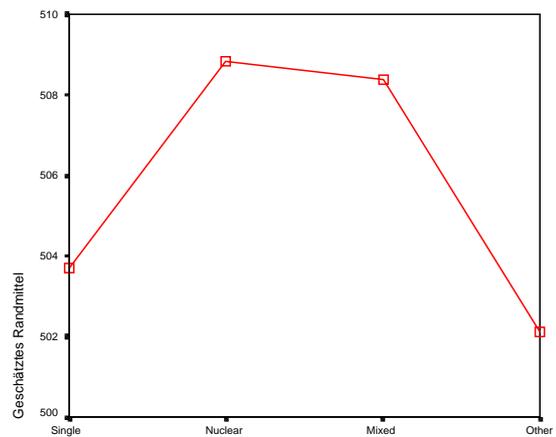
COUNTRY: 246 FINLAND



Family structure

Geschätztes Randmittel von Plausible value in reading

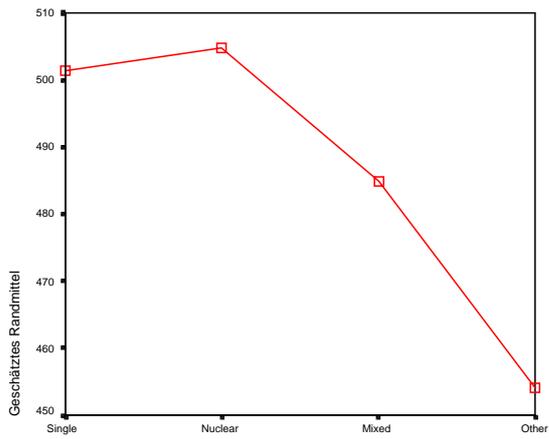
COUNTRY: 250 FRANCE



Family structure

Geschätztes Randmittel von Plausible value in reading

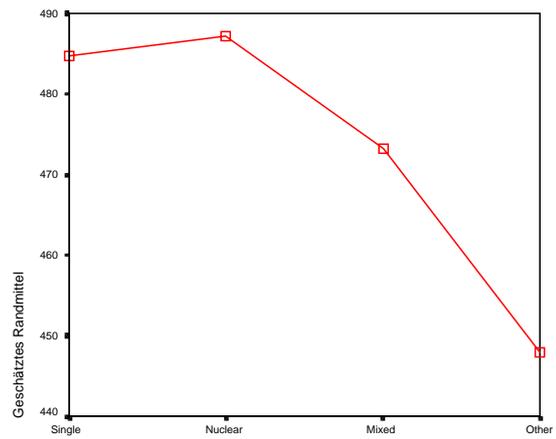
COUNTRY: 276 GERMANY



Family structure

Geschätztes Randmittel von Plausible value in reading

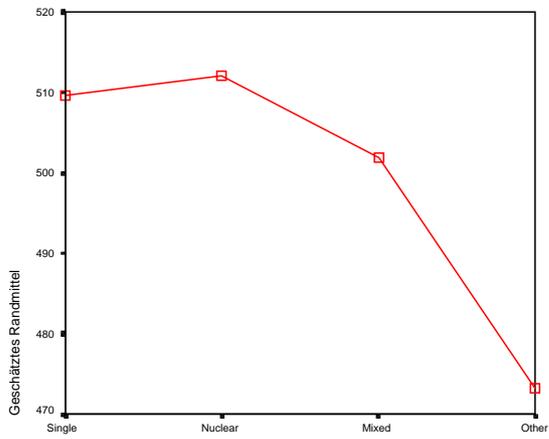
COUNTRY: 348 HUNGARY



Family structure

Geschätztes Randmittel von Plausible value in reading

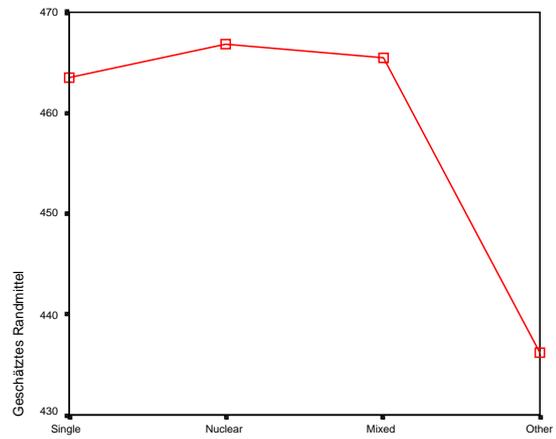
COUNTRY: 352 ICELAND



Family structure

Geschätztes Randmittel von Plausible value in reading

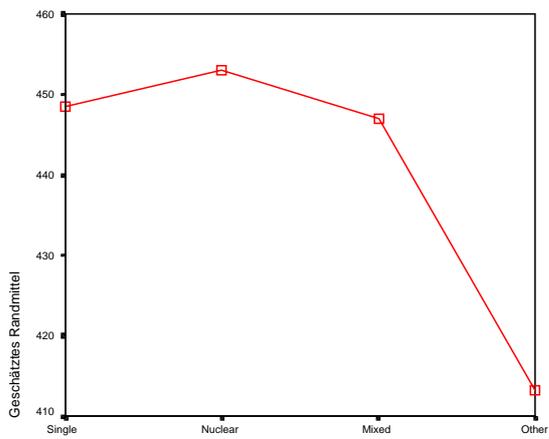
COUNTRY: 428 LATVIA



Family structure

Geschätztes Randmittel von Plausible value in reading

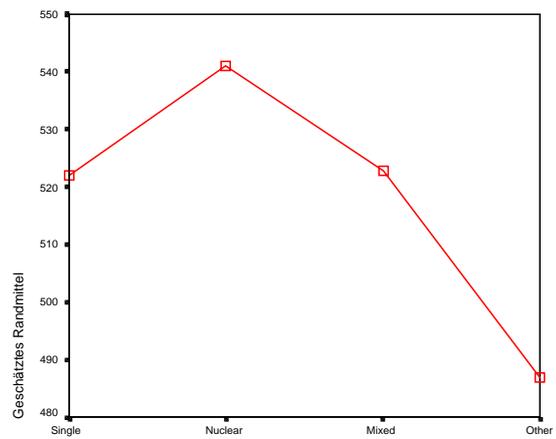
COUNTRY: 442 LUXEMBOURG



Family structure

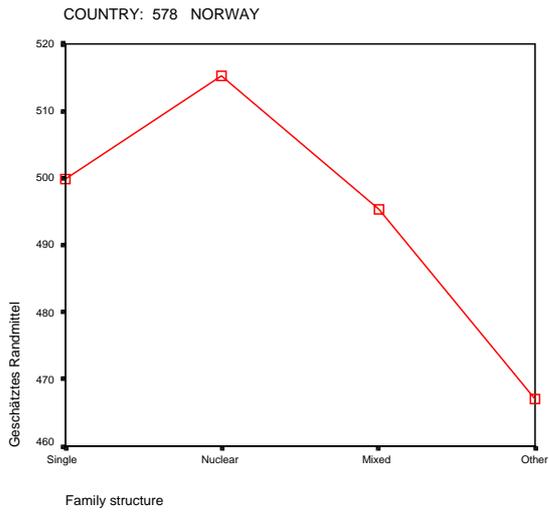
Geschätztes Randmittel von Plausible value in reading

