

© Gaeth 2005
alle Rechte vorbehalten



PISA
(Programme for International Student Assessment)
Eine statistisch-methodische Evaluation



Frank Gaeth

PISA
(Programme for International Student Assessment)
Eine statistisch-methodische Evaluation

Inaugural-Dissertation
Freie Universität Berlin

vorgelegt von Frank Gaeth

Berlin 2005

DANKSAGUNG

Natürlich ist es nicht möglich, ein solches Projekt ohne vielfältige Unterstützung erfolgreich zu beenden. Deshalb möchte ich diese Gelegenheit dazu nutzen, mich ganz herzlich bei allen zu bedanken.

Herrn Prof. Massing danke ich für die keineswegs selbstverständliche Bereitschaft zur Annahme meiner Arbeit. Ich hoffe, zu einem objektiven Bild der Leistungen von Lehrerinnen und Lehrern und der sie ausbildenden Institutionen und Personen beitragen zu können.

Mein Dank gilt Herrn Prof. Büning. Die Seminare am Institut für Statistik und Ökonometrie sowie die Diskussionen zu den Verfahren und Methoden waren ein entscheidender Bestandteil des erfolgreichen Abschlusses dieser Arbeit.

Mein Dank gilt Herrn Prof. Jürgens. Gern denke ich an das eine oder andere Projekt zurück.

Mein Dank gilt Herrn Prof. Raschert. Ich hoffe mit dieser Arbeit einen kleinen Beitrag für die Erziehungswissenschaft leisten zu können.

Nicht zuletzt aber möchte ich mich ganz besonders bei Herrn Akad. Rat Dr. Ryll bedanken. Hier hatte das Projekt begonnen, hier fand es schließlich einen guten Abschluß. Danke.

meiner Familie

Inhaltsverzeichnis

1	EINLEITUNG	17
1.1	GEGENSTAND DER EVALUATION	17
1.2	STUDIE	18
1.2.1	Untersuchungsgegenstand und Durchführung der PISA-Studie.....	18
1.2.2	Ziele von PISA.....	18
1.2.3	Testkonzeption	19
1.2.4	Projekt	20
1.2.5	Koordination	20
1.2.6	Kontroverse	21
2	EVALUATION	21
2.1	STICHPROBENDEFINITIONEN INTERNATIONALER VERGLEICHSTUDIEN	21
2.1.1	TIMSS	22
2.1.1.1	Population I	22
2.1.1.2	Population II	22
2.1.1.3	Population III	22
2.1.2	PISA	25
2.1.3	IGLU	26
2.2	PISA – VERGLEICHBARKEIT DER POPULATIONEN INNERHALB DER STUDIE.....	26
2.2.1	Korrelation.....	30
2.2.2	Regression.....	35
2.2.3	Varianzanalytischer Ansatz.....	36
2.3	BEWERTUNG.....	45
2.4	ALTERNATIVER ANSATZ	45
2.4.1	Vergleichbarkeit der Teilnehmerstaaten.....	45
2.4.2	Rückhaltequote (Retentivität)	47
2.4.3	Ergänzungsunterricht	54
2.4.4	Gewichtung.....	55
2.4.5	Darstellung der Testwerte	57
2.4.6	Interpretation.....	62
2.4.7	Zwischenbewertung.....	63
2.5	SOZIALE DISPARITÄT IN DER LESEKOMPETENZ.....	68
2.5.1	Streuung des Schulalters	70
2.5.2	Bodeneffekte/Deckeneffekte	72
2.6	MIGRATION/DURCHSCHNITTLICHES SCHULALTER	80
2.6.1	Chi-Quadrat-Test/Klassenstufe.....	83
2.7	VERGLEICH IGLU/PISA.....	85

3	BUNDESDEUTSCHER VERGLEICH.....	87
3.1	TESTLEISTUNG BUNDESLAND	87
3.2	NICHTLESER UND TESTLEISTUNG.....	90
3.3	STAGE MIGRATION	97
3.4	MIGRATIONSHINTERGRUND	104
3.5	RISIKOSCHÜLER.....	107
3.5.1	Risikoschüler – Leseschwäche (I)	108
3.5.2	Risikoschüler – Leseschwäche (II)	110
3.5.2.1	Hauptschule.....	111
3.5.2.2	Realschule	112
3.5.2.3	Gymnasium	113
3.5.3	Risikoschüler - Familienstruktur	114
3.5.4	Schlussfolgerung.....	125
3.6	DIAGNOSE UND TRENNSCHÄRFE	126
3.6.1	Begriffsdefinitionen	126
3.6.2	Beispiel	127
3.7	GESAMTSCHULE	130
3.7.1	Abschluss: Hauptschule nach 9. Klasse	134
3.7.2	Abschluss: Hauptschule nach 10. Klasse.....	134
3.7.3	Abschluss: RS Fachoberschule	135
3.7.4	Abschluss: Abitur	135
3.8	SOZIALE HETEROGENITÄT.....	136
4	ZUSAMMENFASSUNG DER ERGEBNISSE	138
4.1	INTERNATIONAL	138
4.2	NATIONAL (BUNDESREPUBLIK)	139
5	ABSCHLIEBENDE BEWERTUNG	140
5.1	INTERNATIONAL	140
5.2	NATIONAL (BUNDESREPUBLIK)	143
5.3	WEITERFÜHRENDE PERSPEKTIVEN.....	144
5.3.1	Testkonstruktion.....	144
5.3.2	Curriculare Validität.....	145
5.4	RESÜMEE	145
6	ANHANG TABELLEN	146
7.	DOKUMENTATION (PROGRAMMBEISPIELE).....	160
8.	LITERATUR.....	165

Abbildungsverzeichnis

ABBILDUNG 1 TARGET POPULATION TIMSS/III, BERECHNUNG UNGEWICHTET	24
ABBILDUNG 2 VERTEILUNG KLASSENSTUFE	27
ABBILDUNG 3 KORRELATION KLASSENSTUFE LESEWERT FALLBASIERTE	30
ABBILDUNG 4 KORRELATION KLASSENSTUFE LESETESTWERT JE TEILNEHMERSTAAT (I)	32
ABBILDUNG 5 KORRELATION KLASSENSTUFE LESETESTWERT JE TEILNEHMERSTAAT (II)	33
ABBILDUNG 6 KORRELATION KLASSENSTUFE LESETESTWERT (AGGREGIERT)	34
ABBILDUNG 7 REGRESSION KLASSENSTUFE, GESCHLECHT, SOZIOÖKON. INDEX AUF LESETESTWERT ..	35
ABBILDUNG 8 VARIANZANALYSEN KLASSENSTUFE UND LESETESTWERT	42
ABBILDUNG 9 VARIANZANALYSE KLASSENSTUFE LESETESTWERT/SCHEFFÉ-TEST, DEUTSCHLAND	43
ABBILDUNGEN 10 SCHEFFÉ-TEST/DIAGRAMM DER MITTELWERTE KLASSENSTUFE/TESTWERT LESEN ...	44
ABBILDUNG 11 VERTEILUNG GESCHLECHT/TEILNEHMERSTAAT (ANGABEN IN PROZENT)	46
ABBILDUNG 12 TCI-INDEX LESETESTWERT	49
ABBILDUNG 13 TCI – LESETESTWERT/KORRELATION.....	50
ABBILDUNG 14 KORRELATION TIMSS TESTWERT – TCI - BSP	51
ABBILDUNG 15 RÜCKHALTEQUOTE UNTER 15-JÄHRIGEN IN PISA	52
ABBILDUNG 16 RÜCKHALTEQUOTE (RETENTIVITÄT) UND TESTLEISTUNG IN PISA (I)	53
ABBILDUNG 17 RÜCKHALTEQUOTE (RETENTIVITÄT) UND TESTLEISTUNG IN PISA (II)	54
ABBILDUNG 18 ERGÄNZUNGSUNTERRICHT	55
ABBILDUNG 19 GEWICHTUNG/DIFFERENZ DER TESTERGEBNISSE.....	57
ABBILDUNG 20 LESEVERSTÄNDNIS KLASSENSTUFE 11 (GEWICHTET)	58
ABBILDUNG 21 LESEVERSTÄNDNIS KLASSENSTUFE 10 (GEWICHTET)	58
ABBILDUNG 22 LESEVERSTÄNDNIS KLASSENSTUFE 9 (GEWICHTET).....	59
ABBILDUNG 23 LESEVERSTÄNDNIS (RANDMITTEL UNGEWICHTET)	59
ABBILDUNG 24 LESEVERSTÄNDNIS (RANDMITTEL - DATENSATZ GEWICHTET)	60
ABBILDUNG 25 SCIENCE (RANDMITTEL UNGEWICHTET).....	60
ABBILDUNG 26 SCIENCE (RANDMITTEL - DATENSATZ GEWICHTET)	61
ABBILDUNG 27 MATH (RANDMITTEL UNGEWICHTET)	61
ABBILDUNG 28 MATH (RANDMITTEL - DATENSATZ GEWICHTET).....	62
ABBILDUNG 29 KLASSENGRÖÖE – EXTRA UNTERRICHT	67
ABBILDUNG 30 LERNGELEGENHEITEN – SOZ.ÖKON. INDEX/KORRELATION	68
ABBILDUNG 31 DISPARITÄT	69
ABBILDUNG 32 DISPARITÄT IM LESEN – STREUUNG DES SCHULALTERS	71
ABBILDUNG 33 DISPARITÄT - SCHIEFE TESTWERTEVERTEILUNG (BODENEFFEKT)	72
ABBILDUNG 34 DISPARITÄT - SYMMETRISCHE TESTWERTEVERTEILUNG	73
ABBILDUNG 35 DECKENEFFEKT/BODENEFFEKT (ABBILDUNG I)	74
ABBILDUNG 36 DECKENEFFEKT/BODENEFFEKT (ABBILDUNG II)	75

ABBILDUNG 37 DISPARITÄT/KORRELATION.....	76
ABBILDUNG 38 DISPARITÄT LESETESTWERT KLASSENSTUFE (GRADE) 9.....	76
ABBILDUNG 39 DISPARITÄT DIFFERENZ/KORRELATION KLASSENSTUFE (GRADE) 9.....	77
ABBILDUNG 40 TESTWERT – SCHIEFE/KORRELATION (KLASSENSTUFE 10).....	78
ABBILDUNG 41 LESETESTWERT (SÄMTLICHE KLASSENSTUFEN) SCHIEFE DER VERTEILUNG.....	79
ABBILDUNG 42 SOZIOÖKONOMISCHE DIFFERENZ – SCHULALTER DIFFERENZ/KORRELATION.....	81
ABBILDUNG 43 CHI-QUADRAT-TEST/KLASSENSTUFE.....	84
ABBILDUNG 44 BILDUNGSBETEILIGUNG DER 15-JÄHRIGEN PISA TESTTEILNEHMER.....	88
ABBILDUNG 45 ANTEIL NICHTLESER UND MITTLERE LESEKOMPETENZ.....	91
ABBILDUNG 46 NICHTLESER IM INTERNATIONALEN VERGLEICH.....	92
ABBILDUNG 47 NICHTLESER.....	93
ABBILDUNG 48 NICHTLESER - TESTWERT LESEN INTERNATIONAL/KORRELATION.....	94
ABBILDUNG 49 NICHTLESER BERUFSSCHÜLER NATIONAL/KORRELATION.....	95
ABBILDUNG 50 STAGE MIGRATION.....	97
ABBILDUNG 51 STAGE MIGRATION.....	98
ABBILDUNG 52 SCHÜLER IN GYMNASIEN – MATHEMATISCHE KOMPETENZ.....	99
ABBILDUNG 53 PERZENTILBÄNDER TESTWERT LESEN GYMNASIEN.....	100
ABBILDUNG 54 ANTEIL GYMNASIASTEN – RELATIVER SOZIOÖKON. INDEX/KORRELATION.....	101
ABBILDUNG 55 ZUSAMMENHANG ZWISCHEN RELATIVEM GYMNASIALBESUCH UND LEISTUNGSNIVEAU.....	102
ABBILDUNG 56 RELATIVES SOZIOÖKONOMISCHES NIVEAU REALSCHULE/GYMNASIUM.....	103
ABBILDUNG 57 MIGRATION – SCHULFORM QUELLE: STATISTISCHES BUNDESAMT.....	106
ABBILDUNG 58 LESESCHWÄCHE DIAGNOSE (I) HAUPTSCHULE.....	109
ABBILDUNG 59 LESESCHWÄCHE DIAGNOSE (I) SONDERSCHULE.....	109
ABBILDUNG 60 LESESCHWÄCHE DIAGNOSE(III) HAUPTSCHULE.....	111
ABBILDUNG 61 LESESCHWÄCHE DIAGNOSE(III) REALSCHULE.....	112
ABBILDUNG 62 LESESCHWÄCHE DIAGNOSE(III) GYMNASIUM.....	113
ABBILDUNG 63 ANTEIL KERNFAMILIE IN ABHÄNGIGKEIT ZUM SOZIOÖKONOMISCHEN INDEX.....	115
ABBILDUNG 64 ANTEIL KERNFAMILIE IN ABHÄNGIGKEIT ZUM SOZIOÖKONOMISCHEN INDEX.....	115
ABBILDUNG 65 ANTEIL NEUER PARTNERSCHAFTEN.....	116
ABBILDUNG 66 HAUSHALTSTYP.....	117
ABBILDUNGEN 67 TESTLEISTUNG UND HAUSHALTSTYP, GESCHÄTZTE RANDMITTEL.....	121
ABBILDUNG 68 FAMILIE LÄNGSSCHNITTANALYTISCHE BETRACHTUNG.....	123
ABBILDUNG 69 FAMILIE LÄNGSSCHNITTANALYTISCHE BETRACHTUNG/WOHLSTANDSINDIKATOREN....	125
ABBILDUNG 70 PRÄDIKTIVER WERT IN ABHÄNGIGKEIT VON DER PRÄVALENZ.....	128
ABBILDUNG 71 BILDUNGSGANG UND TESTLEISTUNG.....	131
ABBILDUNG 72 BILDUNGSGANG - SCHULFORM.....	132
ABBILDUNG 73 SCHULFORM - SCHULABSCHLUSS.....	133
ABBILDUNG 74 SIGNIFIKANZBERECHNUNG SPLIT: HAUPTSCHULE NACH 9. KLASSE.....	134
ABBILDUNG 75 SIGNIFIKANZBERECHNUNG SPLIT: HAUPTSCHULE NACH 10. KLASSE.....	134

ABBILDUNG 76 SIGNIFIKANZBERECHNUNG SPLIT: ABSCHLUSS: RS FACHOBERSCHULE.....	135
ABBILDUNG 77 SIGNIFIKANZBERECHNUNG SPLIT: ABSCHLUSS: ABITUR.....	135
ABBILDUNG 78 SOZIALE HETEROGENITÄT/SCHULFORM.....	137

Tabellenanhang

TABELLE 1	PISA 2000 LESEVERSTÄNDNIS (DATENSATZ GEWICHTET)	146
TABELLE 2	PISA 2000 LESEVERSTÄNDNIS (DATENSATZ UNGEWICHTET)	147
TABELLE 3	VERTEILUNG KLASSENSTUFE PISA 2000.....	148
TABELLE 4	PISA 2003 LESEVERSTÄNDNIS (DATENSATZ GEWICHTET)	149
TABELLE 5	PISA 2003 LESEVERSTÄNDNIS (DATENSATZ UNGEWICHTET)	150
TABELLE 6	PISA 2003 VERTEILUNG KLASSENSTUFE	151
TABELLE 7	PISA 2003 MATHEMATIKTEST (DATENSATZ GEWICHTET)	152
TABELLE 8	PISA 2003 MATHEMATIKTEST (DATENSATZ UNGEWICHTET)	153
TABELLE 9	PISA2000/PISA2003 LESETEST	154
TABELLE 10	TAB. L 204.....	155
TABELLE 11	TAB. L 204.....	156
TABELLE 12	BIP-LR1	157
TABELLE 13	AUSGEWÄHLTE KENNZIFFERN DER BILDUNGSSYSTEME	158
TABELLE 14	SOZIOÖKONOMISCHE DIFFERENZ UND SCHULALTERDIFFERENZ.....	159

Notation

x_i	Meßwert
\bar{x}	arithmetischer Mittelwert der Meßgröße
σ_x	Standardabweichung der Messung
μ	Lageparameter
$n(x)$	Häufigkeit des Merkmals x
H_0	Ausgangshypothese
H_1	Gegenhypothese
p	Fehler erster Ordnung (Signifikanz)
α_i	unbekannter Effekt des i-ten Treatments
π	Testanteil
π_i	$P(A_i)$, Wahrscheinlichkeit für Ausprägung A_i
π_{ij}	$P(A_i, B_j)$, Wahrscheinlichkeit für Ausprägungspaar (A_i, B_j)
F, G	Verteilungsfunktionen F, G
$E(X)$	Erwartungswert
$Var(X)$	Varianz von X, (σ_x^2)
$x_{0,25}, x_{0,5},$	
$x_{0,75}$	Quartile
β_i	Beta-Koeffizient im regressionsanalytischen Modell
r_p	Korrelationskoeffizient nach Pearson
r_s	Korrelationskoeffizient nach Spearman
ρ	Korrelation (rho)
\hat{y}	regressionsanalytisch bestimmter Funktionswert
K, N	Zahl der Beobachtungswerte, Stichprobenumfang
J	Zahl der Regressoren
R^2	Bestimmtheitsmaß
x_{\min}	kleinster Wert eines Merkmals
x_{\max}	grösster Wert eines Merkmals
$R(x_i)$	Rang des i-ten Meßwertes von X
d_i^2	Quadrat der i-ten Rangplatzdifferenz $d_i = R(x_i) - R(y_i)$
z_{Fisher}	transformierter Wert von r_p als Fisher-Z-Wert
Z	standardnormalverteilte ZV mit $Z \sim N(0,1)$
ζ	Erwartungswert der Fisher-z-transformierten Zufallsvariable ρ für kleine Stichproben
$\frac{\mu_3}{\sigma^3}$	Schiefe der Verteilung

Δ	Differenz
df	Freiheitsgrade (Degree of Freedom)
χ_{emp}^2	Wert des empirisch ermittelten Chi-Quadrat Wertes
$\chi_{df;1-\alpha}^2$	theoretischer Wert der durch Anzahl der Freiheitsgrade df definierten Chi-Quadrat Verteilung zum Niveau $1 - \alpha$
\tilde{n}_{ij}	erwartete Zellbesetzung in Zeile i, Spalte j
n_{ij}	Zellbesetzung in Zeile i und Spalte j
$t_{df;1-\alpha}$	Wert der durch Anzahl der Freiheitsgrade (df) bestimmten t-Verteilung zum Niveau $1 - \alpha$
u_{ij}	standardisiertes Residuum
t_0	Zeitpunkt 0, Ausgangszeit
t_1	Zeitpunkt 1
L	Levene Statistik
W_k	Summe der Fallgewichtungen in Gruppe k
n_k	Anzahl der Fälle in Gruppe k
x_{ki}	Wert des i-ten Falls in Gruppe k
w_i	Gewichtung von Fall i
w_{ki}	Gewichtung des i-ten Falls in Gruppe k
W	Summe der Gewichte

Abkürzungen

ACER	Australian Council for Educational Research
BC	Boston College, Lynch School of Education
BPC	Board of Participating Countries
IEA	International Association for the Evaluation of Educational Achievement
IGLU	Internationale Grundschul-Lese-Untersuchung
INES	Indicators of Education Systems
KMK	Kultusministerkonferenz der Länder
MPIB	Max-Planck-Institut für Bildungsforschung, Berlin
OECD	Organisation for Economic Co-operation and Development
PISA	Programme for International Students Assessment
TIMSS	Third International Mathematics and Science Study
WESTAT	Research Corporation, Rockville, Maryland

Variablen

AL_ERZ	Anteil Alleinerziehende
ANT_GYM	Gymnasialbesuch anteilig
B_GANG	Bildungsgang
BIP_JEWP	Bruttoinlandsprodukt zu jeweiligen Preisen
BIP_KONP	Bruttoinlandsprodukt zu konstanten Preisen
BSP_JEWP	Bruttosozialprodukt zu jeweiligen Preisen
BSP_KONP	Bruttosozialprodukt zu konstanten Preisen
CENTRAL	Anteil zentral geprüfter Schüler (PISA 2003)
COUNTRY – ISO CODE	Teilnehmerstaat
DIFF	Differenz der Mittelwerte (vgl. <i>Disparität</i>)
DIFF_9	Differenz der Mittelwerte Klassenstufe 9 (vgl. <i>Disparität</i>)
EXTRA	Ergänzungsunterricht in Prozent der Nennungen
G_DIFF	Differenz der durchschnittlichen Klassenstufen zwischen Schülern mit und ohne Migrationshintergrund
GRADE – Q2	Klassenstufe
GYM_Z	Durchschnittlicher relativer höchster sozioökonomischer Index der Gymnasiasten eines Bundeslandes, z-standardisiert
H_DIFF	Differenz der Mittelwerte des sozioökonomischen Index zwischen Schülern mit und ohne Migrationshintergrund
HISEI	Höchster sozioökonomischer Index (internationaler Datensatz)
ID-Stud	Identitätsnummer des Probanden (TIMSS)
ISEI	International Socio-Economic Index

ISEI_F	Höchster SES in der Familie
KL_GROES	Durchschnittliche Anzahl der Schüler pro Klasse
MEAN_TIM	Landesmittel mathematisch-naturwissenschaftliches Grundbildungsniveau in der TIMSS
Migration 2 Gruppen (4,8)	Migrationshintergrund Gruppen 4,8 (<i>MIG/Non_MIG 2 Gruppen</i>)
MIXED	Anteil <i>MIXED FAMILY</i> (neue Partnerschaften)
N_LESER	<i>NICHTLESER</i>
NUCLEAR	Anteil Kernfamilie
PV(1)READ	Lesetestwert PISA2000 International (Plausible value in reading) Wert (1) des mehrdimensionalen Rasch-Modells
PV10READ	Lesetestwert PISA2000 International (Plausible value in reading) Wert (1) des mehrdimensionalen Rasch-Modells, Auswahl 10. Klassenstufe
QUANT1	Arithmetisches Mittel der Testwerte über den Bereich $x_{\min} - x_{0,25}$ einer Verteilung
RETENT	Retentivität/Rückhaltequote
SCHIEFE, SKEWNESS	Schiefe der Verteilung der Testwerte eines Teilnehmerstaates
SES	Socio-Economic Status
SINGLE	Anteil Alleinerziehende
STD_G_N	Standardabweichung über Variable GRADE (Klassenstufe)
STD_G_W	Standardabweichung über Variable GRADE (Klassenstufe), Fälle gewichtet über Variable W_FSTUWT (Fallgewichtung)
TCI	TIMSS Coverage Index
VOLKEINK	Volkseinkommen
W_FSTUWT	Gewichtungsvariable

Hinweise

Die vorliegende Evaluation basiert auf den Public-Use Files als Download vom Server der OECD und der KMK. In Hinblick auf das im Test eingesetzte probabilistische Modell und der daraus resultierenden Probleme bezüglich der Auswertung mit in der Wissenschaft genutzten Standardverfahren und Programmen folgt die Evaluation den Empfehlungen des *Technical Report*, der Dokumentation der Erhebungsinstrumente, dem *Manual for the PISA Database*, der Vorgehensweise der PISA-Studie sowie der Rücksprache mit C. Artelt (MPI Berlin) und R. Adams (ACER Australien).

Auf die ausführliche Darstellung der statistischen Modelle und Verfahren wurde der Übersichtlichkeit und leichteren Lesbarkeit wegen nur jeweils einmal eingegangen.

Als *t-Test bei Ungleichheit der Varianzen* wird der Welch-Test eingesetzt.

Die Codierungen der Variablen wurden unverändert aus den Public-Use-Files der Studie übernommen.

Abbildungen wurden unter Angabe der Quelle ohne Berücksichtigung eventueller rechtlicher Beschränkungen in dieser Arbeit veröffentlicht. Die mangelhafte graphische Qualität einzelner Abbildungen folgt aus dem Kopierverfahren, insofern keine Alternative zur Verfügung stand.

Die Auswahl der in der vorliegenden Evaluation eingesetzten statistischen Verfahren orientiert sich an der PISA-Studie sowie weiteren thematisch verbundenen Vergleichsuntersuchungen (IGLU, TIMSS). Sämtliche nicht gesondert gekennzeichnete Tabellen und Abbildungen basieren auf eigenen Berechnungen.

Zum mathematischen Hintergrund der Berechnung geschätzter Randmittel siehe: SPSS Inc. Chicago: SPSS Statistical Algorithms GLM UNIVARIATE AND MULTI-VARIATE S.5 ff (<http://support.spss.com/tech/default.asp>).

Zur Messung des sozioökonomischen Index siehe: *PISA 2000 Dokumentation der Erhebungsinstrumente*, S.226.

1 Einleitung

1.1 Gegenstand der Evaluation

Die Kernthesen der PISA 2000-Studie, sowohl des internationalen Vergleichs als auch der nationalen Ergänzungsstudie, werden anhand der Public Use Files der TIMSS, PISA-I und PISA-E evaluiert. Der Fokus der vorliegenden Arbeit liegt ausschließlich auf den Statistiken, Verfahren und Methoden in Hinblick auf die zentralen Thesen. Der Fokus ist bewusst eng gewählt und beschränkt sich allein auf die Frage, ob mit dem erhobenen Datenmaterial die Kernthesen der beiden PISA-Studien zu belegen sind.

Die Kernthesen betreffen

- a. die geringe Performanz der Schülerinnen und Schüler der Bundesrepublik Deutschland,
- b. deren hohe soziale Disparitäten im Lesen,
- c. die über die Bundesländer systematisch differierenden Testergebnisse,
- d. die so genannte Diagnosefähigkeit der Lehrkräfte,
- e. den Zusammenhang von sozioökonomischem Hintergrund der Familie und Testleistung in der PISA-Studie im Vergleich der Schüler des integrierten und des gegliederten Systems,
- f. den rückläufigen Anteil der Kernfamilie an der Gesamtbevölkerung der Bundesrepublik Deutschland und dessen Bedeutung für die gemessenen Testergebnisse,
- g. den positiven Vorbildcharakter der Schulsysteme der Teilnehmerstaaten Japan und Korea,
- h. die einzelnen Studien (TIMSS, PISA und IGLU) als längsschnittanalytischer Vergleich zur Messung von Leistungsfortschritten,
- i. die Entwicklung einer international sinnvollen Stichprobendefinition.

1.2 Studie

PISA steht für *Programme for International Student Assessment*. Die Studie ist Teil des Indikatorenprogramms INES der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD). Die OECD startete die PISA-Studien im Jahr 1997.

PISA represents a desire by governments to monitor the outcomes of education systems in terms of student achievement on a regular basis and within an internationally accepted common framework. An international consortium, led by the Australian Council for Educational Research (ACER), manages the design and implementation of PISA. Other consortium partners are the National Institute for Educational Measurement (CITO) in the Netherlands, Westat and the Educational Testing Service (ETS) in the United States, and the National Institute for Educational Policy Research (NIER) in Japan. ¹

1.2.1 Untersuchungsgegenstand und Durchführung der PISA-Studie

Im Mittelpunkt des Interesses stehen in PISA so genannte *Basiskompetenzen*, die in modernen Staaten für eine Teilhabe am gesellschaftlichen, wirtschaftlichen und politischen Leben als notwendig erachtet werden. Neben der Frage, inwieweit Jugendliche diese Kompetenzen erworben haben und soziale Ungleichheiten im Bildungserfolg bestehen, interessieren Aspekte schulischer und außerschulischer Lern- und Lebensbedingungen.

1.2.2 Ziele von PISA

PISA ist ein langfristig angelegtes Projekt in drei Erhebungszyklen, in denen drei inhaltlich abgegrenzte Bereiche, so genannte Kompetenzbereiche, *Lesekompetenz* (reading literacy), *mathematische Grundbildung* (mathematical literacy) und *naturwissenschaftliche Grundbildung* (scientific literacy) mit wechselndem Schwerpunkt untersucht werden.

Der erste Zyklus im Jahr 2000 umfasst die *Lesekompetenz* als Schwerpunkt, der zweite Zyklus (Erhebung im Jahr 2003) die *mathematische* und der dritte Zyklus (Erhebung im Jahr 2006) die *naturwissenschaftliche Grundbildung*.

¹ OECD (Hrsg.) Thomson, Cresswell, De Bortoli: Facing the Future PISA 2003 S.27

1.2.3 Testkonzeption

PISA zielt nicht auf reproduzierendes Faktenwissen ab. Mit der Erhebung soll vielmehr untersucht werden, inwieweit die Jugendlichen in der Lage sind, ihre Kenntnisse und Fähigkeiten in realistischen Situationen anzuwenden und zur Bewältigung von Alltagsproblemen zu nutzen. Der Test zielt auf ein tieferes Verständnis für zentrale Konzepte, indem er aufzeigen soll, ob Schüler Prozesse wie das *Modellieren von Situationen* beherrschen, das *Kommunizieren von Ergebnissen* oder das kritische Beurteilen von Informationen verstehen und ob sie in der Lage sind, *Konzept- und Prozesswissen* in unterschiedlichen Kontexten anzuwenden.

Die Rahmenkonzeption wurde in enger Zusammenarbeit zwischen internationalen und nationalen Expertengruppen entwickelt. Im Lesetest wurde eine breite Palette verschiedener Arten von Texten eingesetzt, die neben kontinuierlichen Texten wie Erzählungen, Beschreibungen oder Anweisungen auch nichtkontinuierliches Material wie Tabellen, Diagramme oder Formulare umfasst.

Mathematische Grundbildung bedeutet die Fähigkeit, mathematische Begriffe als Werkzeuge in einer Vielfalt von Kontexten einzusetzen. Naturwissenschaftliche Grundbildung wird als Verständnis grundlegender naturwissenschaftlicher Konzepte, wie etwa Energieerhalt, *Vertrautheit* mit naturwissenschaftlichen Denk- und Arbeitsweisen sowie als Fähigkeit, dieses *Konzept- und Prozesswissen* vor allem bei der Beurteilung naturwissenschaftlich-technischer Sachverhalte anzuwenden, verstanden.

PISA besteht aus einer Mischung von Multiple-Choice Aufgaben und Aufgaben mit offenem Format. Insgesamt umfasst der Test Items für eine Gesamtdauer von sieben Stunden, wovon die Schüler jeweils unterschiedliche Kombinationen in einer zweistündigen Testsitzung bearbeiten. Die Testleistungen sämtlicher Schüler wurden auf eine Skala mit $\mu = 500$ und $\sigma = 100$ standardisiert.

Hintergrundmerkmale von Schülern zur sozialen Herkunft, Aspekte der Beziehung der Jugendlichen zu ihren Eltern, Einstellungen der Schüler zum Lesen sowie ihre privaten Lesegewohnheiten wurden zusätzlich erhoben. Den Schulen wurden Fragen zur finanziellen und personellen Ausstattung, Größe von Lerngruppen, Organisationsstrukturen und Entscheidungsprozessen gestellt. Eine Befragung der Eltern diente schließlich zur Verifikation der Schülerangaben.

1.2.4 Projekt

Weltweit nahmen im Frühsommer 2000 rund 180.000 Jugendliche aus 28 OECD-Mitgliedsstaaten sowie aus Brasilien, Lettland, Liechtenstein und der Russischen Föderation an der Erhebung teil. In Deutschland bestand die Stichprobe für den internationalen Vergleich aus 5.073 Schülern in 219 Schulen, wobei im Durchschnitt 23 15-Jährige pro Schule getestet wurden.

Auf Beschluss der Kultusminister der Länder in der Bundesrepublik Deutschland wurde die PISA-Stichprobe zusätzlich erweitert, so dass ein Vergleich der Länder möglich wurde. Insgesamt wurden 45.899 Schülerinnen und Schüler (zwei überlappende Stichproben von 33.809 15-Jährigen und 33.766 Schülern der 9. Klasse) in 1.466 Schulen untersucht. Die Tests fanden von Mai bis Juni 2000 in den Schulen statt.²

Die Beteiligungsquoten der ausgewählten Schulen wurden vorab festgelegt, so dass in jedem Teilnehmerstaat mindestens 85 Prozent der für die Stichprobe gezogenen Schulen und mindestens 80 Prozent der ausgewählten Schüler an der Untersuchung teilnehmen mussten. Die Niederlande verfehlten diese Mindestbeteiligungsquoten und wurden vom internationalen Vergleich ausgeschlossen. In Deutschland wurde auf Schülerebene eine Teilnahmequote von 86 Prozent erreicht.

Die internationalen Kriterien für die Beteiligungsquoten wurden auch auf den nationalen Ländervergleich [PISA-E] übertragen. Diese Mindeststandards wurden in Berlin und Hamburg deutlich unterschritten. Ein Vergleich der Ergebnisse an Gymnasien war jedoch möglich.

1.2.5 Koordination

PISA versteht sich als kooperatives Unternehmen, das wissenschaftliche Expertise aus allen Teilnehmerstaaten zusammenführt und von den Regierungen der Teilnehmerstaaten gelenkt wird. Die wesentlichen Entscheidungen trifft das Board of Participating Countries (BPC) der OECD. Mit der Planung und wissenschaftlichen Koordination der Studie wurde vom BPC ein internationales Konsortium unter Federführung des Australian Council for Educational Research (ACER) beauftragt. In Deutschland wurde die PISA-Studie von der Kultusministerkonferenz in Auftrag gegeben. Verantwortlich für ihre Durchführung und Erweiterung war ein nationales Konsortium unter der Leitung des Max-Planck-Instituts für Bildungsforschung (MPIB) in Berlin.

² Angaben des MPIB. Die Stichprobenumfänge der Public-Use-Files weichen geringfügig ab: 34.561 15-Jährige, 31.751 Schüler der 9. Klasse.

1.2.6 Kontroverse

Die internationalen Studien, TIMSS und PISA sind zum Teil kontrovers zur Kenntnis genommen worden.

In Großbritannien wäre Prais³ zu nennen. In der Bundesrepublik wies Kraus bereits in der Mathematikstudie (TIMSS) auf die unterschiedliche Vorauswahl der Schüler hin.⁴ Collani verwies auf statistische Sachverhalte, die die internationalen Testergebnisse in Frage stellten.⁵ Brügelmann wies auf Inkonsistenzen zwischen den einzelnen Studien hin.⁶ Klemm stellte das relativ gute Abschneiden der Bundesländer Bayern und Baden-Württemberg in PISA-E in Frage,⁷ Fuchs sprach das Problem öffentlicher Kontrolle der OECD-Studien an.⁸

2 Evaluation

Ausgangspunkt vergleichender, empirischer Untersuchungen ist die präzise Bestimmung des Untersuchungsgegenstandes und die damit verbundene Frage nach einer geeigneten Stichprobendefinition. In der Schulvergleichsforschung sind insbesondere das Schulalter (Klassenstufe) und das Lebensalter von Bedeutung. Folglich wird die Stichprobendefinition an erster Stelle zum Gegenstand dieser Evaluation.

2.1 Stichprobendefinitionen internationaler Vergleichsstudien

Die PISA-Studien stehen in Zusammenhang vorangegangener und nachfolgender internationaler Vergleichsuntersuchungen, insbesondere der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS) - mit unterschiedlichen Populationsdefinitionen sowie der Internationalen Grundschulstudie (IGLU).