COUNTRY: 554 NEW ZEALAND

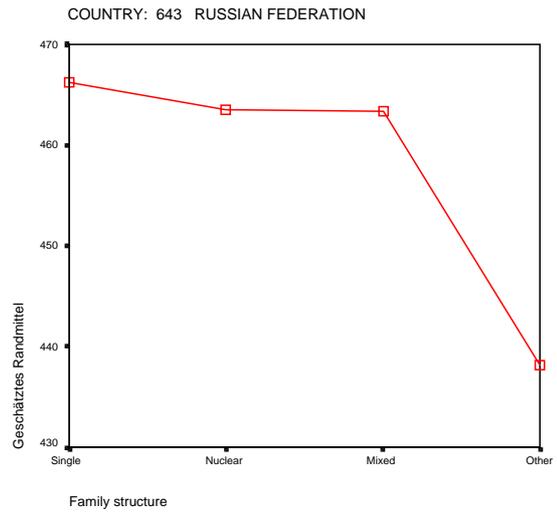


Family structure

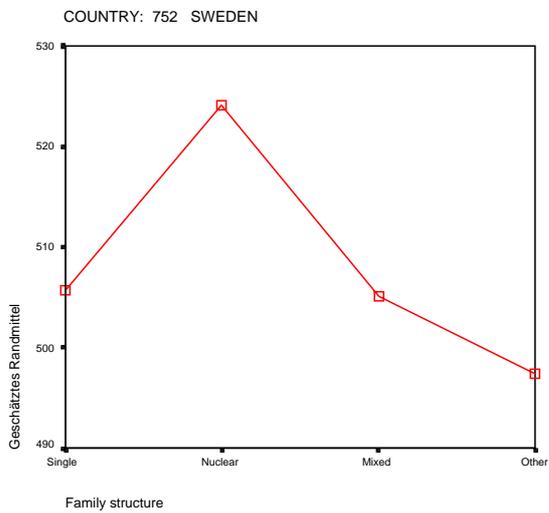
Geschätztes Randmittel von Plausible value in reading



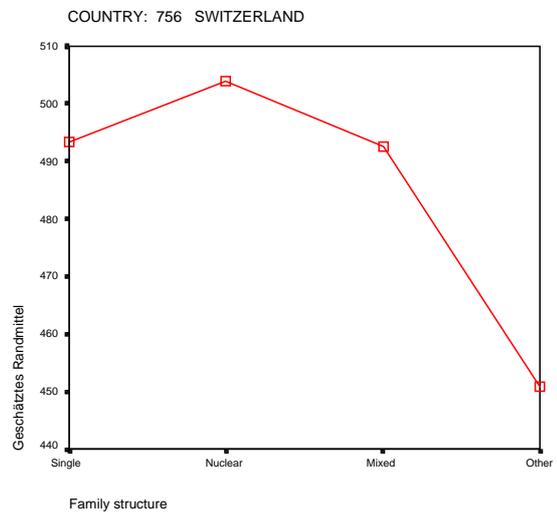
Geschätztes Randmittel von Plausible value in reading



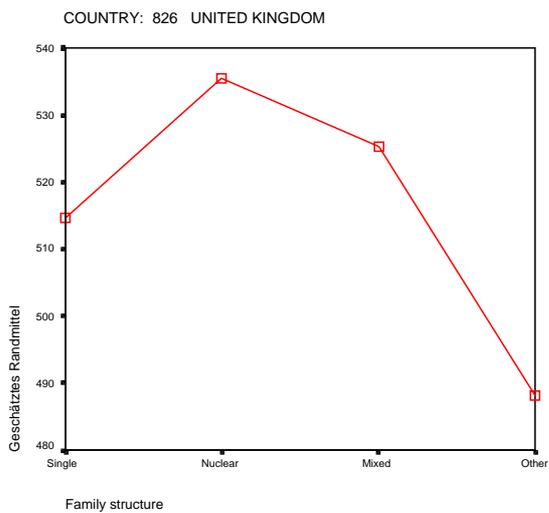
Geschätztes Randmittel von Plausible value in reading



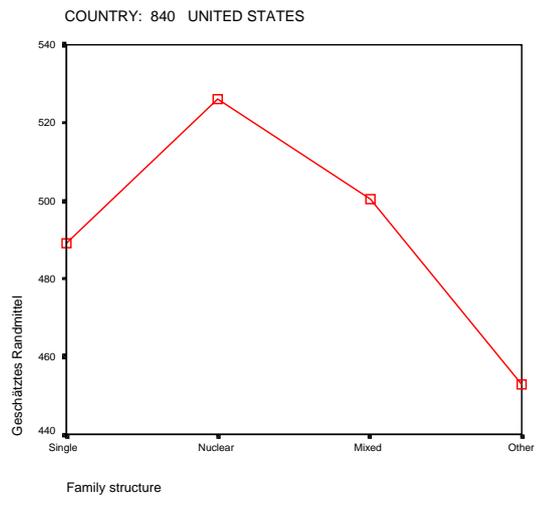
Geschätztes Randmittel von Plausible value in reading



Geschätztes Randmittel von Plausible value in reading



Geschätztes Randmittel von Plausible value in reading



Abbildungen 67 Testleistung und Haushaltstyp, geschätzte Randmittel

Die Ergebnisse folgen über sämtliche dargestellte Teilnehmerstaaten dem gleichen Muster, wenn auch die Differenzen variieren. Schülerinnen und Schüler aus Kernfamilien erreichen nach Kontrolle der Variablen *Klassenstufe*, *Geschlecht* und *höchstem sozioökonomischem Index der Familie* (HISEI) durchweg die höchsten mittleren Werte. Die Unterschiede sind zumeist zu beiden relevanten Vergleichsgruppen (mixed, single) statistisch signifikant. Zu berücksichtigen ist das Verhältnis zwischen Effektgröße und Stichprobenumfang.

Für die Bundesrepublik ergibt sich ein Unterschied der Testleistung zwischen Schülern aus Kernfamilien und Alleinerziehenden sowie ein Unterschied zwischen der Kernfamilie und Neuer Partnerschaft (*mixed*). Das Muster der Bundesrepublik entspricht dem der übrigen Staaten.

Die Autoren hatten den längsschnittanalytischen Aspekt der soziodemographischen Entwicklung angesprochen.¹³³ Die *Langen Reihen* des Statistischen Bundesamtes beinhalten Kennziffern soziodemographischer Entwicklung. Die Zeitreihe der *Familien mit Kindern in der Familie nach Familienstand der Bezugsperson und Altersgruppen der Kinder*¹³⁴ zeigt über die Jahre 1975 bis 2001 einen systematischen Wandel in der Zusammensetzung der Haushaltsstrukturen in der Bundesrepublik Deutschland. Die folgende Abbildung zeigt die Zeitreihen der beiden Gruppen von 1975 bis 2001 sowie einen weiteren Split für *Familien mit Kindern in der Familie nach Familienstand der Bezugsperson und Altersgruppen der Kinder* mit mindestens einem Kind unter 15 Jahren.

¹³³ Schümer, Weiß, Steinert, Baumert, Tillmann, Meier: Lebens- und Lernbedingungen von Jugendlichen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 481

¹³⁴ Quelle: Statistisches Bundesamt Lange Reihen Tab. L204

Die beiden Jahre 1961 und 1970 werden aus methodischen Gründen nicht ausgewiesen. 1961 und 1970: Ergebnis der Volkszählung; sonst Ergebnis des Mikrozensus (1975 aus der EG-Arbeitskräftestichprobe).- 1961 wohnberechtigte Bevölkerung, ab 1970 Bevölkerung am Familienwohnsitz.- Als Alleinerziehende zählen auch Väter und Mütter mit volljährigen Kindern.- Einschl. der allein Erziehenden, die Partner/in einer nichtehelichen Lebensgemeinschaft sind.

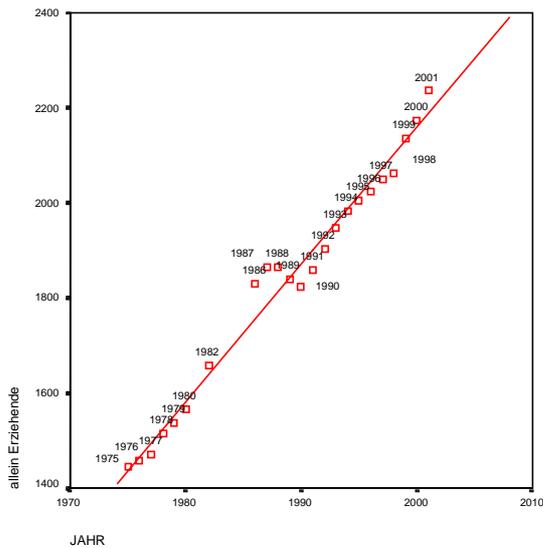
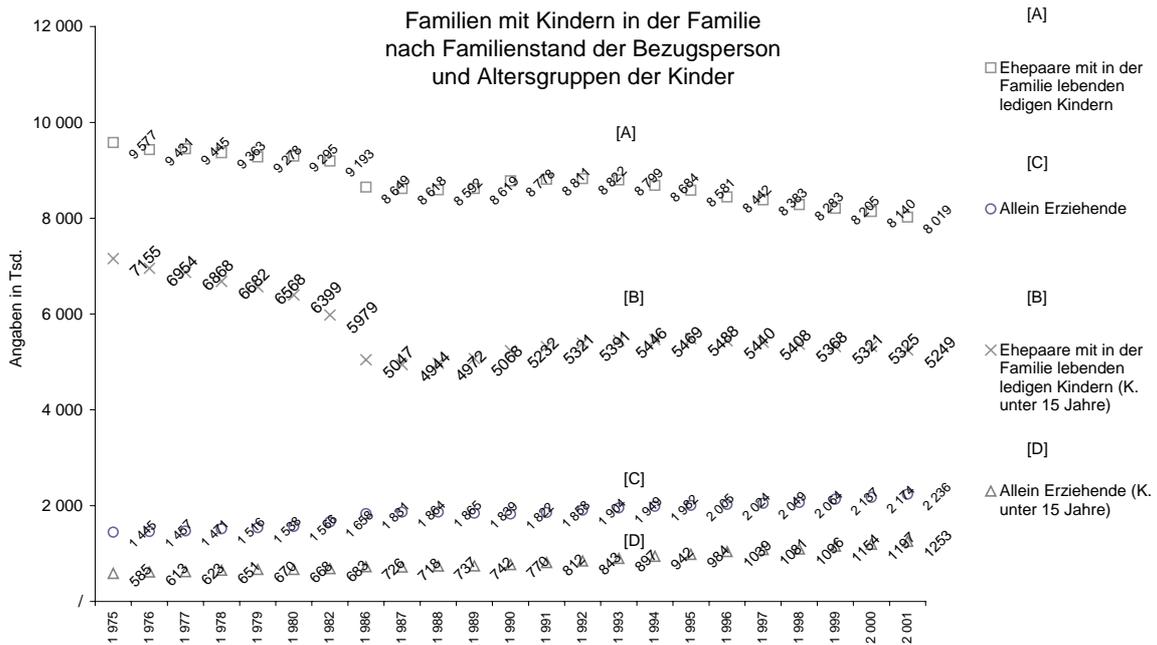
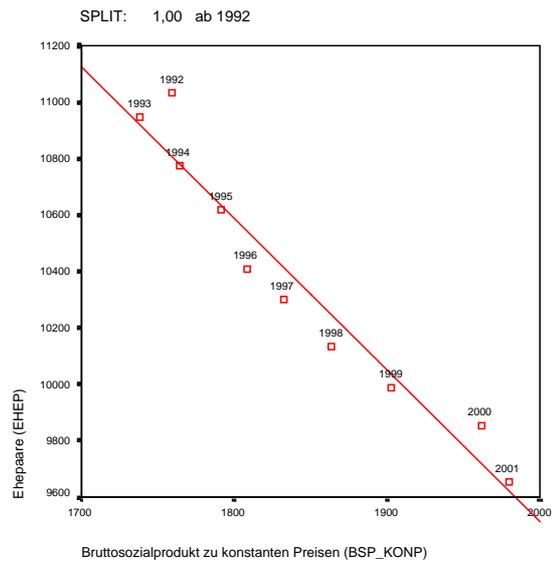
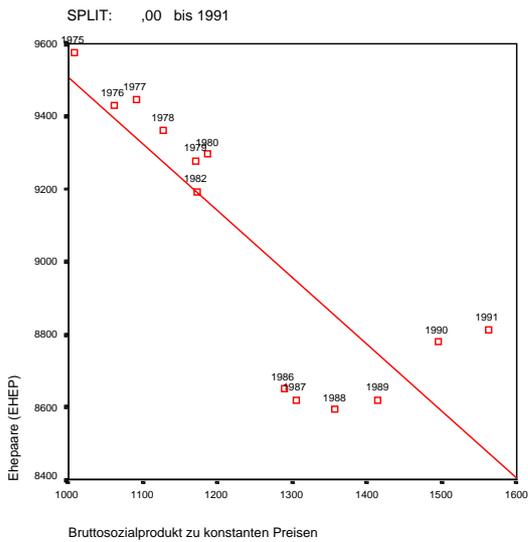
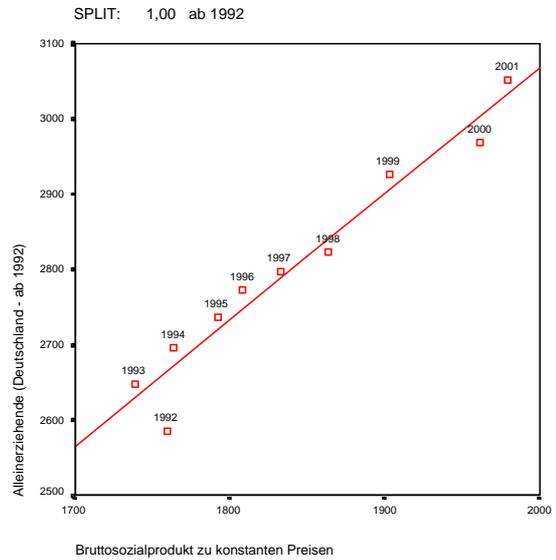
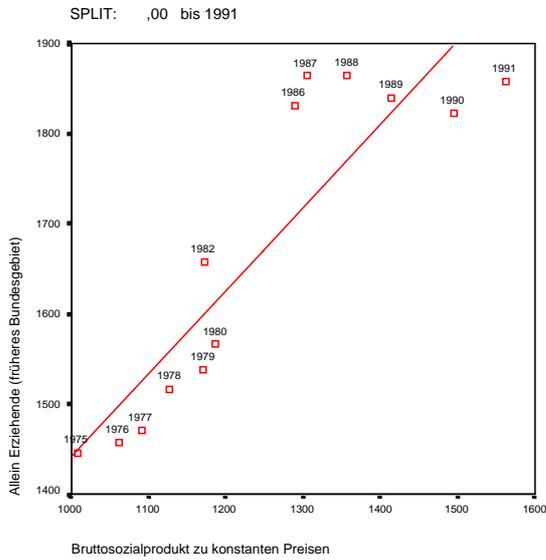