³ S. J. Prais, Oxford Review of Education, Vol. 29, June, 2003. Prais: National Institute of Economic and Social Research (NIESR)

⁴ Kraus: Apokalyptische Inszenierung und dilettantischer Leichtsin. In: DEUTSCHER LEHRERVERBAND (DL)

⁵ Collani, E.v.: OECD PISA – An Example of Stochastic Illiteracy?, Economic Quality Control Vol 16 (2001), No.2, S. 240

⁶ Brügelmann, H.: Besserwisser und Alleskönner, Gewerkschaft Erziehung und Wissenschaft 11/12/2001

⁷ Klemm, K.: PISA E – Erste Einschätzungen (Frankfurter Rundschau, 26.7.2002)

⁸ Fuchs, H.-W.: Auf dem Weg zu einem Weltcurriculum? Zeitschrift für Pädagogik. (2003) S.161ff

2.1.1 TIMSS

Die Dritte Internationale Mathematik- und Naturwissenschaftsstudie (TIMSS) untersucht drei verschiedene Populationen.⁹

2.1.1.1 Population I

Population I entspricht der Grundschule. Ihr gehören Schülerinnen und Schüler derjenigen beiden angrenzenden Klassenstufen an, die zum Testzeitpunkt den größten Anteil der 9-Jährigen umfasst. Diese Population wurde in Deutschland nicht untersucht.

*Population 1: all students enrolled in the two adjacent grades that contain the largest proportion of students of age 9 years at the time of testing.*¹⁰

2.1.1.2 Population II

Die Population II entspricht der Sekundarstufe I. Ihr gehören Schülerinnen und Schüler der beiden angrenzenden Klassenstufen an, die zum Testzeitpunkt den größten Anteil der 13-Jährigen aufweisen. In Deutschland wurde jedoch wegen der relativ späten Einschulung der Schüler auf die Altersvorschrift verzichtet und stattdessen die Schuljahrgänge 7 und 8 gezogen.

*Population 2: all students enrolled in the two adjacent grades that contain the largest proportion of students of age 13 years at the time of testing.*¹¹

2.1.1.3 Population III

Population III entspricht der Sekundarstufe II.

Population 3: all students in their final year of secondary education, including students in vocational education programs. Population 3 has two optional subpopulations:
- Students taking advanced courses in mathematics

⁹ Baumert, Köller, Lehrke, Brockmann: Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS/III)- Technische Grundlagen. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie, 2000, Band 1. S.31

¹⁰ Martin, Kelly (Hrsg.): TIMSS Technical Report 1996 Vol. 1. S. 16

¹¹ Martin, Kelly (Hrsg.): TIMSS Technical Report 1996 Vol. 1. S. 16

- *Students taking advanced courses in physics.*¹²

(Hier) *mussten für die Population III breitere Rahmendefinitionen festgelegt werden, die jeweils national angemessene Präzisierungen erlaubten. Die Schulsysteme der an TIMSS teilnehmenden Staaten sind in der Sekundarstufe II zu unterschiedlich organisiert, als dass sich eine enge Definition hätte finden lassen.*¹³

Die national angemessenen Präzisierungen umfassen ein mittleres Alter (arithmetisches Mittel des ungewichteten Datensatzes) von 17,1 Jahren/mittleres Geburtsjahr 1977,9 (Russische Föderation) bis 20,45 Jahren/mittleres Geburtsjahr 1974,55 (Island). Das Alter des jüngsten Teilnehmers betrug 14 Jahre (Geburtsjahr April 1981, ID-Stud 21581005, Russische Föderation), das des ältesten 59 Jahre (Geburtsjahr August 1936, ID-Stud 3240010, Dänemark).¹⁴

¹² Martin, Kelly (Hrsg.): TIMSS Technical Report 1996 Vol. 1. S. 16

¹³ Baumert, Köller, Lehrke, Brockmann: Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS/III)- Technische Grundlagen. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S.32

¹⁴ Berechnung aus den Rohdaten TIMSS/III bezogen auf das Jahr 1995, ohne Berücksichtigung des für den einzelnen Schüler individuellen Geburts-, Testmonats sowie Testjahrs. Laut TIMSS Technical Report fand die Erhebung im August 1995 (südliche Hemisphäre) und Februar bis Mai 1995 (nördliche Hemisphäre) statt (TIMSS Technical Report, Vol.1. S. 16). In Island erfolgte die Untersuchung laut Datensatz im März 1996. Dies würde zu einem numerischen Durchschnittsalter von 21,4 Jahren führen. Der internationale Test fand nach Datenlage zwischen Februar 1995 und Mai 1996 statt. Der Testmonat in Österreich wird nicht genannt.

STUDENT'S DATE OF BIRTH\YEAR				
	Häufigkeit	Prozent	Gültige Prozenze	Kumulierte Prozenze
Gültig	36	,0	,0	,0
	40	,0	,0	,0
	41	,0	,0	,0
	43	,0	,0	,0
	45	,0	,0	,0
	46	,0	,0	,0
	47	,0	,0	,0
	48	,0	,0	,0
	49	,0	,0	,0
	50	,0	,0	,0
	51	,0	,0	,0
	52	,0	,0	,0
	53	,0	,0	,0
	54	,0	,0	,0
	55	,0	,0	,1
	56	,0	,0	,1
	57	,0	,0	,1
	58	,0	,0	,1
	59	,0	,0	,1
	60	,0	,0	,1
	61	,0	,0	,1
	62	,0	,0	,2
	63	,0	,0	,2
	64	,0	,0	,2
	65	,0	,0	,2
	66	,0	,1	,3
	67	,1	,1	,3
	68	,1	,1	,4
	69	,1	,1	,5
	70	,2	,2	,7
	71	,3	,3	,9
	72	,5	,5	1,5
	73	,9	,9	2,3
	74	2,6	2,7	5,0
	75	9,5	9,6	14,7
	76	32,4	33,0	47,6
	77	35,8	36,4	84,1
	78	14,2	14,4	98,5
	79	1,4	1,4	100,0
	80	,0	,0	100,0
	81	,0	,0	100,0
	Gesamt	87619	98,4	100,0
Fehlend	98	1,1		
	System	482	,5	
	Gesamt	1461	1,6	
Gesamt	89080	100,0		

Bericht			
STUDENT'S DATE OF BIRTH\YEAR			
COUNTRY ID	Mittelwert	N	Standardabweichung
Australia	77,57	3025	,695
Austria	76,11	3076	1,690
Canada	76,45	9532	,757
Cyprus	77,13	1161	,418
Czech Republic	76,72	3819	,673
Denmark	75,48	4363	2,130
France	76,32	3377	1,072
Germany	76,36	4873	1,610
Greece	78,20	803	,445
Hungary	77,01	4901	1,064
Iceland	74,55	1673	3,689
Israel	77,95	2390	,405
Italy	76,01	1974	1,167
Latvia (LSS)	76,83	708	,430
Lithuania	76,98	3620	,512
Netherlands	76,09	1454	1,202
New Zealand	77,36	1763	,608
Norway	75,32	3558	2,156
Russian Federation	77,90	4911	,382
South Africa	74,94	2629	2,726
Sweden	75,83	4572	1,071
Switzerland	75,17	5313	1,331
United States	76,67	10834	,606
Slovenia	76,14	3290	,621
Insgesamt	76,41	87619	1,532

Abbildung 1 Target Population TIMSS/III, Berechnung ungewichtet

Weiterhin unterscheiden sich die getesteten Schüler erheblich hinsichtlich ihrer Vorauswahl.¹⁵

¹⁵ Baumert, Köller, Lehrke, Brockmann: Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS/III) – Technische Grundlagen. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S.41

2.1.2 PISA

*In PISA wurden Schülerinnen und Schüler untersucht, die zum Beginn des Testzeitraums zwischen 15 Jahren/drei Monaten und 16 Jahren/zwei Monaten alt waren – unabhängig von der besuchten Jahrgangsstufe oder Art der Bildungseinrichtung.*¹⁶

Der Technical Report der PISA-Studie definiert wie folgt:

The desired base PISA target population in each country consisted of 15-year-old students attending educational institutions located within the country. This meant that countries were to include (i) 15-year-olds enrolled full-time in educational institutions, (ii) 15-year-olds enrolled in educational institutions who attended on only a part-time basis, (iii) students in vocational training types of programmes, or any other related type of educational programmes, and (iv) students attending foreign schools within the country (as well as students from other countries attending any of the programmes in the first three categories). It was recognised that no testing of persons schooled in the home, workplace or out of the country would occur and therefore these students were not included in the International Target Population. The operational definition of an age population directly depends on the testing dates.

(...)

The 15-year-old International Target Population was slightly adapted to better fit the age structure of most of the Northern Hemisphere countries. As the majority of the testing was planned to occur in April, the international target population was consequently defined as all students aged from 15 years and 3 (completed) months to 16 years and 2 (completed) months at the beginning of the assessment period. This meant that in all countries testing in April 2000, the national target population could have been defined as all students born in 1984 who were attending a school or other educational institution. Further, a variation of up to one month in this age definition was permitted. For instance, a country testing in March or in May was still allowed to define the national target population as all

¹⁶ Baumert, Stanat, Demmrich: PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 34

*students born in 1984. If the testing was to take place at another time, the birth date definition had to be adjusted and approved by the Consortium.*¹⁷

2.1.3 IGLU

In der Internationalen Grundschul-Lese-Untersuchung (IGLU) wird die Zielgruppe wie folgt definiert:

*Die Zielgruppe der Untersuchung war „the upper of the two adjacent grades with the most 9-year-olds“ (...) In den meisten Ländern und auch in Deutschland entspricht diese Definition der vierten Jahrgangsstufe.*¹⁸

Die Studie orientiert sich an der Klassenstufe (zumeist 4. Klasse). In England, Neuseeland und Schottland wurden die Schüler jedoch am Ende des fünften Schuljahres getestet.¹⁹

Grob vereinfachend lässt sich feststellen, dass IGLU mit der Untersuchung von Schülern am Ende der vierten Jahrgangsstufe eine am Schulalter orientierte Definition vorsieht. Die PISA-Studie hingegen orientiert sich am Lebensalter der Probanden. Die Dritte Internationale Mathematik- und Naturwissenschaftsstudie (TIMSS) wiederum stellt eine Reihe unterschiedlicher Definitionen auf, in der Regel eine Kombination beider Varianten. Die Vergleichbarkeit der Studien wird später Gegenstand der Diskussion sein.

2.2 PISA – Vergleichbarkeit der Populationen innerhalb der Studie

Die im Rahmen der PISA-Studie fokussierte Gruppe der 15-Jährigen weist eine ungleiche Verteilung hinsichtlich der Variablen Schulalter über die teilnehmenden Länder auf. So findet sich in Deutschland der überwiegende Teil der 15-Jährigen in der 9. Klassenstufe.²⁰ Die folgende Abbildung gibt die Schulalterverteilung nach Teilnehmerstaat, entsprechend dem öffentlich zugänglichen Datensatz (Public Use File) wieder. Die Schulalterverteilung

¹⁷ Adams, Wu (Hrsg.): PISA 2000 Technical Report. S. 39

¹⁸ Lankes, Bos, Mohr, Plaßmeier, Schwippert, Sibberns, Voss: Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU) und ihrer Erweiterung um Mathematik und Naturwissenschaften (IGLU-E) Bos, Lankes, Prenzel, Schwippert, Walther, Valtin (Hrsg.): Erste Ergebnisse aus IGLU. S. 7

¹⁹ Lankes, Bos, Mohr, Plaßmeier, Schwippert: Lehr- und Lernbedingungen in den Teilnehmerländern Bos, Lankes, Prenzel, Schwippert, Walther, Valtin (Hrsg.): Erste Ergebnisse aus IGLU 2003. S. 38

der bundesdeutschen Schülerinnen und Schüler steht in Kontrast zu einer Anzahl weiterer Teilnehmerstaaten. Diese weisen einen zum Teil erheblichen Anteil in höheren Klassenstufen auf.

Country three-digit ISO code * Grade - Q2 Kreuztabelle

% von Country three-digit ISO code		Grade - Q2										Gesamt
Country three-digit ISO code		4	5	6	7	8	9	10	11	12	13	
AUSTRALIA						,1%	6,3%	76,7%	16,9%			100,0%
AUSTRIA				,0%	,3%	4,8%	42,1%	52,7%	,1%			100,0%
BELGIUM					,2%	4,2%	26,3%	68,2%	1,0%	,0%		100,0%
BRAZIL				23,8%		34,5%	34,5%	7,2%				100,0%
CANADA				,0%	,2%	1,7%	16,4%	80,4%	1,2%	,0%	,0%	100,0%
CZECH REPUBLIC				,1%	,3%	1,8%	38,2%	59,6%				100,0%
DENMARK						6,0%	91,7%	2,2%				100,0%
FINLAND					,2%	10,8%	89,0%					100,0%
FRANCE					,3%	7,5%	37,9%	51,7%	2,6%			100,0%
GERMANY				,0%	1,2%	13,2%	63,3%	22,3%	,1%			100,0%
GREECE					,1%	,8%	2,0%	79,8%	17,3%			100,0%
HUNGARY				,2%	1,0%	2,1%	59,1%	37,6%				100,0%
ICELAND								100,0%				100,0%
IRELAND					,0%	3,3%	61,6%	15,8%	19,2%	,0%		100,0%
ITALY					,5%	,6%	16,0%	77,2%	5,7%			100,0%
JAPAN								100,0%				100,0%
KOREA, REPUBLIC OF							1,2%	98,2%	,6%			100,0%
LATVIA	,0%	,1%	,1%	1,6%	7,0%	35,3%	55,4%	,5%				100,0%
LIECHTENSTEIN				1,0%	16,2%	79,2%	3,6%					100,0%
LUXEMBOURG				1,7%	16,6%	56,2%	25,6%					100,0%
MEXICO				2,8%	10,4%	27,2%	59,4%	,2%				100,0%
NETHERLANDS				,3%	4,9%	45,5%	49,2%	,1%				100,0%
NEW ZEALAND							7,0%	88,1%	4,8%	,1%		100,0%
NORWAY						,0%	,7%	98,9%	,4%			100,0%
POLAND							100,0%					100,0%
PORTUGAL			,2%	,7%	5,1%	11,9%	25,3%	56,5%	,3%			100,0%
RUSSIAN FEDERATION				,0%	,0%	1,8%	25,0%	72,6%	,6%			100,0%
SPAIN				,0%	,0%	2,0%	24,6%	73,4%	,0%			100,0%
SWEDEN					,0%	2,0%	97,8%	,2%				100,0%
SWITZERLAND				,7%	14,0%	77,0%	8,3%	,0%	,0%			100,0%
UNITED KINGDOM						,0%	33,5%	62,7%	3,8%			100,0%
UNITED STATES				,1%	1,8%	40,2%	57,4%	,4%				100,0%
Gesamt		,0%	,0%	,0%	1,1%	4,7%	32,4%	54,2%	7,2%	,3%	,0%	100,0%

Abbildung 2 Verteilung Klassenstufe

Eine der ersten Fragen wird die Vergleichbarkeit der Gruppen betreffen. Schümer kommt zu folgender Bewertung:

*Dass die in der 10. Klasse getesteten 15-Jährigen ein Jahr länger beschult worden sind als die meisten der in Deutschland erfassten Schülerinnen und Schüler ist allerdings keine hinreichende Erklärung für ihren Leistungsvorsprung, denn es gibt auch Länder, in denen die Mehrheit der 15-Jährigen ebenfalls erst auf der 9. Klassenstufe ist und trotzdem überdurchschnittliche Testleistungen erbracht hat.*²¹

²⁰ 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. 413

²¹ Schümer, Weiß, Steiner, Baumert, Tillmann, Meier: Institutionelle Bedingungen schulischen Lernens im internationalen Vergleich. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 415

*Dass sich die Schülerleistungen in beiden Ländern [Island, Österreich d. Verf.] nicht in der erwarteten Richtung voneinander unterscheiden, bestätigt erneut, dass die nominelle Unterrichtszeit wenig Erklärungskraft besitzt, solange die Stundentafeln, die Zeitnutzung und andere Kontextfaktoren außer Acht bleiben.*²²

Die von Schümer geäußerte These widerspricht den Darstellungen von Mullis, Martin, Beaton, Gonzalez, Kelly und Smith in der Dritten Internationale Mathematik- und Naturwissenschaftsstudie (TIMSS), die in Bezug auf den Mathematiktest von substantziellen Unterschieden ausgehen.

*Since for most children the opportunity to learn mathematics is anchored in the school, even one year can make a substantial difference.*²³

Jürges, Schneider und Büchel berichten ebenfalls einen statistisch signifikanten Einfluss der Klassenstufe auf die in TIMSS erzielten Leistungen (student achievement in mathematics – percentile effects).²⁴

Wößmann widerspricht ebenfalls der Einschätzung von Schümer:

In der PISA-Studie, deren Zielpopulation in allen Ländern die 15-Jährigen waren, gehörten die deutschen 15-Jährigen vergleichsweise niedrigen Jahrgangsstufen an: Während zwei Drittel der 15-Jährigen aller OECD-Länder der 10. bis 13. Jahrgangsstufe angehörten, waren über drei Viertel der deutschen 15-Jährigen in der 7. bis 9. Jahrgangsstufe. Bei den PISA-Ergebnissen fällt auf, dass deutsche Schüler bei einem Vergleich innerhalb jeder Jahrgangsstufe zwischen drei und fünf Prozent über der OECD-Durchschnittsleistung der jeweiligen Jahrgangsstufe liegen. Das heißt, 15-jährige Neuntklässler aus Deutschland sind besser als der

²² Schümer, Weiß, Steinert, Baumert, Tillmann, Meier: Institutionelle Bedingungen schulischen Lernens im internationalen Vergleich. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 418

²³ Martin, M.O., Beaton, A.E., Gonzalez, E.J., Kelly, D.L., Smith, T.A.: Mathematics Achievement in the Primary School Years (TIMSS), Boston College 1997, S. 30

²⁴ Jürges, H., Schneider, K., Büchel, K.: The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany, 2003 S. 19

OECD-Durchschnitt etc. Nur befinden sich die meisten deutschen 15-Jährigen im Vergleich zum OECD-Durchschnitt in niedrigeren Jahrgangsstufen.²⁵

Der Einfluss des Schulalters auf die Testleistungen soll nun an dieser Stelle Gegenstand der Untersuchung sein. Als inferenzstatistische Verfahren können die Varianzanalyse sowie die Regressionsanalyse eingesetzt werden. Die Autoren Bos, Lankes, Schwippert, Valtin, Voss, Badel und Plaßmeier verwenden in der Internationalen Grundschul-Lese-Untersuchung (IGLU) die Korrelationsanalyse im Rahmen einer analogen Fragestellung.²⁶ Sie untersuchen den Zusammenhang von Schuljahr und Testleistung, der sowohl mit $r=0,23$ ²⁷ als auch mit $r=0,34$ berichtet wird.^{28 29} Daher soll diese Verfahrensweise hier Berücksichtigung finden. Der sinnvolle Einsatz der Korrelation setzt die Vergleichbarkeit der Teilnehmerstaaten hinsichtlich wesentlicher struktureller Merkmale voraus. Diese Voraussetzung dürfte mit zunehmender Heterogenität der Teilnehmerstaaten, insbesondere bei einer Erweiterung der Teilnehmerstaaten wie in PISA 2003, kaum erfüllt sein.