Abbildung 68 Familie längsschnittanalytische Betrachtung

Mit $r_p = 0,99$ / $r_s = 0,98$ (Korrelation: Anzahl Alleinerziehende/Jahr) und $r_p = -0,95$ / $r_s = -0,899$ (Korrelation: Anzahl Ehepaare/Jahr) zeigt sich ein statistischer Zusammenhang. Dieser vorgefundene Zusammenhang korreliert signifikant mit Wohlstandsindikatoren, hier gemessen am *Bruttonationaleinkommen* (Bruttosozialprodukt in Mrd. Euro) in konstanten Preisen¹³⁵ (bis 1991 früheres

¹³⁵ Daten: Statistisches Bundesamt LR1

Bundesgebiet, ab 1992 Deutschland). Zusätzliche Indikatoren (*Bruttoinlandsprodukt – zu jeweiligen Preisen* beziehungsweise zu konstanten Preisen sowie *Volkseinkommen*) werden hier nicht dargestellt, weisen jedoch fast identische Ergebnisse auf.



Korrelationen²

	BIP_JEWP	BIP_KONP	BSP_JEWP	BSP_KONP	VOLKEINK
AL_ERZ Korrelation nach Pearson	,936*	,894*	,936*	,893*	,936*
Signifikanz (2-seitig)	,000	,000	,000	,000	,000
N	13	13	13	13	13

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

a. SPLIT = bis 1991

Korrelationen^a

	BIP_JEWP	BIP_KONP	BSP_JEWP	BSP_KONP	VOLKEINK
AL_ERZ Korrelation nach Pearson	,984*	,979*	,988*	,972*	,982*
Signifikanz (2-seitig)	,000	,000	,000	,000	,000
N	10	10	10	10	10

** - Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

a. SPLIT = ab 1992

Abbildung 69 Familie längsschnittanalytische Betrachtung/Wohlstandsindikatoren

3.5.4 Schlussfolgerung

Die vorliegenden Daten zeigen einen Zusammenhang von (zunehmendem) sozioökonomischem Niveau und (sich verringerndem) Anteil der Kernfamilie sowohl für die Bundesrepublik Deutschland (Längsschnitt) als auch über die gesamten Teilnehmerstaaten (Querschnitt). Dabei lässt sich ein statistischer Zusammenhang zwischen diskontinuierlicher Elternschaft und Testleistung feststellen. Dieser Befund stimmt sowohl mit den von der OECD publizierten Daten als auch mit den berichteten Ergebnissen von Fertig und Schmidt überein

Students living in an intact family, i.e. together with both parents, perform significantly better throughout all quantiles.¹³⁶

¹³⁶ Fertig, M., Schmidt, C.M.: The Role of Background Factors for Reading Literacy: Straight National Scores in the PISA 2000 Study Discussion Paper No. 545 August 2002 S. 6 http://ssrn.com/abstract_id=323599

3.6 Diagnose und Trennschärfe

Die mangelnde Trennschärfe der Beurteilungen durch Lehrkräfte wurde von Artelt, Stanat, Schneider und Schiefele angesprochen.¹³⁷ Die Testwerteverteilungen der als Risikoschüler eingestuften Fälle überlagert die der nicht als Risikoschüler eingestuften Fälle. Daraus ließe sich die These ableiten, dass zwar in der Tendenz leseschwache Schülerinnen und Schüler erkannt werden, wie oben gezeigt sogar mit deutlichen Unterschieden der Mittelwerte, dennoch eine Vielzahl an Fehlklassifikationen vorliegen. Im Folgenden wird nun gezeigt, dass diese Interpretation auf der Gleichsetzung von *Testeigenschaften* und *Testergebnis* beruht.

Aus den empirisch gewonnenen Daten ziehen Artelt, Stanat, Schneider und Schiefele anhand eines Verfahrens (PISA-Studie) rückwirkend Schlüsse auf die Sensitivität und Spezifität der Schülerklassifikationen durch Lehrkräfte (als statistische Größe). Vom Prinzip her wird also ein Testverfahren mit einem weiteren Testverfahren verglichen und bewertet. Die nun folgenden Ausführungen zeigen, dass ohne genaue Kenntnis der Vergleichsverfahren und der ihnen zugrunde liegenden Voraussetzungen, insbesondere der Prävalenz der zu spezifizierenden Fälle und prinzipiell auch ohne Kenntnis wesentlicher Daten zu Sensitivität und Spezifität des Meßinstruments, kein Rückschluß auf diagnostische Eigenschaften rückwirkend möglich ist.

3.6.1 Begriffsdefinitionen

Sensitivität Unter Sensitivität (diagnostische Sensitivität¹³⁸) versteht man das Ansprechen eines Tests beispielsweise auf eine Infektionskrankheit, die –katamnestisch– als solche diagnostiziert wird. $RP/(RP+FN)$

Spezifität Unter Spezifität (diagnostische Spezifität¹³⁹) eines Tests wird die Qualität eines Tests verstanden, bei negativer Diagnose auch einen negativen Zustand korrekt vorauszusagen. $RN/(RN+FP)$

Positiv prädiktiver Wert Der positiv prädiktive Wert ist der Anteil der Individuen mit richtig positivem Testresultat gegenüber allen im Test positiven Individuen: $RP/(RP+FP)$.

¹³⁷ Artelt, Stanat, Schneider und Schiefele: Lesekompetenz: Testkonstruktion und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 120

¹³⁸ Man unterscheidet zwischen analytischer und diagnostischer Sensitivität.

¹³⁹ Man unterscheidet zwischen analytischer und diagnostischer Spezifität.

Negativ prädiktiver Wert Anteil Individuen mit richtig negativem Testresultat, gegenüber allen im Test negativen Individuen: $RN/(RN+FN)$.