²⁵ Wößmann: Familiärer Hintergrund, Schulsystem und Schülerleistungen im internationalen Vergleich, aus Politik und Zeitgeschichte (B 21-22/2003)

²⁶ Bos, Lankes, Schwippert, Valtin, Voss, Badel, Plaßmeier: Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In: Erste Ergebnisse aus IGLU 2003. S. 105

²⁷ Der eingesetzte Schätzer wurde im Original nicht weiter präzisiert. Die Kennwerte werden in vorliegender Studie ohne Rücksicht auf weitergehende Präzisierung des Schätzers übernommen.

²⁸ Bos, Lankes, Schwippert, Valtin, Voss, Badel, Plaßmeier: Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In: Erste Ergebnisse aus IGLU 2003. S. 105

²⁹ Bos, Valtin, Lankes, Schwippert, Voss, Badel, Plaßmeier: Lesekompetenzen am Ende der vierten Jahrgangsstufe in einigen Ländern der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. In: Bos, Lankes, Prenzel, Schwippert, Valtin, Walther (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich, 2004. S. 65

2.2.1 Korrelation

Als lineares Zusammenhangsmaß kann der Korrelationskoeffizient nach Pearson beziehungsweise als Rangstatistik der Korrelationskoeffizient nach Spearman eingesetzt werden. Modell: Es seien X, Y zwei Zufallsvariablen mit den Erwartungswerten μ_x und μ_y , den Varianzen σ_x^2, σ_y^2 sowie der Kovarianz σ_{xy} mit $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$,

Korrelation $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ und folgenden Statistiken

Korrelationskoeffizient nach Pearson

$$r_p = \frac{s_{xy}}{\sqrt{s_x^2 s_y^2}}$$

Korrelationskoeffizient nach Spearman (Rangstatistik)

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{(n-1)n(n+1)}$$

Zu testen ist nun die Hypothese $H_0 : \rho_{xy} = 0$ gegen $H_1 : \rho_{xy} \neq 0$ mit

$$\sqrt{n-2} \frac{|r|}{\sqrt{1-r^2}} > t_{n-2; 1-\alpha/2} \text{ als Ablehnungsbereich von } H_0^{30}.$$

Die zwischen Testleistung im Lesen und Schulalter über sämtliche Teilnehmerstaaten fallbasiert, über den ungewichteten Datensatz gemessene Korrelation beträgt $r_p = 0,33$. Dem entspricht ein positiver linearer Zusammenhang zwischen Schulalter und Lesetestleistung im PISA-Test³¹. Die Ausgangshypothese H_0 wird für $p < 0,01$ verworfen.

Correlations

		Plausible value in reading
Grade - Q2	Pearson Correlation	,333**
	Sig. (2-tailed)	,000
	N	172412

** . Correlation is significant at the 0.01 level (2-tailed).

Abbildung 3 Korrelation Klassenstufe Lesewert fallbasiert

Von Interesse ist nun die Differenzierung nach Teilnehmerstaat. Die Ergebnisse variieren erheblich. Das Schulalter wird in einigen Ländern, insbesondere in Korea und Japan, entweder nur sehr geringfügig oder nicht variiert.

³⁰ Bosch: Statistik 1996, S. 409

³¹ Auf die Darstellung von Teststärke (Power), Zusammenhang von Signifikanzniveau, Effektgröße und Stichprobenumfang wird verzichtet.

Mit $r_p = 0,419$ (Deutschland), $r_p = 0,429$ (Luxemburg), $r_p = 0,451$ (Niederlande), $r_p = 0,55$ (Spanien), $r_p = 0,594$ (Belgien), $r_p = 0,643$ (Frankreich) und $r_p = 0,71$ (Portugal) lässt sich ein positiver linearer statistischer Zusammenhang nachweisen. Sämtliche Korrelationen sind statistisch signifikant ($p < 0,01$). Dieser Zusammenhang lässt sich für jeden Teilnehmerstaat, in dem das Schulalter variiert wird, nachweisen ($p < 0,01$).

Correlations

Country three-digit ISO code			Plausible value in reading
AUSTRALIA	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,246** ,000 5140
AUSTRIA	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,266** ,000 4565
BELGIUM	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,594** ,000 6609
BRAZIL	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,550** ,000 4820
CANADA	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,312** ,000 29026
CZECH REPUBLIC	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,203** ,000 5183
DENMARK	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,185** ,000 4120
FINLAND	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,213** ,000 4864
FRANCE	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,643** ,000 4648
GERMANY	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,419** ,000 4997
GREECE	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,195** ,000 4586
HUNGARY	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,286** ,000 4840
ICELAND	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	, ^a , 3317
IRELAND	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,233** ,000 3829
ITALY	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,354** ,000 4952
JAPAN	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	, ^a , 5256
KOREA, REPUBLIC OF	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,029* ,041 4976

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Abbildung 4 Korrelation Klassenstufe Lesetestwert je Teilnehmerstaat (I)

Correlations

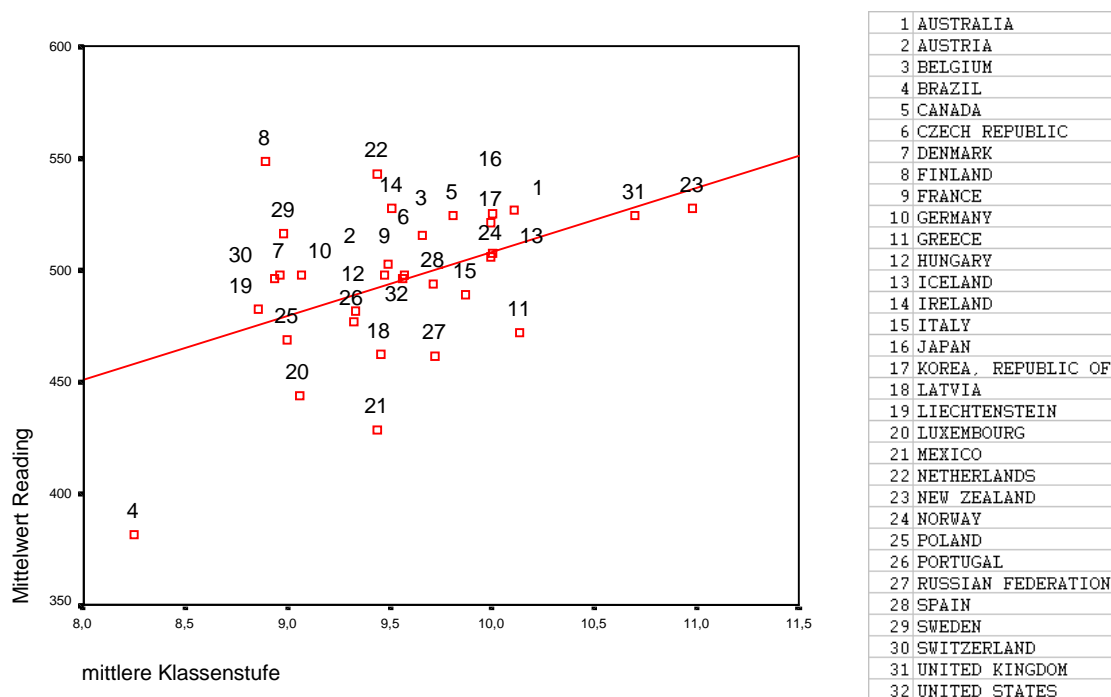
Country three-digit ISO code			Plausible value in reading
LATVIA	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,318** ,000 3846
LIECHTENSTEIN	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,403** ,000 308
LUXEMBOURG	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,429** ,000 3482
MEXICO	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,554** ,000 4297
NETHERLANDS	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,451** ,000 2470
NEW ZEALAND	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,273** ,000 3633
NORWAY	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,088** ,000 4099
POLAND	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	, ^a , 3654
PORTUGAL	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,710** ,000 4504
RUSSIAN FEDERATION	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,175** ,000 6688
SPAIN	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,550** ,000 6160
SWEDEN	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,174** ,000 4376
SWITZERLAND	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,320** ,000 6017
UNITED KINGDOM	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,098** ,000 9305
UNITED STATES	Grade - Q2	Pearson Correlation Sig. (2-tailed) N	,334** ,000 3845

** . Correlation is significant at the 0.01 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Abbildung 5 Korrelation Klassenstufe Lesetestwert je Teilnehmerstaat (II)

Eine weitere, wenngleich wie in der fallbasierten Korrelationsrechnung unter der Einschränkung vergleichbarer Schulsysteme und vergleichbarer sozioökonomischer Niveaus sinnvolle Möglichkeit zum Nachweis eines linearen Zusammenhangs besteht auf aggregiertem Datenniveau in Form eines Vergleichs mittlerer Testleistung und mittlerer Klassenstufe (grade) eines Landes (jeweils arithmetisches Mittel). Dieser führt mit $r_p = 0,464$ zu einem statistisch signifikanten Ergebnis ($p < 0,01$).³²



Correlations

		Mittelwert Reading
mittlerer Grad	Pearson Correlation	,464**
	Sig. (2-tailed)	,007
	N	32

** . Correlation is significant at the 0.01 level

Abbildung 6 Korrelation Klassenstufe Lesetestwert (aggregiert)

Die Teilnehmerstaaten Vereinigtes Königreich, Neuseeland und Australien weisen die höchste mittlere Klassenstufe und gleichzeitig die höchsten mittleren Lesewerte auf. Die niedrigste mittlere Klassenstufe und gleichzeitig den niedrigsten Lesewert weist Brasilien auf. Es besteht ein signifikant positiver linearer statistischer Zusammenhang.

³² Korrelationsberechnungen über den gesamten Datensatz (aggregiert oder fallweise) sind prinzipiell nur über vergleichbare Teilnahmestaaten sinnvoll, was hier –insbesondere jedoch in PISA 2003- in Frage gestellt werden kann.

2.2.2 Regression

Um den Einfluss weiterer Variablen im Rahmen statistischer Zusammenhangsanalysen zu berücksichtigen, können multivariate Verfahren, wie die der Regressionsanalyse, eingesetzt werden. Das eingesetzte Modell bezieht die Variablen *höchster sozioökonomischer Index* in der Familie (*HISEI*) sowie *Geschlecht* als dummycodierte Zufallsvariable nach folgendem Modell mit ein:

$$T = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

wobei sich die empirisch ermittelten Testwerte T aus der Klassenstufe X_1 , der dummycodierten Variable *Geschlecht* X_2 , dem höchsten sozioökonomischen Index der Familie X_3 dem Achsenabschnitt β_0 und dem Fehlerterm ε mit $\varepsilon \sim N(0, \sigma^2)$ ergibt. In folgender Tabelle gibt $B = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)'$ die Schätzungen für $\beta_0, \beta_1, \beta_2, \beta_3$ an.

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderungsstatistiken				
					Änderung in R-Quadrat	Änderung in F	df1	df2	Änderung in Signifikanz von F
1	,458 ^a	,210	,210	87,68889	,210	14408,532	3	162594	,000

a. Einflussvariablen : (Konstante), Highest In. Socio-Econ. Index, Sex - Q3, Grade - Q2

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	103,947	2,921		35,589	,000
	Grade - Q2	37,164	,294	,282	126,495	,000
	Sex - Q3	-29,118	,435	-,148	-66,866	,000
	Highest In. Socio-Econ. Index	1,744	,013	,291	130,509	,000

a. Abhängige Variable: Plausible value in reading

Abbildung 7 Regression *Klassenstufe, Geschlecht, Sozioökon. Index* auf Lesetestwert

Das relativ geringe korrigierte R^2 von 0,21 muss vor dem Hintergrund der Modellvoraussetzungen bewertet werden. Der Einfluss sämtlicher Variablen erweist sich in diesem Modell als statistisch signifikant ($p < 0,01$). Unter Berücksichtigung der Codierung der Zufallsvariable *Geschlecht* lässt der negative Beta-Koeffizient die geringeren Lesetestwerte der männlichen Schüler erkennen. Das Schulalter (grade) weist auch unter Kontrolle weiterer Variablen einen statistisch signifikanten, positiven Zusammenhang zur Testleistung auf.

2.2.3 Varianzanalytischer Ansatz

Ein varianzanalytischer Ansatz ist insofern sinnvoll, als einerseits der Faktor *Klassenstufe* (grade) in wenigen Stufen vorliegt, andererseits die Lesetestwerte standardisiert wurden und annähernd eine Normalverteilung aufweisen.^{33 34}

Das Modell ist einfaktoriell mit Klassenstufe als Faktor und wie folgt beschrieben: $X_{ij} = \mu + \alpha_i + E_{ij}$, $i=1, \dots, c$, $j=1, \dots, n_i$ worin μ das (unbekannte) Gesamtmittel und α_i der (unbekannte) Effekt des i -ten Treatments bedeutet. Die Zufallsvariablen E_{ij} sind unabhängig und $N(0, \sigma^2)$ -verteilt, d.h., X_{ij} ist $N(\mu_i, \sigma^2)$ -verteilt mit $\mu_i = \mu + \alpha_i$. Die Beobachtungen x_{ij} sind also nach *einem*, so genannten α -Effekt, klassifiziert.

Die Ausgangshypothese lautet $H_0 : \mu_1 = \mu_2 = \dots = \mu_c$ gegen $H_1 : \mu_k \neq \mu_l$ für mindestens ein Paar $(k, l), k \neq l$. Die Teststatistik für den Test auf Gleichheit der c Mittelwerte (d.h. sämtliche $\alpha_i = 0$) lautet:

$$F = \frac{(N - c) \sum_{i=1}^c n_i (\bar{X}_i - \bar{X})^2}{(c - 1) \sum_{i=1}^c \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}, \quad \sum_{i=1}^c n_i = N,$$

mit \bar{X}_i als Mittelwert der i -ten Stichprobe und \bar{X} als Mittelwert aller N Beobachtungen.

Die Teststatistik F hat unter H_0 eine F-Verteilung mit $(c - 1, N - c)$ Freiheitsgraden.³⁵

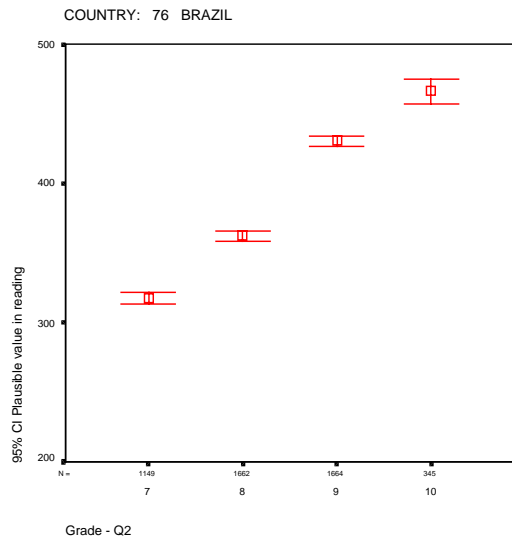
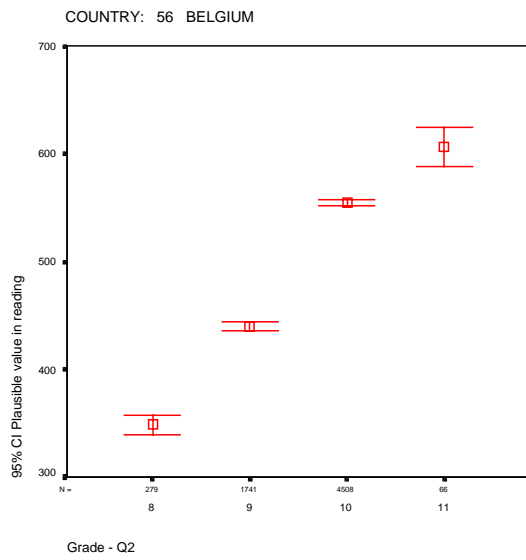
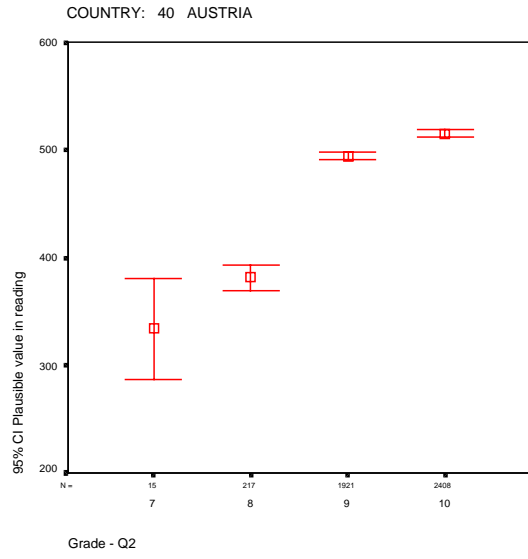
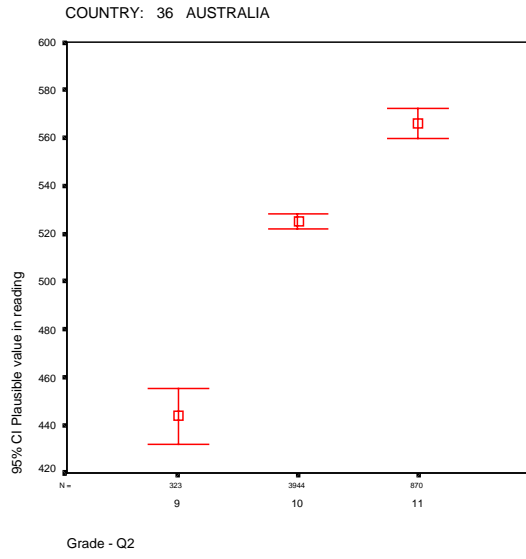
Im Folgenden sollen zunächst Fehlerbalken-Graphiken (arithmetisches Gruppenmittel mit Konfidenzintervall zum Niveau $1 - \alpha$) für jeden einzelnen Teilnehmerstaat dargestellt werden. Nur in einem Fall werden die Signifikanztests der Varianzanalyse sowie der Scheffé-Test als a posteriori Einzelvergleichstest dargestellt. Wie jedoch bereits aus den Konfidenzintervallen (95% KI) deutlich ersichtlich ist, sind die Ergebnisse jeweils statistisch signifikant ($p < 0,01$).³⁶