Gruppe (tatsächlich)	Predicted Group	
	krank	Gesund
krank	RP (richtig positiv)	FN (falsch negativ)
gesund	FP (falsch positiv)	RN (richtig negativ)

Prävalenz Die Prävalenz bezieht sich auf das Vorhandensein eines Merkmals in einer Population (Populationsprävalenz). Sie wäre beispielsweise die epidemiologische Häufigkeit aller Fälle einer bestimmten Krankheit in einer Population zum Zeitpunkt der Untersuchung. Die Prävalenz ist für die Qualität der Vorhersage eines Verfahrens von Bedeutung. Dass das Testergebnis nicht nur von den Eigenschaften des Tests (hier: von der so genannten Diagnosefähigkeit der Lehrkräfte) abhängt, sondern eng mit der Prävalenz in Beziehung steht, zeigt folgende Rechnung aus dem medizinischen Bereich.

3.6.2 Beispiel

Angenommen sei bei $N = 100.000$ Personen ein Anteil von einmal 1.000 erkrankten Individuen (Prävalenz 1%), und ein zweites Mal von 10.000 erkrankten Individuen (Prävalenz 10%) sowie ein Anteil $RP = RN$ eines Testverfahrens von 95%. Damit ergibt sich folgende Rechnung:

N = 100.000 RP = RN = 95%	<u>Diagnose</u>		
	Positive Diagnose	Negative Diagnose	Summe
Prävalenz 10 %	Krank	9500	500
	Gesund	4500	85500
	Summe	14000	86000
Prävalenz 1%	Krank	950	50
	Gesund	4950	94050
	Summe	5900	94100
WKT bei 10%		67,86%	
WKT bei 1%		16,10%	

Obwohl das Testverfahren in beiden Fällen (Prävalenz 1%, 10%) mit $RP = RN = 95\%$ einen gleich hohen Wert aufweist, folgen zwei deutlich verschiedene Ergebnisse. Die Wahrscheinlichkeit einer Erkrankung unter der Voraussetzung einer positiven Diagnose (positiv prädiktiver Wert) beläuft sich im ersten Fall bei einer Prävalenz von 10% auf 67,86%, bei einer Prävalenz von 1% hingegen auf nur 16,1%. Obwohl das Testverfahren mit $RP = RN = 95\%$ einen sehr hohen Wert aufweist, ist die Wahrscheinlichkeit einer Erkrankung unter der Bedingung einer positiven Diagnose deutlich geringer als die meßtheoretischen Eigenschaften des Verfahrens zunächst erwarten lassen.

Das Ergebnis eines Tests ist also nicht nur Funktion seiner Sensitivität oder Spezifität, sondern ebenfalls abhängig von Eigenschaften der Testpopulation. Nicht immer sind Kenntnisse a priori vorhanden, so dass prinzipiell nicht notwendigerweise vom Testergebnis auf die Qualität des Tests, sozusagen rückwärts, geschlossen werden kann. Dies wird an oben dargestellter Vergleichsrechnung, in der die Testparameter bei unterschiedlicher Prävalenz konstant sind, deutlich. Der Zusammenhang lässt sich auch funktional wie in folgender Graphik darstellen.

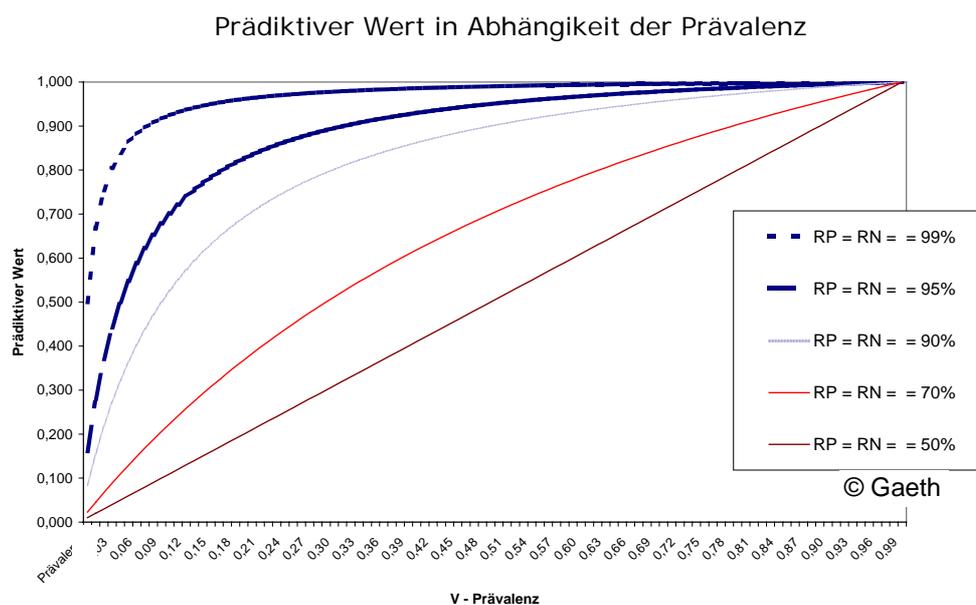


Abbildung 70 Prädiktiver Wert in Abhängigkeit von der Prävalenz bei verschiedenen Sensitivitäten und Spezifitäten

Die Graphik stellt in Abhängigkeit der Prävalenzen von 1% bis 100% die Kurvenverläufe der positiv prädiktiven Werte dar, wobei der Einfachheit halber jeweils $RP = RN$ gesetzt wurde bei stufenweiser Variation (50%, 70%, 90%, 99%). Deutlich wird, dass bei niedriger Prävalenz sehr hohe Sensitivität/Spezifität einer korrekten Klassifikation

vorausgesetzt werden muss, um bei positivem/negativem Testausfall auch ein positives/negatives Ergebnis erwarten zu können.

Bei einer Prävalenz von 1% liegt die Wahrscheinlichkeit einer Erkrankung unter der Bedingung eines positiven Testausfalls unter Voraussetzung von $RP = RN = 95\%$ nur knapp über 16%. Bei $RP = RN = 50\%$ entspricht das Testergebnis mit einer Wahrscheinlichkeit von 1% genau einer Zufallsziehung (bezogen auf die Prävalenz von 1%). Im Falle einer Prävalenz von 10% muss das Testverfahren eine Sensitivität und Spezifität von 99% aufweisen, damit ein positiv klassifizierter Schüler mit einer Wahrscheinlichkeit von 91,7% tatsächlich leseschwach ist. Bereits eine Reduzierung von $RP = RN = 99\%$ auf $RP = RN = 95\%$, also um lediglich 4 Prozentpunkte, senkt diese Wahrscheinlichkeit auf 67,9%. Eine weitere Senkung auf $RP = RN = 90\%$ (ein immerhin noch sehr hoher Wert) reduziert die korrekte Klassifikation auf 50% und entspricht damit dem Münzwurf (Laplace).

Damit ist die Betrachtung jedoch noch nicht vollständig. In den vorangegangenen Beispielen wird generell von einer Gleichheit der Anteile RP und RN und damit der Sensitivität und Spezifität ausgegangen. Tatsächlich aber sind Sensitivität und Spezifität konkurrierende Größen. Die Sensitivität sinkt in dem Maße, in dem die Spezifität steigt. Das Festlegen der Sensitivität und Spezifität ist letztendlich immer eine Optimierungsaufgabe. In den Ausführungen des deutschen PISA-Konsortiums finden sich jedoch keine näheren Darstellungen, weder zur Sensitivität noch zur Spezifität des PISA-Testkonstrukts noch zur ROC-Kurve (receiver operating characteristic) zur Bestimmung eines optimalen Trennwerts (cutoff-value). Im Technical Report konnte ebenfalls keine Information zur Problematik gefunden werden.

3.7 Gesamtschule

Gegenstand des Interesses ist nun die berichtete Unterschiedlichkeit hinsichtlich sozialer Zusammensetzung und Testergebnis zwischen den einzelnen Schulformen, insbesondere die der Gesamtschule im Verhältnis zum gegliederten Schulsystem.

*Zunächst wird deutlich, dass entgegen weit verbreiteter Vorstellungen das Gymnasium mittlerweile die sozial heterogenste Schulform ist, aber dennoch die hinsichtlich ihrer kognitiven Grundfähigkeiten homogenste Schülerschaft besitzt. Das Pendant stellt die Integrierte Gesamtschule dar, deren Schülerschaft sozial homogener, aber in ihren kognitiven Grundfähigkeiten weitaus heterogener ist. Die Leistungsheterogenität der Schülerschaft dieser Schulform, die sie deutlich von den anderen Bildungsgängen unterscheidet, hat unmittelbare Konsequenzen für die Kopplung von Sozialschichtzugehörigkeit und Leistungsergebnissen. Die Kovariation von sozialer Herkunft und erworbenen Kompetenzen ist an Gesamtschulen weitaus enger, als dies an anderen Schulformen der Fall ist.*¹⁴⁰

Baumert und Schümer vergleichen Schulformen, nicht jedoch Schüler mit ihren jeweils verschiedenen Abschlusszielen (Schülerinnen und Schüler aus *Schulen mit mehreren Bildungsgängen* wurden auf die Schulformen je nach Bildungsgang aufgeteilt).¹⁴¹ Da integrierte Systeme Schüler unterschiedlicher gewünschter Schulabschlüsse aufweisen, wäre die Berücksichtigung der Abschlussziele sinnvoll. Es wären Schüler zu vergleichen, die übereinstimmende Abschlüsse in den unterschiedlichen Schulformen anstreben, beispielsweise Schüler mit Abitur als Abschlussziel im gegliederten und integrierten Schulsystem. Der Vergleich von Schulformen hingegen umfaßt in integrierten Systemen Schüler unterschiedlicher Abschlussziele. Die Testergebnisse hingen lediglich von der unterschiedlich gewichteten Zusammensetzung der Schülerschaft ab.

¹⁴⁰ Baumert, Schümer: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 371

¹⁴¹ Artelt, Stanat, Schneider und Schiefele: Lesekompetenz: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 120

Lesekompetenz nach Bildungsgang

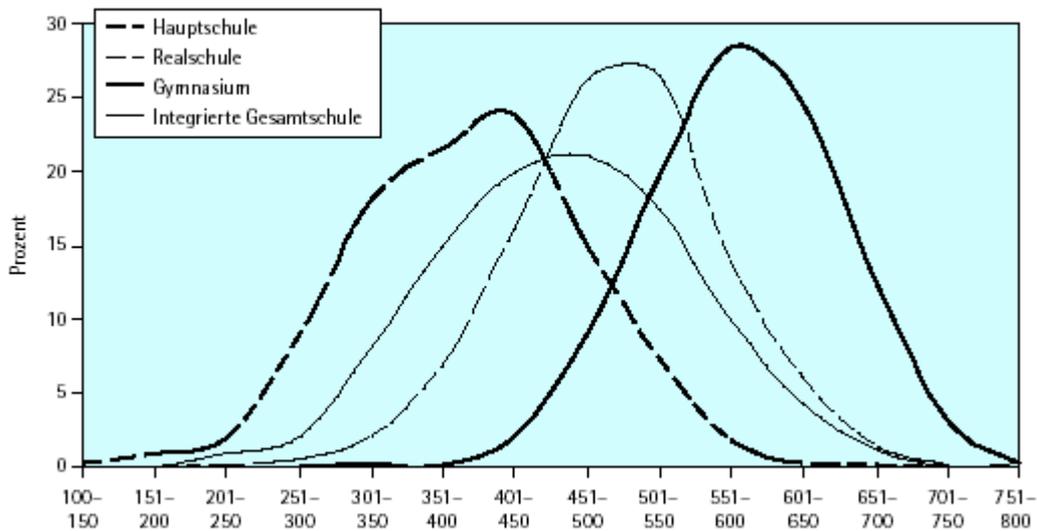


Abbildung 71 Bildungsgang und Testleistung (Quelle: PISA 2000)

Baumert und Schümer lösen sich mit ihrer Vorgehensweise vom Ansatz der Dritten Internationalen Mathematik und Naturwissenschaftsstudie (TIMSS).¹⁴²

Allgemein bildende Schulabschlüsse sind im deutschen Schulsystem nicht notwendigerweise an bestimmte Schulformen gebunden. Dass nennenswerte Anteile eines Altersjahrgangs an Realschulen und insbesondere Gymnasien nach dem Absolvieren der Vollzeitschulpflicht das jeweilige Schulziel nicht erreichen und mit niedrigerem Abschluss die Schule verlassen, ist bekannt. In Regionen mit schwach ausgebauten Realschulen ist das Gymnasium darüber hinaus immer noch ein Äquivalent der mittleren Schulform (Arbeitsgruppe Bildungsbericht, 1994; Köller, Baumert & Schnabel, 1999). Ein beträchtlicher Teil des Jahrgangs verlässt nach dem erfolgreichen Abschluss der 10. Klasse das Gymnasium mit dem Realschulabschluss - und zwar in Übereinstimmung mit den individuellen Bildungsplänen. Darüber hinaus aber wurde der Zusammenhang zwischen Schulabschluss und Schulform auch durch bildungspolitische Maßnahmen entkoppelt. Der Realschulabschluss und die Fachoberschulreife können,

¹⁴² Watermann, Baumert: Mathematische und naturwissenschaftliche Grundbildung beim Übergang von der Schule in den Beruf. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 205

wenngleich nicht in allen Bundesländern, auch an Hauptschulen, Gesamtschulen und in beruflichen Bildungsgängen erworben werden.¹⁴³

Die Codierung der Variablen *Bildungsgang* erscheint zumindest in den Public Use Files der Ergänzungsstudie dem gegenüber als nicht konsistent. Die folgende Abbildung zeigt als Kreuztabelle die Variablen *Bildungsgang* und *Schulform*. Die Codierung wurde im Original übernommen.