³³ Adams, Wu (Hrsg.): PISA 2000 Technical Report Kap. 13

³⁴ Zu Effizienzverlusten parametrischer Verfahren siehe Büning: Nichtparametrische Statistische Methoden 1994, insbesondere S. 285 ff

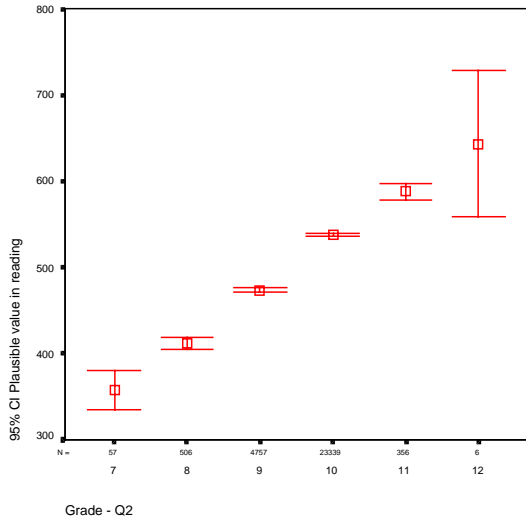
³⁵ Büning, H.: Nichtparametrische statistische Methoden 1994, S.183

³⁶ Hier werden lediglich die Berechnungen des ungewichteten Datensatzes berichtet. Berechnungen über den gewichteten Datensatz führen aufgrund der Erhöhung des Stichprobenumfangs durch die Gewichtungvariable (ohne Einsatz von Korrektur- und Replikationsvariablen) um 9525% zu progressiven Entscheidungen. N (ungew.)=174896, N (gew.) 16659562,5712. Gewichtungvariable w_{fstuw} : „The final weight variable on the

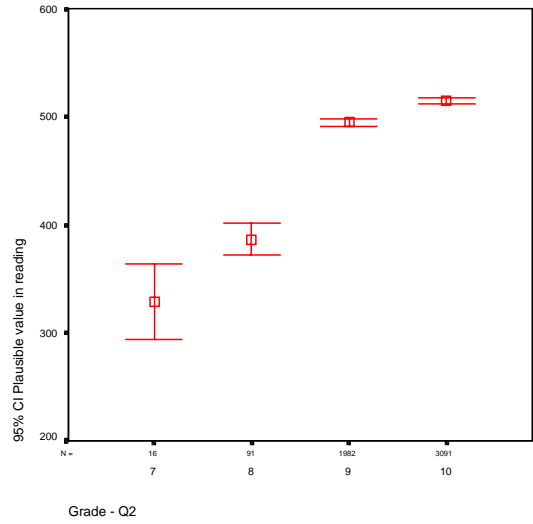


data file was called w_fstuwt, which is the final student weight that incorporates any studentlevel trimming." Vgl. Adams, Wu (Hrsg.): PISA 2000 Technical Report. S. 95

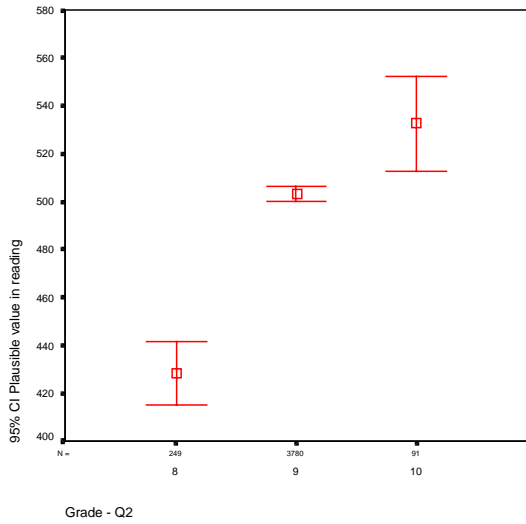
COUNTRY: 124 CANADA



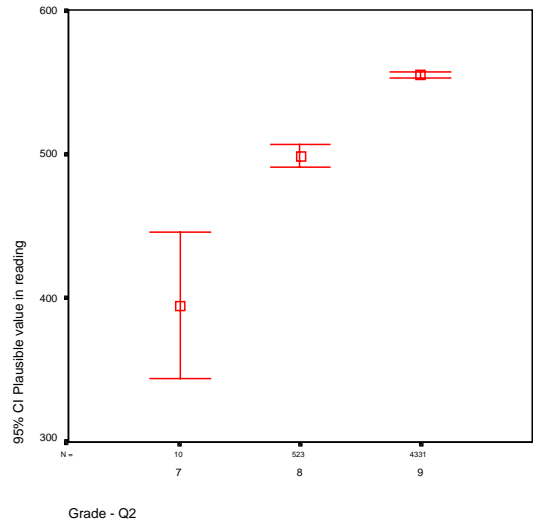
COUNTRY: 203 CZECH REPUBLIC



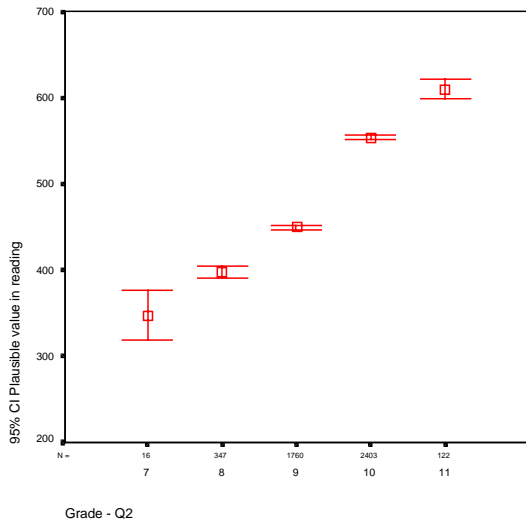
COUNTRY: 208 DENMARK



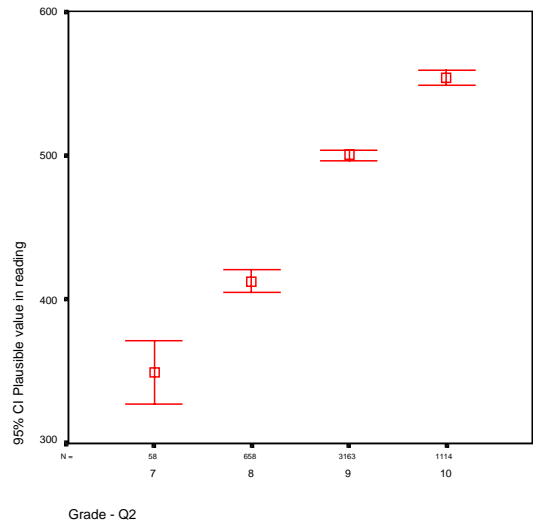
COUNTRY: 246 FINLAND

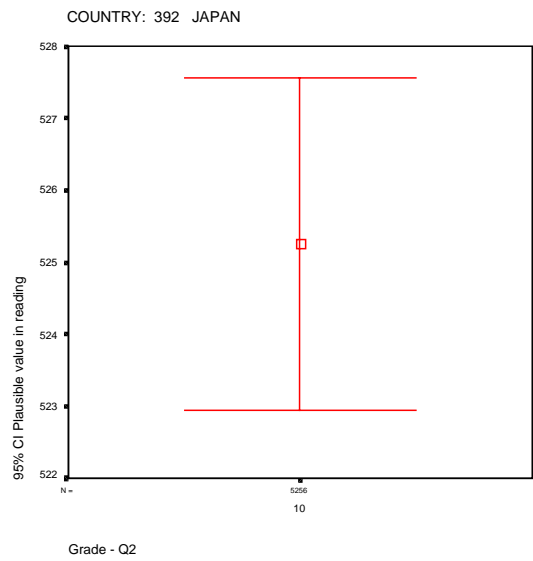
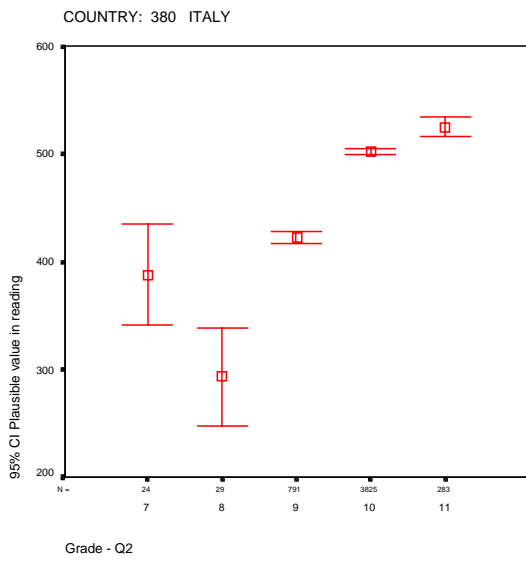
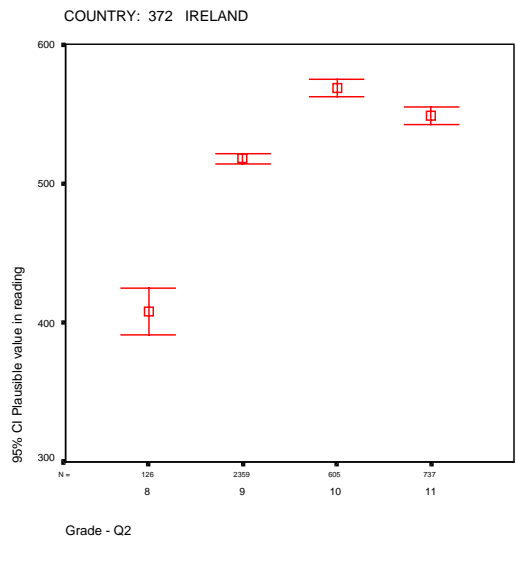
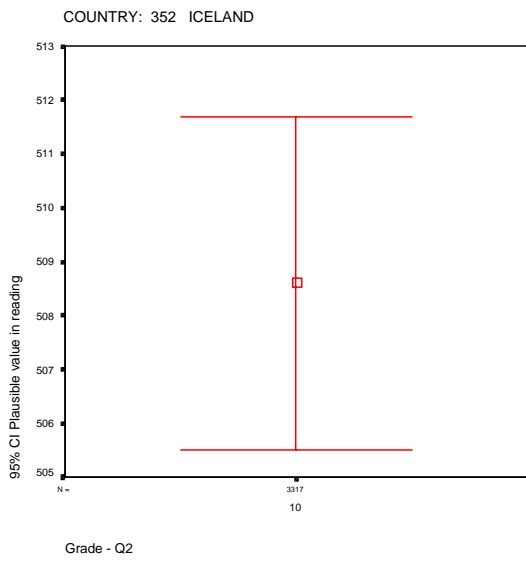
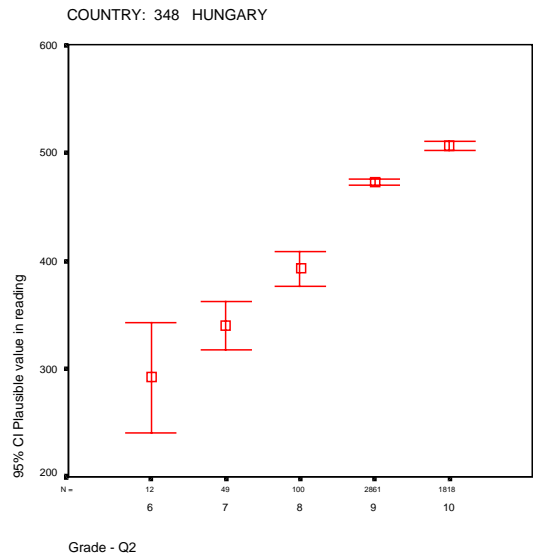
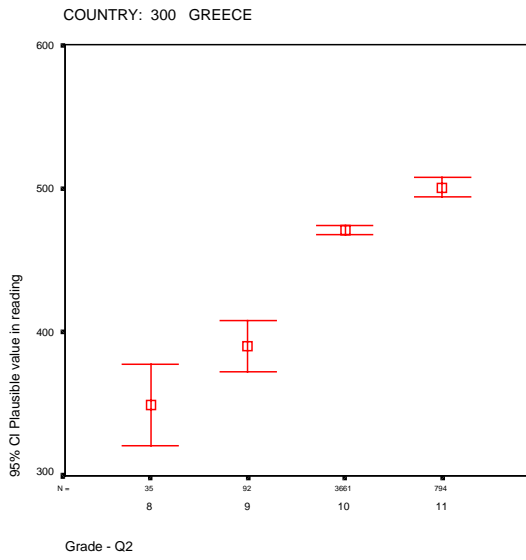