Schulform * Bildungsgang Kreuztabelle

			Bildungsgang				Gesamt
			HS	RS	GY	IGS	
Schulform	Hauptschule	% von Schulform	99,9%	,1%			100,0%
		% von Bildungsgang	90,5%	,0%			23,1%
	Schule mit mehreren Bildungsgängen	% von Schulform	24,3%	74,7%		1,0%	100,0%
		% von Bildungsgang	9,3%	20,3%		1,1%	9,7%
	Realschule	% von Schulform	,2%	99,8%			100,0%
		% von Bildungsgang	,2%	79,6%			28,5%
	Integrierte Gesamtschule	% von Schulform				100,0%	100,0%
		% von Bildungsgang				98,9%	8,5%
	Gymnasium	% von Schulform			100,0%		100,0%
		% von Bildungsgang			100,0%		30,1%
Gesamt		% von Schulform	25,5%	35,8%	30,1%	8,6%	100,0%
		% von Bildungsgang	100,0%	100,0%	100,0%	100,0%	100,0%

Abbildung 72 Bildungsgang - Schulform

Schüler des Gymnasiums befinden sich hier zu 100% im *Bildungsgang GY*. Schüler der Realschulen befinden sich zu 0,2% im *Bildungsgang HS* und zu 99,8% im *Bildungsgang RS*. Unter Berücksichtigung des geplanten Schulabschlusses hingegen zeigt sich ein differenziertes Bild.

¹⁴³ Watermann, Baumert: Mathematische und naturwissenschaftliche Grundbildung beim Übergang von der Schule in den Beruf. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 205

Schulform * Schulabschluss Kreuztabelle

			Schulabschluss					Gesamt
			HS nach 9. Kl	HS nach 10. Kl	RS, Fachober schule	Fachhoch schulreife	Abitur	
Schulform	Hauptschule	% von Schulform	43,5%	22,9%	27,1%	3,2%	3,3%	100,0%
		% von Schulabschluss	81,1%	63,4%	14,4%	12,9%	1,8%	20,8%
	Schule mit mehreren Bildungsgängen	% von Schulform	15,4%	8,2%	70,0%	2,4%	4,0%	100,0%
		% von Schulabschluss	13,7%	10,8%	17,8%	4,7%	1,1%	9,9%
	Realschule	% von Schulform	,4%	1,5%	77,2%	9,3%	11,6%	100,0%
		% von Schulabschluss	1,2%	5,9%	57,8%	53,0%	9,2%	29,3%
	Integrierte Gesamtschule	% von Schulform	5,4%	17,3%	36,5%	9,7%	31,2%	100,0%
		% von Schulabschluss	3,9%	18,5%	7,5%	15,2%	6,8%	8,0%
	Gymnasium	% von Schulform	,0%	,3%	3,2%	2,3%	94,0%	100,0%
		% von Schulabschluss	,1%	1,4%	2,6%	14,2%	81,1%	31,9%
Gesamt		% von Schulform	11,2%	7,5%	39,2%	5,1%	37,0%	100,0%
		% von Schulabschluss	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Abbildung 73 Schulform - Schulabschluss

Im Folgenden sollen der zweiten Alternative entsprechend Schülerinnen und Schüler mit gleichen Abschlusszielen verglichen werden, wobei sich die Frage stellt, ob sich in der Gesamtschule im Verhältnis zum gegliederten Schulsystem ein stärkerer Zusammenhang zwischen sozioökonomischem Status der Schüler und ihrer Testleistung nachweisen lässt. Als Variable kommt einerseits der höchste sozioökonomische Index der Familie (X_1), andererseits die Testleistung (X_2) in Frage. Berechnet werden zunächst die Korrelationskoeffizienten zwischen beiden Variablen bei gleichem Abschlußwunsch jeweils für diejenigen Schüler im dreigliedrigen Schulsystem ($r_{p(\text{gegliedert})} = r_1$) und in der Integrierten Gesamtschule ($r_{p(\text{Gesamtschule})} = r_2$). Sodann sind die Korrelationskoeffizienten der beiden Gruppen zu vergleichen. Geprüft wird die Hypothese der Gleichheit der Korrelation: $H_0 : \rho_1 = \rho_2$ gegen $H_1 : \rho_1 \neq \rho_2$.

Gestetet wird über die Fishers-Z-Transformation $z_{Fisher} = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$

mit $Z = \frac{z_{Fisher1} - z_{Fisher2}}{\sigma_{(z_{Fisher1} - z_{Fisher2})}}$ und $\sigma_{(z_{Fisher1} - z_{Fisher2})} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$.¹⁴⁴

Dargestellt werden zunächst die Korrelationskoeffizienten getrennt nach Abschlußziel und Schulform.

¹⁴⁴ Bortz: Statistik, 1993, S.202 ff

3.7.1 Abschluss: Hauptschule nach 9. Klasse

Schulform recodiert			Reading Gesamtscore 1. PV Landesspezifische Metrik	Mathe nat. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften Biologie 1. PV Landesspezifische Metrik	Mathe int. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften int. Gesamtscore 1. PV Landesspezifische Metrik
Haupt - Real - Gymn	Höchster SES in der Familie	Korrelation nach Pearson	0,136	0,067	0,087	0,068	0,136
		Signifikanz (2-seitig)	0,000	0,012	0,001	0,058	0,000
		N	1495	1392	1432	777	828
Integrierte Gesamtschule	Höchster SES in der Familie	Korrelation nach Pearson	-0,020	0,012	-0,048	0,137	-0,058
		Signifikanz (2-seitig)	0,769	0,862	0,499	0,147	0,547
		N	208	200	201	113	111
Fisher z (1)			0,136	0,067	0,087	0,068	0,137
Fisher z (2)			-0,020	0,012	-0,048	0,138	-0,058
Z			2,107	0,722	1,782	0,689	1,902
alpha			1,75%	23,50%	3,73%	24,55%	2,86%

Abbildung 74 Signifikanzberechnung Split: Hauptschule nach 9. Klasse

3.7.2 Abschluss: Hauptschule nach 10. Klasse

Schulform recodiert			Reading Gesamtscore 1. PV Landesspezifische Metrik	Mathe nat. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften Biologie 1. PV Landesspezifische Metrik	Mathe int. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften int. Gesamtscore 1. PV Landesspezifische Metrik
Haupt - Real - Gymn	Höchster SES in der Familie	Korrelation nach Pearson	0,158	0,080	0,140	0,151	0,139
		Signifikanz (2-seitig)	0,000	0,016	0,000	0,001	0,001
		N	952	899	932	509	548
Integrierte Gesamtschule	Höchster SES in der Familie	Korrelation nach Pearson	0,073	0,045	0,036	0,082	0,049
		Signifikanz (2-seitig)	0,190	0,433	0,535	0,274	0,514
		N	323	300	307	180	181
Fisher z (1)			0,159	0,081	0,141	0,152	0,140
Fisher z (2)			0,073	0,045	0,036	0,082	0,049
Z			1,330	0,525	1,592	0,803	1,057
alpha			9,18%	29,97%	5,57%	21,10%	14,52%

Abbildung 75 Signifikanzberechnung Split: Hauptschule nach 10. Klasse

3.7.3 Abschluss: RS Fachoberschule

Schulform recodiert			Reading Gesamtscore 1. PV Landesspezifische Metrik	Mathe nat. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften Biologie 1. PV Landesspezifische Metrik	Mathe int. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften int. Gesamtscore 1. PV Landesspezifische Metrik
Haupt - Real - Gymn	Höchster SES in der Familie	Korrelation nach Pearson	0,115	0,123	0,111	0,110	0,138
		Signifikanz (2-seitig)	0,000	0,000	0,000	0,000	0,000
		N	6170	5784	5922	3295	3385
Integrierte Gesamtschule	Höchster SES in der Familie	Korrelation nach Pearson	0,068	0,055	0,059	0,118	0,058
		Signifikanz (2-seitig)	0,034	0,092	0,066	0,007	0,176
		N	981	944	970	521	552
Fisher z (1)			0,116	0,124	0,112	0,111	0,139
Fisher z (2)			0,068	0,055	0,059	0,118	0,058
Z			1,392	1,954	1,520	0,157	1,773
alpha			8,20%	2,54%	6,42%	43,77%	3,81%

Abbildung 76 Signifikanzberechnung Split: Abschluss: RS Fachoberschule

3.7.4 Abschluss: Abitur

Schulform recodiert			Reading Gesamtscore 1. PV Landesspezifische Metrik	Mathe nat. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften Biologie 1. PV Landesspezifische Metrik	Mathe int. Gesamtscore 1. PV Landesspezifische Metrik	Naturwissenschaften int. Gesamtscore 1. PV Landesspezifische Metrik
Haupt - Real - Gymn	Höchster SES in der Familie	Korrelation nach Pearson	0,184	0,173	0,181	0,168	0,177
		Signifikanz (2-seitig)	0,000	0,000	0,000	0,000	0,000
		N	9331	8669	8923	4845	5185
Integrierte Gesamtschule	Höchster SES in der Familie	Korrelation nach Pearson	0,198	0,203	0,161	0,221	0,184
		Signifikanz (2-seitig)	0,000	0,000	0,000	0,000	0,000
		N	745	698	723	389	434
Fisher z (1)			0,186	0,175	0,183	0,170	0,179
Fisher z (2)			0,200	0,206	0,162	0,224	0,186
Z			0,379	0,783	0,538	1,029	0,146
alpha			35,24%	21,68%	29,52%	15,17%	44,21%

Abbildung 77 Signifikanzberechnung Split: Abschluss: Abitur

Die Ergebnisse sind für die vier Abschlüsse (wie im Datensatz codiert) dargestellt: Hauptschulabschluss nach der 9. Klasse, Hauptschulabschluss nach der 10. Klasse, Realschulabschluss und Fachoberschule, Abitur. Die Korrelationen sind paarweise berechnet: Der jeweils obere Teil der Tabelle stellt für die einzelnen Tests die Korrelation zwischen sozioökonomischem Index und Testleistung für Schülerinnen und Schüler des gegliederten Schulsystems dar, der untere Teil der Tabelle (Schulform recodiert = Integrierte Gesamtschule) jeweils den Zusammenhang für Schülerinnen und Schüler der Gesamtschule mit gleichem Abschlusswunsch. Die Korrelationskoeffizienten sind inklusive Signifikanztest und jeweiligem Stichprobenumfang dargestellt.

Erkennbar ist die allgemein höhere Korrelation zwischen sozialer Herkunft (gemessen am höchsten sozioökonomischen Index der Familie) und Testleistung Lesen im (gesamten) dreigliedrigen Schulsystem. Beispiel: Es weisen die 1495 Schülerinnen und Schüler des dreigliedrigen Schulsystems, die einen Hauptschulabschluss nach der 9. Klasse anstreben, gegenüber den 208 Gesamtschülern, die den gleichen Abschluss anstreben, eine höhere Korrelation zwischen sozialer Herkunft und Testleistung im Lesen auf. Der Unterschied von $r_p = 0,136$ bei $N=1495$ zu $r_p = -0,02$ bei $N=208$ ist mit $p=1,75\%$ statistisch signifikant. Einige Vergleiche kommen zu einem nichtsignifikanten Ergebnis. Für die Schüler mit Abschlusswunsch Abitur ergeben sich zwischen dreigliedrigem Schulsystem und Gesamtschule keine statistisch signifikanten Unterschiede.

3.8 Soziale Heterogenität

In einem zweiten Schritt soll die Hypothese der *sozialen Heterogenität* geprüft werden. Die Frage ist nun, ob bei Berücksichtigung des Perspektivenwechsels *soziale Heterogenität* zwischen den Vergleichsgruppen (Gesamtschüler und Schüler aus dem dreigliedrigen Schulsystem) nachweisbar unterschiedlich ist. Als Test wird hier analog zur PISA-Studie der Levene-Test auf Gleichheit der Varianzen entsprechend folgender Hypothese gewählt

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ gegen } H_1 : \sigma_1^2 \neq \sigma_2^2$$

Die Levene Statistik ist wie folgt definiert:

$$L = \frac{(W - 2) \sum_{k=1}^2 W_k (\bar{Z}_k - \bar{Z})^2}{\sum_{k=1}^2 \sum_{i=1}^{n_k} w_{ki} (Z_{ki} - \bar{Z}_k)^2}, \text{ wobei } Z_{ki} = |X_{ki} - \bar{X}_k|, \bar{Z}_k = \frac{\sum_{i=1}^{n_k} w_{ki} Z_{ki}}{W_k}, \bar{Z} = \frac{\sum_{k=1}^2 W_k \bar{Z}_k}{W_1 + W_2}$$

Von Interesse ist einerseits der Mittelwertunterschied des höchsten sozioökonomischen Index in der Familie im Vergleich der beiden Gruppen (dreigliedriges Schulsystem/Integrierte Gesamtschule) mit Abschlussziel Abitur, andererseits die Streuungsunterschiede (soziale Heterogenität).

Gruppenstatistiken

Schulform recodiert		N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Höchster SES in der Familie	Haupt - Real - Gymn	9363	56,13	15,872	,164
	Integrierte Gesamtschule	755	51,89	15,173	,552

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
Höchster SES in der Familie	Varianzen sind gleich	3,860	,049	7,092	10116	,000	4,25	,599	Untere	Obere
	Varianzen sind nicht gleich			7,369	892,374	,000	4,25	,576	3,115	5,376

Abbildung 78 Soziale Heterogenität/Schulform

Das Ergebnis zeigt einen statistisch signifikant geringeren durchschnittlichen sozioökonomischen Index derjenigen Schüler mit Abitur als Abschlussziel in der Integrierten Gesamtschule ($\bar{x}_{gym} = 56,13$ zu $\bar{x}_{gs} = 51,89$; $p < 0,01$). Die Hypothese der Varianzgleichheit lässt sich zwar mit einem Fehler erster Ordnung von 4,9% nach Konvention verwerfen (konventionelle Grenze: 5%). Die nachgewiesene Streuung ist demnach im (gesamten) dreigliedrigen Schulsystem geringfügig höher als in der Gruppe der Gesamtschüler. In Anbetracht der geringen Unterschiede ($s = 15,872$ zu $s = 15,173$) stellt sich jedoch die Frage nach der Relevanz.