COUNTRY: 250 FRANCE

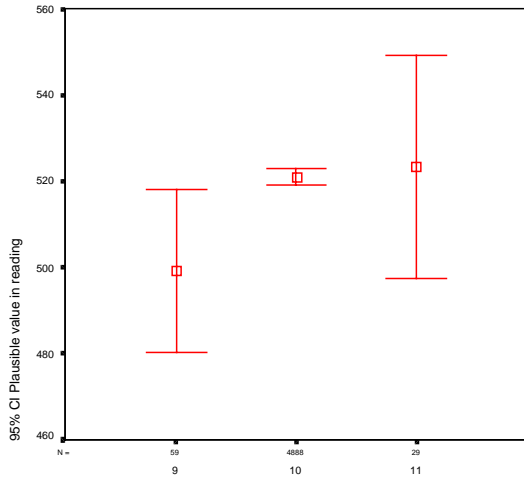


COUNTRY: 276 GERMANY



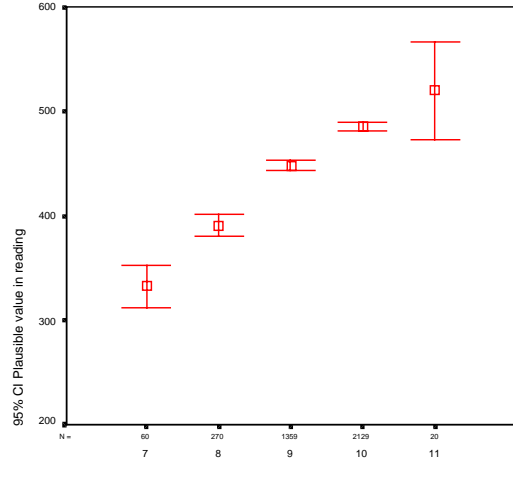


COUNTRY: 410 KOREA, REPUBLIC OF



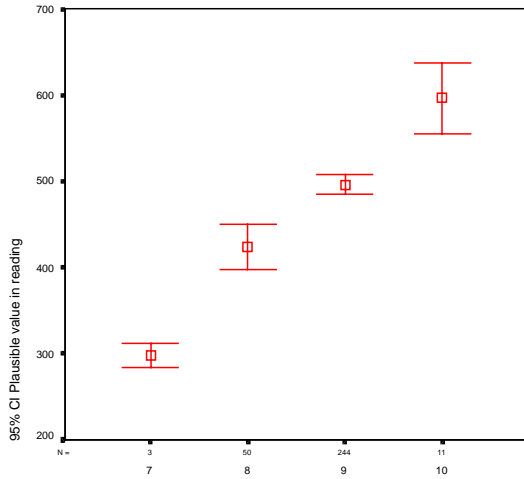
Grade - Q2

COUNTRY: 428 LATVIA



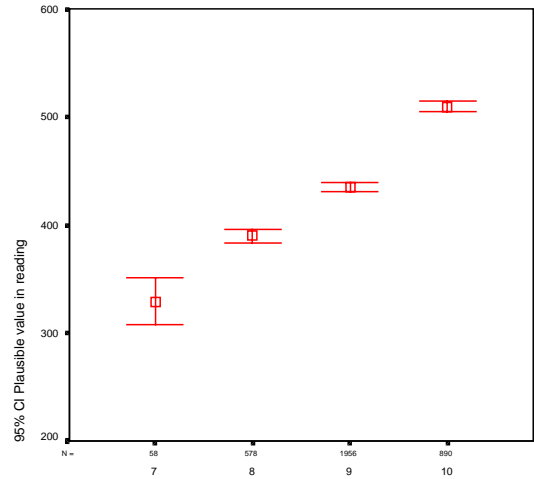
Grade - Q2

COUNTRY: 438 LIECHTENSTEIN



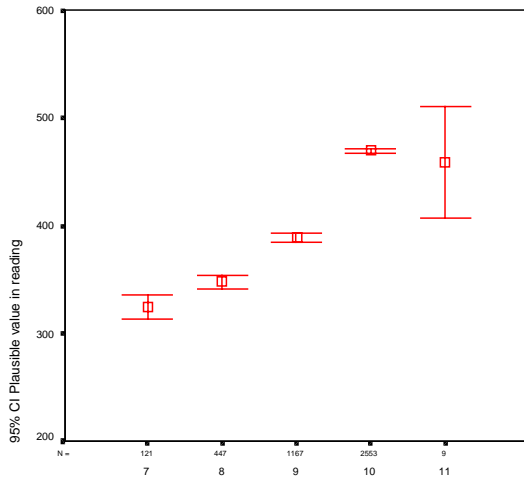
Grade - Q2

COUNTRY: 442 LUXEMBOURG



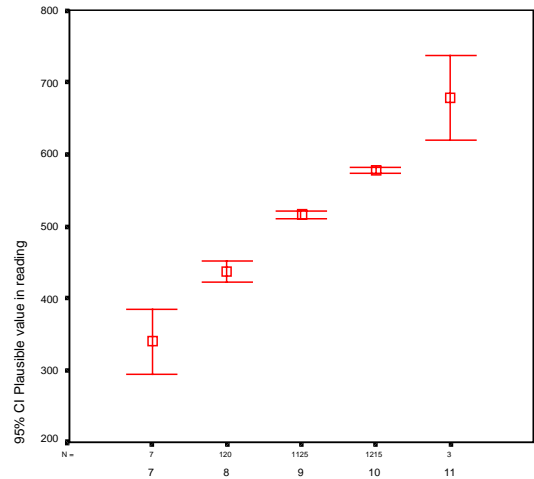
Grade - Q2

COUNTRY: 484 MEXICO

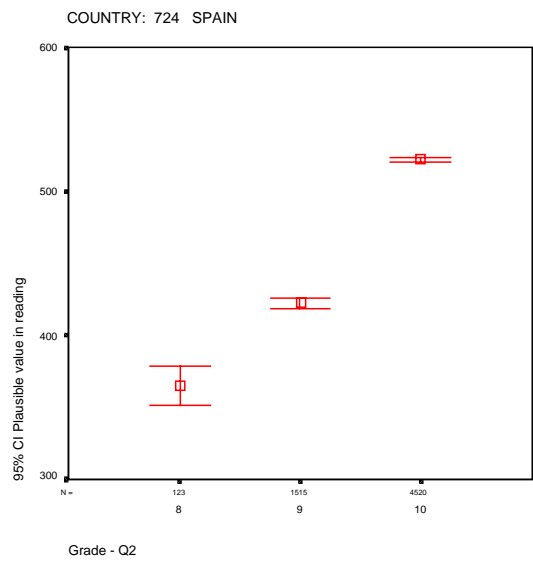
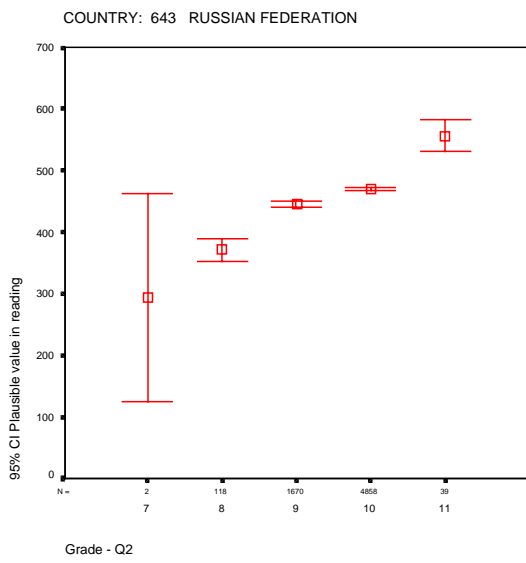
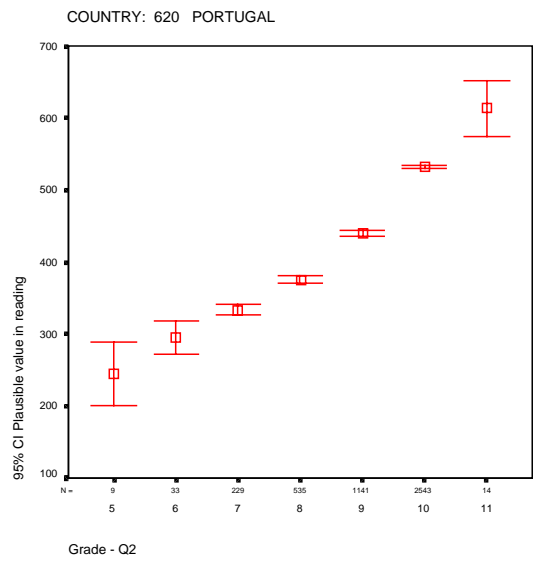
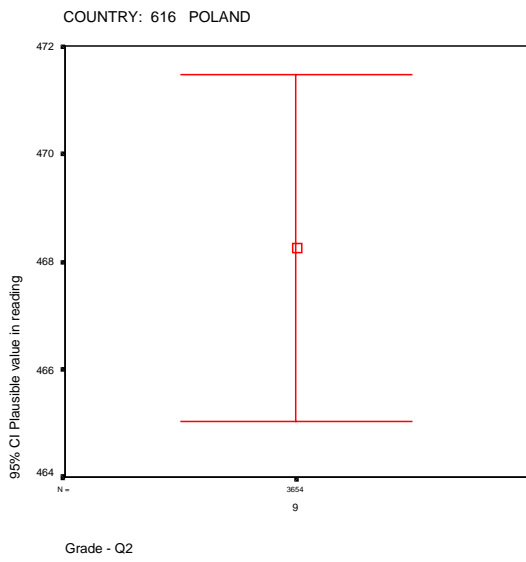
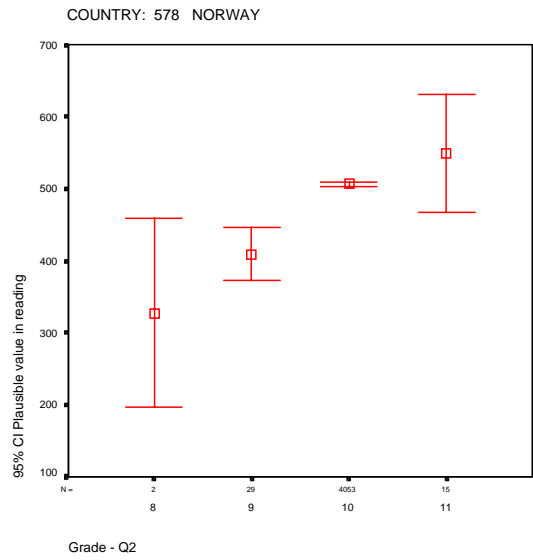
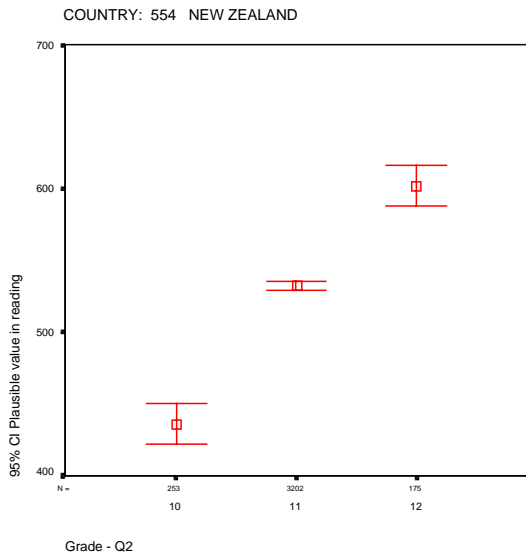


Grade - Q2

COUNTRY: 528 NETHERLANDS



Grade - Q2



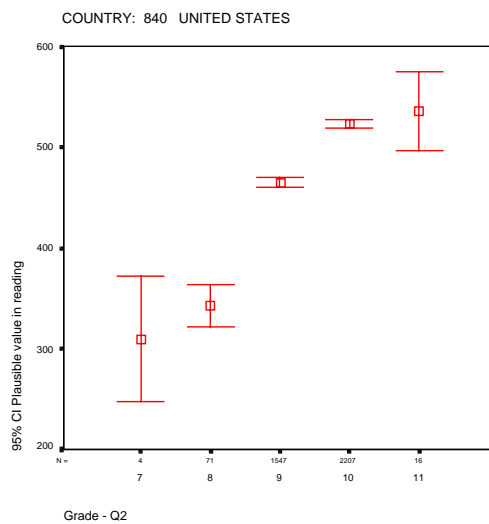
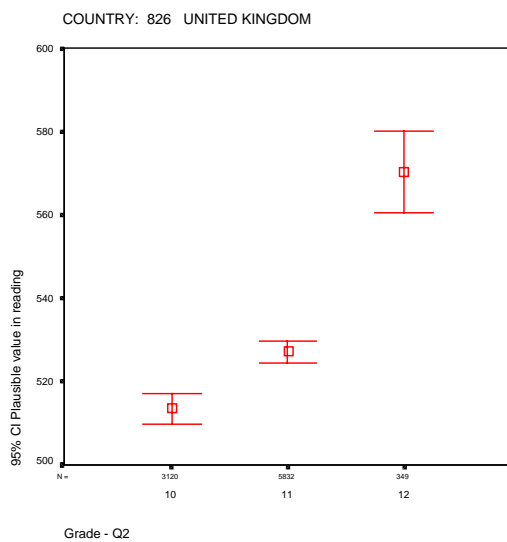
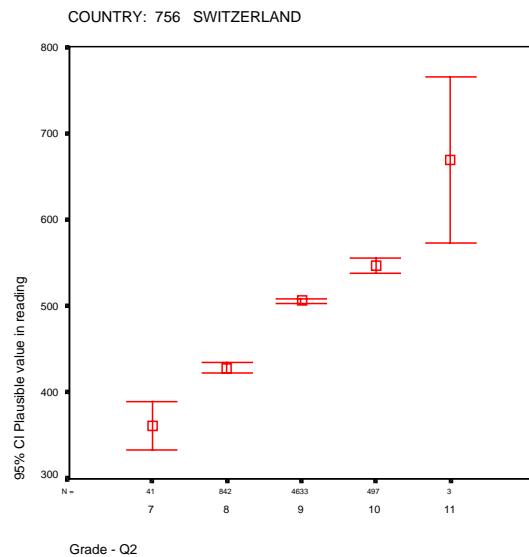
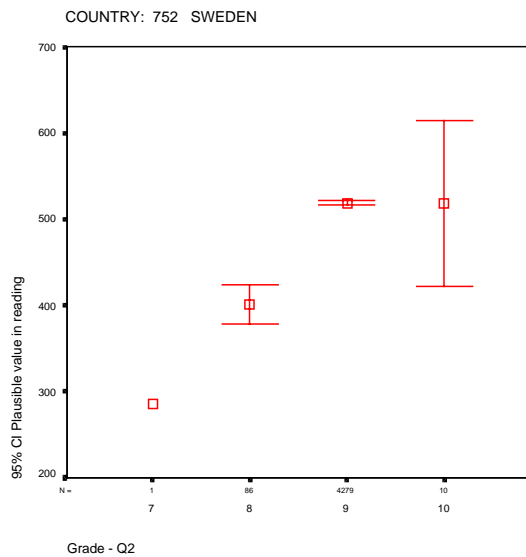


Abbildung 8 Varianzanalysen Klassenstufe und Lesetestwert

In folgender Graphik soll lediglich am Beispiel Deutschlands im internationalen Vergleich die Varianzanalyse und der Scheffé-Test dargestellt werden. Deutlich wird mit $F = 7804$ bei $df = 3$ die Differenz der Mittelwerte ($p < 0,01$). Der Scheffé-Test zeigt zwischen sämtlichen Faktorstufen statistisch signifikante Unterschiede.

Die mittleren Differenzen sind beträchtlich und zeigen mit bis zu 193 Punkten Unterschied eine auch substantziell relevante Effektgröße an. In Bezug auf den internationalen Datensatz lässt sich feststellen, dass das zwischen den Teilnehmerstaaten variierende Schulalter als starker Einfluss auf die Testleistung zu werten ist.

ONEWAY ANOVA

Plausible value in reading

	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	2,0E+08	3	67295899	7804,614	,000
Innerhalb der Gruppen	1,4E+09	159382	8622,579		
Gesamt	1,6E+09	159385			

Mehrfachvergleiche

Abhängige Variable: Plausible value in reading

Scheffé-Prozedur

(I) Grade - Q2	(J) Grade - Q2	Mittlere Differenz (I-J)	Standardfehler	Signifikanz	95%-Konfidenzintervall	
					Untergrenze	Obergrenze
7	8	-68,6247*	2,34808	,000	-75,1888	-62,0606
	9	-156,1933*	2,14450	,000	-162,1883	-150,1983
	10	-193,7751*	2,12999	,000	-199,7295	-187,8207
8	7	68,6247*	2,34808	,000	62,0606	75,1888
	9	-87,5686*	1,10594	,000	-90,6603	-84,4770
	10	-125,1504*	1,07755	,000	-128,1627	-122,1381
9	7	156,1933*	2,14450	,000	150,1983	162,1883
	8	87,5686*	1,10594	,000	84,4770	90,6603
	10	-37,5818*	,49647	,000	-38,9697	-36,1939
10	7	193,7751*	2,12999	,000	187,8207	199,7295
	8	125,1504*	1,07755	,000	122,1381	128,1627
	9	37,5818*	,49647	,000	36,1939	38,9697

*. Die mittlere Differenz ist auf der Stufe .05 signifikant.

Abbildung 9 Varianzanalyse Klassenstufe Lesetestwert/Scheffé-Test, Deutschland³⁷

Korrespondierende Daten liefert der nationale Ergänzungstest (PISA-E), der aufgrund des höheren Stichprobenumfangs auch für die Klassenstufe 7 einen genügend großen Umfang aufweist.³⁸ Die folgende Abbildung zeigt die mittleren Differenzen der in den einzelnen Klassenstufen erreichten Lesetestwerte (Itemauswahl des internationalen Tests) im bundesdeutschen nationalen Vergleich.

³⁷ Internationaler Datensatz

³⁸ Darstellung über die ungewichteten Daten.

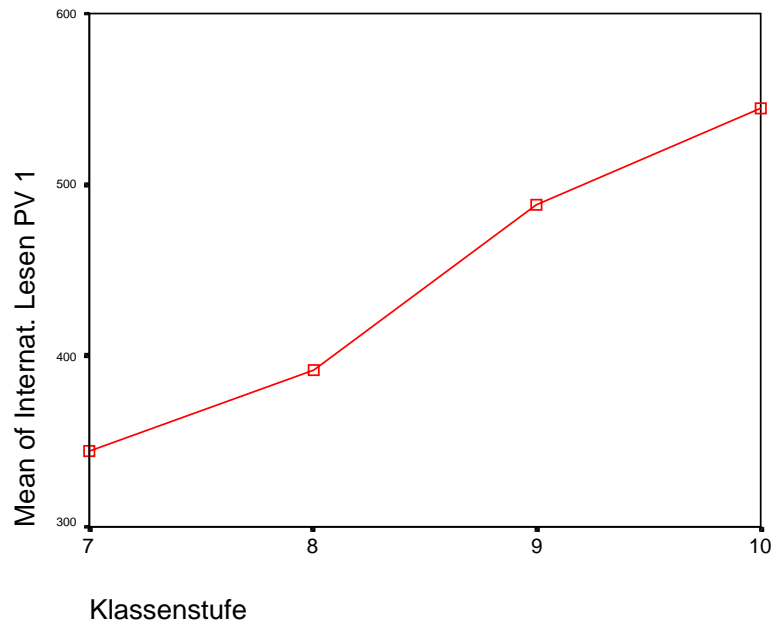
Mehrfachvergleiche

Abhängige Variable: Internat. Lesen PV 1

Scheffé-Prozedur

(I) Klassenstufe	(J) Klassenstufe	Mittlere Differenz (I-J)	Standardfehler	Signifikanz	95%-Konfidenzintervall	
					Untergrenze	Obergrenze
7	8	-46,7446*	4,68597	,000	-59,8448	-33,6444
	9	-143,2974*	4,53344	,000	-155,9712	-130,6236
	10	-199,6304*	4,62599	,000	-212,5630	-186,6979
8	7	46,7446*	4,68597	,000	33,6444	59,8448
	9	-96,5528*	1,48795	,000	-100,7125	-92,3931
	10	-152,8859*	1,74975	,000	-157,7775	-147,9942
9	7	143,2974*	4,53344	,000	130,6236	155,9712
	8	96,5528*	1,48795	,000	92,3931	100,7125
	10	-56,3331*	1,28663	,000	-59,9300	-52,7361
10	7	199,6304*	4,62599	,000	186,6979	212,5630
	8	152,8859*	1,74975	,000	147,9942	157,7775
	9	56,3331*	1,28663	,000	52,7361	59,9300

*. Die mittlere Differenz ist auf der Stufe .05 signifikant.



Abbildungen 10 Scheffé-Test/Diagramm der Mittelwerte Klassenstufe/Testwert Lesen

Das Diagramm der Mittelwerte zeigt steigende Testwerte bei höherer Klassenstufe für die deutschen Schülerinnen und Schüler im Datensatz des nationalen Ergänzungstests. Die Mittelwertdifferenzen sind jeweils statistisch signifikant ($p < 0,01$).

2.3 Bewertung

Sowohl Korrelationsverfahren (fallbasiert/aggregiert), Regressionsanalyse als auch die Varianzanalyse zeigen einen statistisch signifikanten Zusammenhang zwischen dem variierenden Schulalter der Teilnehmerstaaten und deren Lesetestwerte. Somit kann die Hypothese, wonach das zwischen den Teilnehmerstaaten variierende Schulalter keine Erklärung für den Leistungsvorsprung darstellt, verworfen werden.

Der hier nachgewiesene Einfluss von Schulalter (nomineller Lernzeit) auf die Testleistung wirft generell die Frage nach der Vergleichbarkeit verschiedener Länder, ihrer Schulsysteme und Schülerpopulationen auf.

2.4 Alternativer Ansatz

2.4.1 Vergleichbarkeit der Teilnehmerstaaten

Mit der Frage nach der Vergleichbarkeit der Schülerinnen und Schüler der Teilnehmerstaaten ist auch die Problematik der Unterschiedlichkeit der Schulsysteme verbunden. Ein adjustierter, nachträglicher Vergleich beinhaltet sowohl eine noch präzise zu konkretisierende Definition des Untersuchungsgegenstandes als auch den Einsatz einer geeigneten Methode.

Insofern die technische Frage zu klären wäre, läge die Anforderung darin, einerseits eine Vielzahl von Variablen unterschiedlichen Skalenniveaus kontrollieren und andererseits die nach der Kontrolle berechneten geschätzten Leistungstestwerte darstellen zu können. Ein solches Verfahren stellt gegebenenfalls die Kovarianzanalyse dar. Der Einsatz der Kovarianzanalyse wirft natürlich eine Vielzahl neuer Fragen auf. Vergleiche auf der Basis einzelner Klassenstufen werden ebenfalls durchgeführt. Sie berücksichtigen jedoch nicht systematische Unterschiede hinsichtlich struktureller Zusammensetzung der einzelnen Klassenstufen.

Der Einsatz kovarianzanalytischer Verfahren folgt dem Vorgehen der nationalen Ergänzungsstudie (PISA-E).³⁹

³⁹ Baumert, Artelt, Carstensen, Sibberns, Stanat: Untersuchungsgegenstand, Fragestellungen und technische Grundlagen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. S. 35

Die gewichtete koreanische Stichprobe weist eine Differenz von 11,76 (ungewichtet: 10,74) Prozentpunkten zugunsten der männlichen Schüler auf. Dass die Variable Geschlecht überhaupt einen wesentlichen Einfluss auf die Lesetestleistung hat, zeigte sich bereits in der Regressionsanalyse.

Folgende Abbildung zeigt die Verteilung des Anteils der Schülerinnen und Schüler (Variable Geschlecht) über die Teilnehmerstaaten (Datensatz: gewichtet/ungewichtet). Es wird deutlich, dass aufgrund der disproportionalen Verteilung, die durch die Gewichtung nicht ausgeglichen wurde, die Variable Geschlecht im Vergleich der internationalen Testwerte mit in Betracht gezogen werden muss.

SEX*) COUNTRY	Female		Male	
	Female	Male	Female	Male
AUSTRALIA	48,15	51,85	47,50	52,50
AUSTRIA	49,95	50,05	52,09	47,91
BELGIUM	49,40	50,60	47,86	52,14
BRAZIL	52,38	47,62	53,96	46,04
CANADA	49,87	50,13	50,08	49,92
CZECH REPUBLIC	53,59	46,41	51,67	48,33
DENMARK	49,83	50,17	49,69	50,31
FINLAND	51,54	48,46	51,42	48,58
FRANCE	50,74	49,26	51,34	48,66
GERMANY	51,36	48,64	49,65	50,35
GREECE	49,32	50,68	49,80	50,20
HUNGARY	49,14	50,86	49,59	50,41
ICELAND	50,36	49,64	50,43	49,57
IRELAND	51,92	48,08	50,44	49,56
ITALY	51,31	48,69	49,28	50,72
JAPAN	49,54	50,46	50,52	49,48
KOREA, REPUBLIC OF	44,63	55,37	44,12	55,88
LATVIA	52,24	47,76	51,31	48,69
LIECHTENSTEIN	49,68	50,32	49,84	50,16
LUXEMBOURG	50,11	49,89	50,06	49,94
MEXICO	49,89	50,11	50,01	49,99
NETHERLANDS	50,26	49,74	49,29	50,71
NEW ZEALAND	48,96	51,04	49,74	50,26
NORWAY	49,34	50,66	48,99	51,01
POLAND	47,54	52,46	49,13	50,87
PORTUGAL	52,62	47,38	51,96	48,04
RUSSIAN FEDERATION	50,28	49,72	50,08	49,92
SPAIN	50,93	49,07	50,76	49,24
SWEDEN	49,12	50,88	49,24	50,76
SWITZERLAND	49,82	50,18	49,81	50,19
UNITED KINGDOM	49,73	50,27	50,41	49,59
UNITED STATES	52,56	47,44	51,61	48,39

*) percent

PISA 2000 database : http://pisaweb.acer.edu.au/oeod/oeod_pisa_data_s1.html
data: intstud_read.txt

Abbildung 11 Verteilung *Geschlecht*/Teilnehmerstaat (Angaben in Prozent)

2.4.2 Rückhaltequote (Retentivität)

Prinzipiell müssten weitere Variablen berücksichtigt werden. Nicht statistisch kontrolliert werden kann an dieser Stelle, mangels Kenntnis der Grundgesamtheit, die Auswirkung einer variierenden Ausschöpfungsquote der Stichprobe sowie unterschiedliche Rückhaltequoten der Schulsysteme (in der PISA-Studie als *Retentivität* bezeichnet und in der TIMSS anhand eines so genannten TIMSS Coverage Index –TCI- gemessen). Der TCI wird dem Technical Report⁴⁰ entsprechend wie folgt definiert

*The TCI was computed by forming a ratio of the size of the student population covered by the TIMSS sample, as estimated from the sample itself, to the size of the school-leaving age cohort, which was derived from official population census figures supplied by each country. The numerator in this expression is the total enrollment in the grades tested by TIMSS, estimated from the weighted sample data. This estimate corresponds to the size of the population to which the TIMSS results generalize, and makes appropriate provision for student non-response. It does not include students who are no longer attending school, or students who were excluded from the sample on grounds of physical or other disability. It also does not include students who were repeating the final grade. The denominator in the expression is an estimate of the school-leaving age cohort size. Since the age at which upper secondary students may leave school varies, TIMSS estimated the size of the school-leaving age cohort by taking the average of the size of the 1995 age cohorts for 15-, 16-, 17-, 18-, and 19-year-olds in each country. (Although the general procedure was to base the estimate on the 15-19 age group, there were exceptions. For example, in Germany, the estimate was based on the 17-19 age group.)*⁴¹

Die Rückhaltequoten der einzelnen Teilnehmerstaaten unterscheiden sich deutlich. So werden nur noch 85% des Altersjahrgangs in Korea beschult, 95% der österreichischen, 79% der liechtensteiner und sogar nur noch rund 50% der mexikanischen und brasilianischen 15-Jährigen.⁴² In diesen Fällen ist mit einer Überschätzung der Testwerte zu rechnen. Da mit sinkender Rückhaltequote (Retentivität) eines Schulsystems die Selektivität von Population und Stichprobe unter Umständen steigt, ist eine negative

⁴⁰ TIMSS Technical Report / Volume III / Implementation and Analysis S.18

⁴¹ TIMSS Technical Report / Volume III / Implementation and Analysis S.18

⁴² Baumert, Stanat, Demmrich: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 36

Korrelation zwischen Rückhaltequote (Retentivität) und Testergebnis nicht auszuschließen.

Baumert, Bos und Watermann stellen im Gegensatz zu dieser Überlegung die Hypothese auf, dass mit zunehmender Rückhaltequote (Retentivität) eines Schulsystems die Testleistungen eher steigen.

Umso überraschender ist der empirische Befund der TIMSS/III-Grundbildungsuntersuchung, dass sich ein straffer positiver Zusammenhang zwischen Retentivitätsrate beziehungsweise TCI einerseits und Grundbildungsniveau andererseits nachweisen lässt. Die Korrelation zwischen TCI und mittlerem nationalen Grundbildungsniveau liegt bei $r = .56$ (...) (Schließt man Südafrika als potentiellen Ausreißer aus den Berechnungen aus, ändern sich die Korrelationen praktisch nicht.) Je höher die Haltekraft eines Schulsystems in der Sekundarstufe II ist, desto besser fallen die mathematisch-naturwissenschaftlichen Grundbildungsleistungen aus - trotz geringerer Selektivität der Stichprobe.⁴³

(...)

Fasst man die Befunde der Analysen zusammen, so ergibt sich folgendes Bild: Wider Erwarten lässt sich ein straffer Zusammenhang zwischen der Haltekraft eines Bildungssystems und dem erreichten Grundbildungsniveau nachweisen. Dieser Befund ist erwartungswidrig, da mit abnehmender Retentivität des Systems die Untersuchungspopulationen zunehmend positiv selektiert sind.⁴⁴

Da die Rückhaltequote (Retentivität) über die einzelnen Teilnehmerstaaten stark variiert, ist es von Interesse, deren Einfluss auf die Testleistung zu klären. Als Vorlage liegen die Daten der Dritten Internationalen Mathematik und Naturwissenschaftsstudie, TIMSS/III, im Original der deutschen Version wie folgt vor:

⁴³ Klieme, Baumert, Bos: Mathematisch-naturwissenschaftliche Grundbildung im internationalen Vergleich. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 138 ff

⁴⁴ Klieme, Baumert, Bos: Mathematisch-naturwissenschaftliche Grundbildung im internationalen Vergleich. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 142

Land	Bruttonationalprodukt pro Kopf (in 1.000 US-Dollar)	Anteil der öffentlichen Ausgaben für das Schulsystem in Prozent des Bruttonationalprodukts	Retentivität im Abschlussjahr der Sekundarstufe II (in % der Alterskohorte)	TIMSS Coverage Index (TCI)	Relativer Besuch voruniversitärer Bildungswege (in % der Alterskohorte)	Mittleres Alter im Abschlussjahr	Mathematisch-naturwissenschaftliches Grundbildungsniveau (Mittelwert)
Australien	18,0	3,7	72	68	37	17,7	525
Dänemark	28,1	4,8	59	58	33	19,1	528
Deutschland	25,6	2,4	84	78	25	19,5	496
Frankreich	23,5	3,6	85	84	75	18,8	505
Island	24,6	4,8	55	55	45	21,2	541
Italien	19,3	2,9	74	74	55	18,7	475
Kanada	19,6	4,6	78	70	54	18,6	526
Neuseeland	13,2	3,2	71	70	70	17,6	525
Niederlande	22,0	3,3	92	78	34	18,5	559
Norwegen	26,5	5,3	85	84	48	19,5	536
Österreich	25,0	4,2	93	76	33	19,1	519
Schweden	23,6	4,9	71	71	47	18,9	555
Schweiz	37,2	3,7	84	82	19	19,8	531
Slowenien	7,1	4,2	93	88	80	18,8	514
Südafrika	3,0	5,1	49	49	49	20,1	352
Tschechien	3,2	3,8	83	78	36	17,8	476
Ungarn	3,8	4,3	65	65	18	17,5	477
USA	25,9	4,0	65	63	35	18,1	471
Zypern	10,4	3,6	61	48	44	17,7	447

Abbildung 12 TCI-Index Lesetestwert Quelle: Abb. IV.1 TIMSS/III, S. 136⁴⁵

Die Angaben der deutschen Version stimmen nicht in jedem Fall mit denen des Technical Report der TIMSS (Vol.III)⁴⁶ überein.

Italien weist in der tabellarischen Darstellung der deutschen Version einen TCI-Wert von 74%, anstatt 52% entsprechend den Angaben des TIMSS Technical Report (Vol.III) auf. Deutschland weist einen TCI-Wert von 75% im Technical Report auf, anstatt 78% in der deutschen Ausgabe. Russland und Litauen fehlen wiederum in der tabellarischen Darstellung der deutschen Studie. Die Korrelationen berechnen sich wie folgt:

⁴⁵ Original siehe Anhang

⁴⁶ TIMSS Technical Report / Volume III / Implementation and Analysis S.19

Country	Fall- nummer	TIMSS_III	TCI	BSP*	Anteil BSP*
Australia	1	525	68	18	3,7
Austria	2	519	76	25	4,2
Canada	3	526	70	19,6	4,6
Cyprus	4	447	48	10,4	3,6
Czech Republic	5	476	78	3,2	3,8
Denmark	6	528	58	28,1	4,8
France	7	505	84	23,5	3,6
Germany	8	496	75**	25,6	2,4
Hungary	9	477	65	3,8	4,3
Iceland	10	541	55	24,6	4,8
Italy	11	475	52**	19,3	2,9
Lithuania	12	465	43**		
Netherlands	13	559	78	22	3,3
New Zealand	14	525	70	13,2	3,2
Norway	15	536	84	26,5	5,3
Russian Federation	16	476	48**		
Slovenia	17	514	88	7,1	4,2
South Africa	18	352	49	3	5,1
Sweden	19	555	71	23,6	4,9
Switzerland	20	531	82	37,2	3,7
United States	21	471	63	25,9	4

*) Angaben nach TIMSS/III deutsche Ausgabe

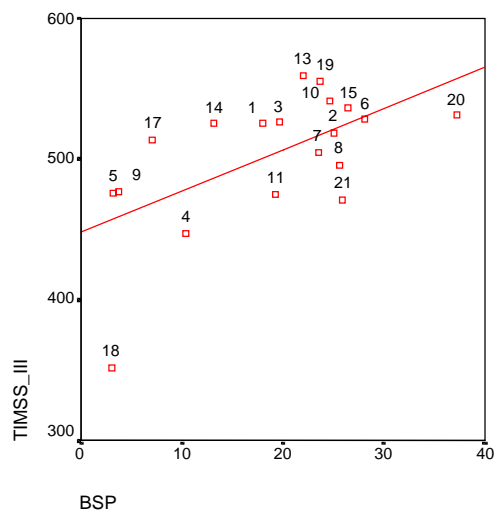
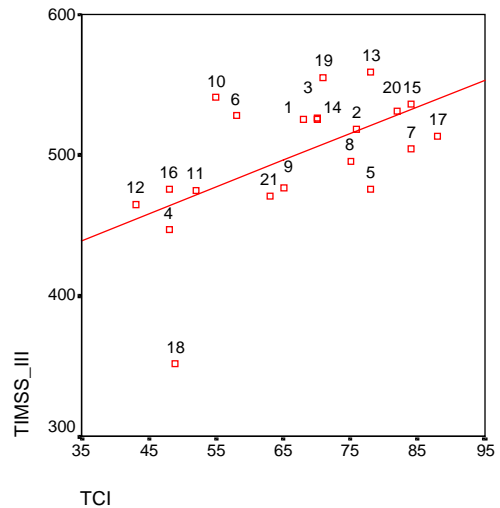
**) korrigierte Angaben nach TIMSS Technical Report Vol. III

TIMSS_III : Testergebnis der TIMSS

TCI : TIMSS Coverage Index

BSP : Bruttosozialprodukt pro Kopf in Tsd. US-Dollar

Anteil BSP: Anteil der öffentlichen Ausgaben für das Schulsystem in % des BSP



Korrelationen

	TCI
TIMSS_III Korrelation nach Pearson	,563**
Signifikanz (2-seitig)	,008
N	21

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

Korrelationen

	TCI
Spearman-Rho TIMSS_III Korrelationskoeffizient	,531*
Sign. (2-seitig)	,013
N	21

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 13 TCI – Lesetestwert/Korrelation

Die Korrelation zwischen mathematisch-naturwissenschaftlichem Testwert (TIMSS_III) und TIMSS Coverage Index (TCI) beträgt $r_p = 0,563$ und $r_s = 0,531$. Beide Werte sind statistisch signifikant ($p < 0,01$ bzw. $p < 0,05$). Demnach besteht ein positiver linearer Zusammenhang zwischen beiden Variablen. Es wird jedoch deutlich, dass die Testwerte der Teilnehmerstaaten mit erfolgsrelevanten Indikatoren, zum Beispiel dem

Bruttosozialprodukt (BSP) der Teilnehmerstaaten, kovariieren mit $r_p = 0,592$ ($p < 0,01$) und $r_s = 0,510$ ($p < 0,05$).

	TCI	BSP
TIMSS_III Korrelation nach Pearson	,563**	,592**
Signifikanz (2-seitig)	,008	,008
N	21	19

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

	TCI	BSP
Spearman-Rho TIMSS_III Korrelationskoeffizient	,531*	,510*
Sig. (2-seitig)	,013	,026
N	21	19

* Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 14 Korrelation TIMSS Testwert – TCI - BSP

Da weitere Variablen mit den TIMSS-Testwerten kovariieren und Angaben zum Schulalter in Bezug auf die TIMSS Studie hier nicht vorliegen sowie Fragen bezüglich der Definition des TCI-Wertes aber insbesondere dessen Datenbasis nicht geklärt werden konnten, soll der Einfluss der Rückhaltequote auf das Testergebnis internationaler Vergleichsstudien anhand der PISA-Studie geprüft werden. Es stellt sich die Frage, ob anhand der PISA-Studie der Zusammenhang von Rückhaltequote und Testwert wiederholt nachgewiesen werden kann.

Der Anteil der noch beschulten 15-Jährigen wird im Technical Report der PISA-Studie veröffentlicht.⁴⁷ Diese Kenngröße kommt der Fragestellung der so genannten Rückhaltequote (Retentivität) des Schulsystem möglicherweise näher als der TCI-Wert. Folgende Darstellung faßt zentrale Daten diesbezüglich zusammen. Die Darstellung umfaßt weitere Teilnehmerstaaten, die in der ursprünglichen Studie (in 2000) noch nicht aufgeführt wurden.

⁴⁷ Adams, Wu (Hrsg.): PISA 2000 Technical Report Tabelle 31 (S. 135-136)/Hinweis Raymond Adams

COUNTRY	PVREAD	HISEI	GRADE	SEX_male	ALL_15	Enroll_15	ENROLL_p	GNI_pc	N
ALBANIA	349,02	45,55	9,71	48,98%				1740	4652
ARGENTINA	417,41	43,35	9,59	43,61%				3810	3608
AUSTRALIA	527,79	52,25	10,10	52,50%	266878	248908	93,27%	21950	4939
AUSTRIA	507,53	49,72	9,43	47,91%	95041	90354	95,07%	26810	4635
BELGIUM	507,31	48,95	9,61	52,14%	121121	119055	98,29%	25760	6371
BRAZIL	395,20	43,93	8,50	46,04%	3464330	1841843	53,17%	2720	4469
BULGARIA	430,54	49,99	9,11	51,46%				2130	4195
CANADA	534,49	52,83	9,84	49,92%	403803	396423	98,17%	24470	28751
CHILE	409,64	39,87	9,48	46,96%				4360	4699
CZECH REPUBLIC	491,94	48,31	9,51	48,33%	134627	132508	98,43%	7150	5273
DENMARK	496,56	49,73	8,97	50,31%	53693	52161	97,15%	33570	3953
FINLAND	546,99	50,00	8,89	48,58%	66571	66561	99,98%	27060	4770
FRANCE	504,38	48,27	9,51	48,66%	788387	788387	100,00%	24730	4389
GERMANY	483,99	48,85	9,06	50,35%	927473	924549	99,68%	25270	4934
GREECE	474,46	47,76	10,05	50,20%	128175	124656	97,25%	13230	4468
HONG KONG	525,15	42,25	9,83	50,15%				25860	4224
HUNGARY	479,83	49,53	9,24	50,41%	120759	115325	95,50%	6350	4746
ICELAND	506,58	52,73	10,00	49,56%	4062	4044	99,56%	30910	3298
INDONESIA	370,76	36,38	9,51	48,86%				810	6811
IRELAND	526,68	48,43	9,50	49,56%	65339	64370	98,52%	27010	3737
ISRAEL	450,96	55,30	9,98	41,78%				16240	3531
ITALY	487,36	47,08	9,86	50,72%	584417	574864	98,37%	21570	4864
JAPAN	522,03	50,54	10,00	49,48%	1490000	1485269	99,68%	34180	2019
KOREA, REPUBLIC OF	525,03	42,80	10,00	55,88%	712812	602605	84,54%	12030	4617
LATVIA	459,14	50,15	9,38	48,69%	38000	35981	94,69%	4400	3713
LIECHTENSTEIN	482,58	47,46	8,85	50,27%	415	326	78,55%		297
LUXEMBOURG	440,40	44,79	9,05	49,94%	4556	4556	100,00%	45740	3201
MEXICO	421,74	42,48	9,36	49,99%	2127504	1098605	51,64%	6230	4222
NETHERLANDS	531,64	50,85	9,43	50,71%	178924	178924	100,00%	26230	2438
NEW ZEALAND	528,60	52,20	10,98	50,26%	54220	51464	94,92%	15530	3523
NORWAY	505,71	53,91	10,00	51,01%	52165	51587	98,89%	43400	4037
PERU	327,23	40,42	9,31	50,10%				2140	4152
POLAND	480,01	46,03	9,00	50,87%	665500	643528	96,70%	5280	3396
PORTUGAL	469,89	43,85	9,25	48,04%	132325	127165	96,10%	11800	4426
ROMANIA	427,19	45,53	8,89	47,32%				2260	4514
RUSSIAN FEDERATION	461,91	49,38	9,69	49,92%	2268566	2259985	99,62%	2610	6512
SPAIN	492,84	44,99	9,70	49,24%	462082	451685	97,75%	17040	5923
SWEDEN	515,39	50,57	8,98	50,76%	100940	100940	100,00%	28910	4313
SWITZERLAND	494,80	49,21	8,93	50,19%	81350	79232	97,40%	40680	5880
THAILAND	431,18	33,02	9,38	41,18%				2190	4798
MACEDONIA	372,96	46,79	9,25	51,16%				1980	4266
UNITED KINGDOM	523,54	51,26	10,68	49,59%	731743	705875	96,46%	28320	8843
UNITED STATES	503,71	52,40	9,53	48,39%	3876000	3836000	98,97%	37870	3242

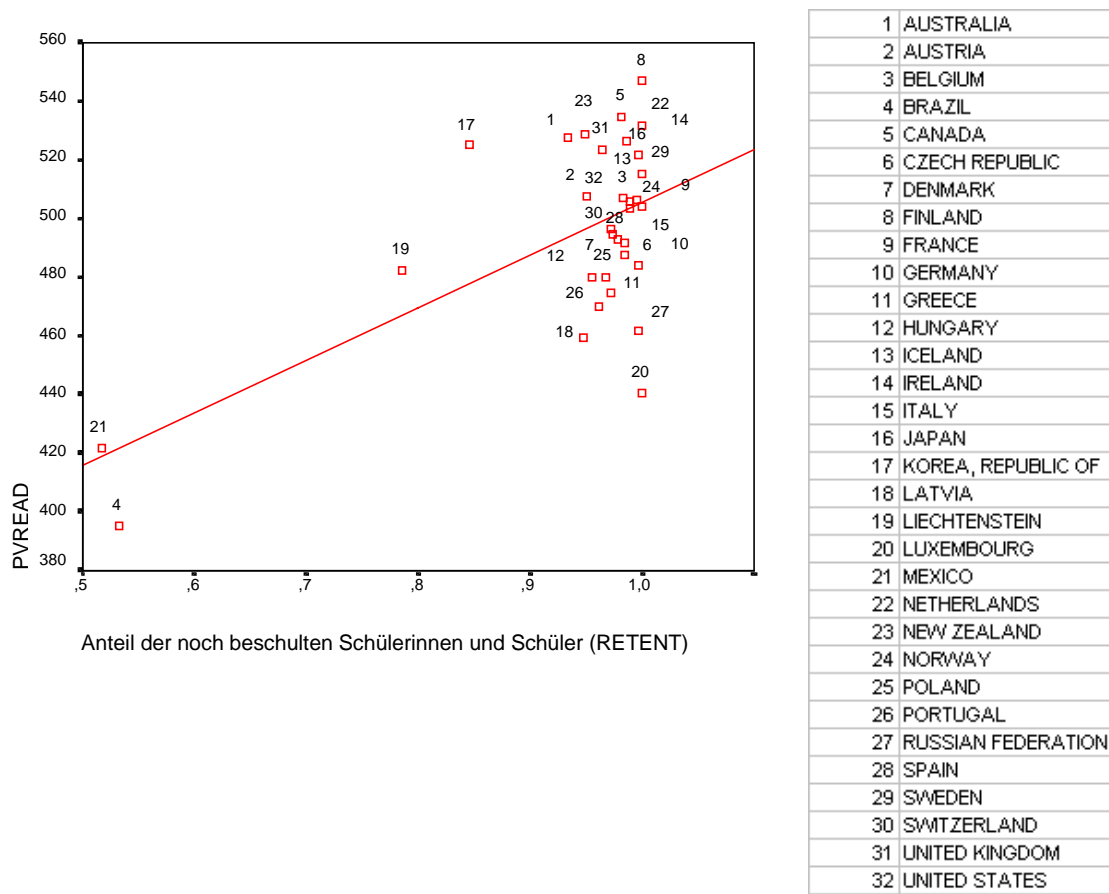
PVREAD	Mean reading score PISA2000 (pv1read)
HISEI	Highest socio-economic index of family (PISA 2000)
GRADE	Mean grade
SEX_male	Percentage of male students
ALL_15	All 15-year-olds
Enroll_15	Enrolled 15-year-olds
ENROLL_p	Percentage of enrolled 15-year-olds
GNI_pc	GNI \$ per capita, 2003 WORLDBANK Atlas method
N	(Sub-) Sample Size

Abbildung 15 Rückhaltequote unter 15-Jährigen in PISA

Insgesamt sind neun Variablen über den erweiterten Datensatz dargestellt. Gemeinsam mit dem aus dem Datensatz berechneten mittleren Lesewert wird der durchschnittliche sozioökonomische Index (HISEI), die mittlere Klassenstufe (GRADE), der Anteil

männlicher Schüler (SEX_MALE), sämtliche 15-Jährige des Teilnehmerstaates, die Anzahl der die Schule besuchenden 15-Jährigen, der daraus berechnete prozentuale Anteil der noch beschulten 15-Jährigen, das von der Weltbank publizierte Pro-Kopf-Einkommen in US\$⁴⁸ sowie die Anzahl der in der Stichprobe berücksichtigten Schülerinnen und Schüler dargestellt.

Sollte ein positiver Zusammenhang zwischen der Rückhaltequote (Retentivität) des Schulsystems und der Testleistung existieren, müsste dieser in PISA erneut nachweisbar sein. Folgende Abbildung trägt den prozentualen Anteil der noch beschulten Altersgruppe gegen die Leseleistung ab.



Correlations		PVREAD
RETENT	Pearson Correlation	,630**
	Sig. (2-tailed)	,000
	N	32

** . Correlation is significant at the 0.01 level

Correlations			PVREAD
Spearman's rho	RETENT	Correlation Coefficient	,252
		Sig. (2-tailed)	,164
		N	32

Abbildung 16 Rückhaltequote (Retentivität) und Testleistung in PISA (I)

⁴⁸ (Total GNI 2003, Atlas method) Download www.worldbank.org/data/databytopic/GNIPC.pdf

Zunächst lässt sich mit dem gegenüber Extremwerten empfindlichen $r_p = 0,63$ ($p < 0,01$) ein statistisch signifikanter Zusammenhang nachweisen. Jedoch der in dieser Hinsicht robuste Korrelationskoeffizient nach Spearman weist mit $r_s = 0,252$ ($p = 0,164$) keinen statistisch signifikanten Zusammenhang auf. Werden Mexiko und Brasilien entfernt, zeigt auch der parametrische Korrelationskoeffizient (Pearson) keinen signifikanten linearen Zusammenhang.

		PVREAD
RETENT	Pearson Correlation	,028
	Sig. (2-tailed)	,885
	N	30

		PVREAD	
Spearman's rho	RETENT	Correlation Coefficient	,092
		Sig. (2-tailed)	,627
		N	30

Abbildung 17 Rückhaltequote (Retentivität) und Testleistung in PISA (II)

Mit $r_p = 0,028$ ($p = 0,885$) und $r_s = 0,092$ ($p = 0,627$) muss von einem so genannten Ausreißerproblem ausgegangen werden.

Anhand der PISA-Daten konnte der positive Zusammenhang von Rückhaltequote (Retentivität) des Schulsystems eines Teilnehmerstaates und dessen Testleistungsergebnis nicht nachgewiesen werden.

2.4.3 Ergänzungsunterricht

Weitere Probleme bezüglich Vergleichbarkeit der Schulsysteme einzelner Staaten ergeben sich in Hinblick auf zusätzliche Bildungsangebote. Hier zeigen sich systematische Unterschiede. So besuchen in Japan 57,5% der 15-jährigen Schülerinnen und Schüler neben der regulären Schule regelmäßig Ergänzungsunterricht, in Deutschland lediglich 3,0%⁴⁹.

⁴⁹ 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. 417

Land	Ergänzungs- unterricht	Land	Ergänzungs- unterricht
AUSTRALIA	6,1%	KOREA, REPUBLIC OF	36,0%
AUSTRIA	2,2%	LATVIA	14,8%
BELGIUM	4,4%	LIECHTENSTEIN	4,2%
BRAZIL	3,7%	LUXEMBOURG	6,7%
CANADA	3,3%	MEXICO	11,2%
CZECH REPUBLIC	6,8%	NETHERLANDS	k.A.
DENMARK	3,1%	NEW ZEALAND	3,5%
FINLAND	0,5%	NORWAY	1,1%
FRANCE	8,6%	POLAND	18,1%
GERMANY	3,0%	PORTUGAL	5,7%
GREECE	k.A.	RUSSIAN FEDERATION	15,1%
HUNGARY	25,0%	SPAIN	13,1%
ICELAND	3,9%	SWEDEN	1,3%
IRELAND	8,6%	SWITZERLAND	2,4%
ITALY	0,6%	UNITED KINGDOM	6,3%
JAPAN	57,5%	UNITED STATES	5,1%

Abbildung 18 Ergänzungsunterricht Quelle: PISA 2000

Der Anteil zusätzlichen Unterrichts variiert erheblich. Die Frage der Vergleichbarkeit wird später noch einmal Gegenstand der Diskussion.

2.4.4 Gewichtung

Die Testrohre wurden in der PISA-Studie nachträglich durch eine Gewichtungsvariable adjustiert. Voraussetzung für den Einsatz von Gewichtungsvariablen sind sowohl präzise Kenntnisse der für die Untersuchung relevanten Variablen als auch der strukturellen Zusammensetzung der Grundgesamtheit.

Die Gewichte wurden bei WESTAT berechnet. Ziehungswahrscheinlichkeiten sowie systematische Unterschiede zwischen Stichprobe und Grundgesamtheit können durch Gewichtungen ausgeglichen werden.⁵⁰

*Eine Überprüfung der bei WESTAT errechneten Gewichte ergab, dass die Populationsverhältnisse durch die Gewichte sehr gut abgebildet werden.*⁵¹

Diese Aussage entspricht nicht den Ausführungen des Technical Report. In Österreich beispielsweise ergeben sich vergleichbare Probleme wie in den Bundesländern Bayern

⁵⁰ Adams, Wu (Hrsg.): PISA 2000 Technical Report. Kapitel 8

⁵¹ Sibberns, Baumert: Stichprobenziehung und Stichprobengewichtung. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 515

und Baden-Württemberg, die jeweils hohe Anteile an Schülern in Berufsschulen aufweisen. In Österreich fehlen genaue Kenntnisse über das vermutlich testleistungsschwächste Segment.

*In Austria, students in vocational schools are enrolled on a part-time/part-year basis. Since the PISA assessment was conducted only at one point in time, these students are captured only partially. Thus, it is not possible to assess how well the students sampled from vocational schools represent the universe of students enrolled in vocational schools, and so those students not attending classes at the time of the PISA assessment are not represented in the PISA results.*⁵²

In Griechenland sind wesentliche Vergleichsdaten nicht verfügbar.

*The sampling frame for Greece had deficient data on the number of 15-year-olds in each school. This resulted in selecting many schools in the sample with no 15-year-olds enrolled, or that were closed (34 of a sample of 175 schools). Often the replacement school was then included, which was not the correct PISA procedure for such cases. Also quite a few sampled schools (a further 34) had relatively few 15-year-olds. (...) Rather than being assessed or excluded, these schools were also replaced, often by larger schools. This resulted in some degree of over-coverage in the PISA sample, as there were 18 replacement schools in the sample for schools that should not have been replaced, and 32 replacements for small schools where the replacement was generally larger than the school it replaced. It is not possible to evaluate what kinds of students might be over-represented (...).*⁵³

Dieses Defizit stellt ein gravierendes Problem zur Bestimmung der adäquaten Repräsentation von Schüleranteilen dar. Der Vergleich gewichteter und ungewichteter Testergebnisse zeigt den Einfluss der GewichtungsvARIABLE auf das Endresultat. Während Österreichs Ergebnis trotz Unkenntnis des vermutlich testleistungsschwächsten Segments durch die Gewichtung um eine Größenordnung von ca. 10 Punkten aufgewertet wird, reduzieren sich die Testwerte der Bundesrepublik Deutschland um ca. 14 Punkte (maximale Abwertung).

⁵² PISA 2000 International Technical Report. S. 183

⁵³ PISA 2000 International Technical Report. S. 185

PV1READ Differenz (gewichtet - ungewichtet)

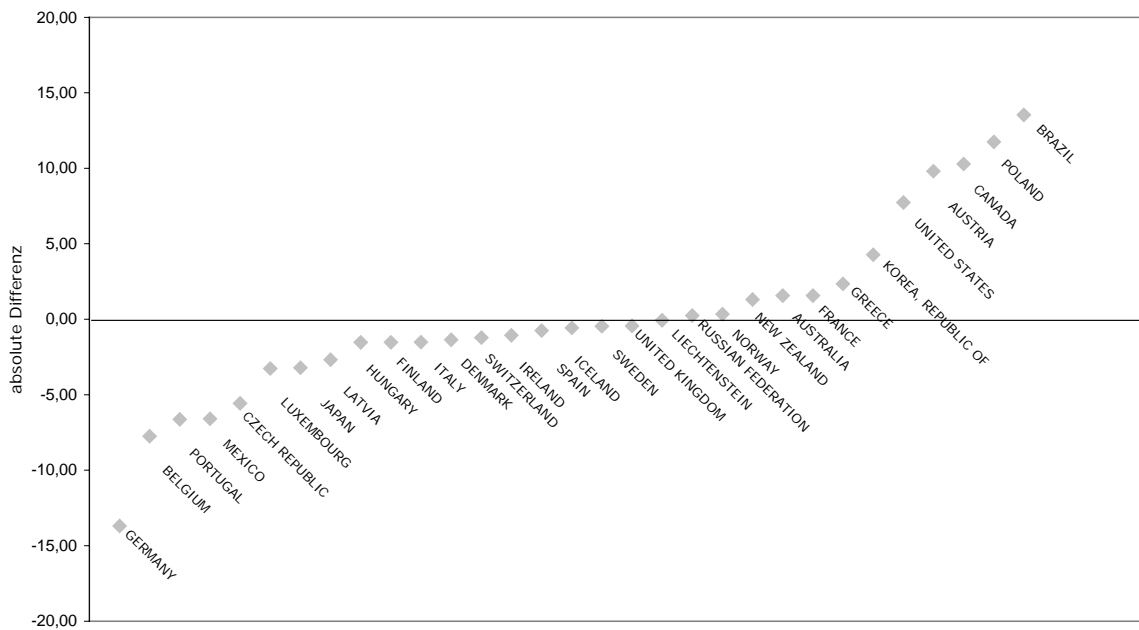


Abbildung 19 Gewichtung/Differenz der Testergebnisse

2.4.5 Darstellung der Testwerte

Aus den bisherigen Ausführungen wird deutlich, dass als Voraussetzung zur Vergleichbarkeit unterschiedlich strukturierter Systeme weitere Variablen, insbesondere Schulalter (*grade*) und *Geschlecht*, kontrolliert werden müssen. Die statistische Kontrolle unterschiedlicher Rückhaltequoten sowie außerschulischen regulären Unterrichts auf das Testergebnis der Schüler ist ein erhebliches methodisches Problem und bleibt einer späteren Betrachtung außerhalb dieser Arbeit vorbehalten.

Im Folgenden werden diejenigen Testwerte dargestellt, die sich unter Kontrolle von lediglich zwei weiteren Variablen, *Klassenstufe* und *Geschlecht*, im Rahmen einer Kovarianzanalyse als so genannte Randmittel schätzen lassen.⁵⁴ Zusätzlich werden die Testwerte nach Klassenstufe berichtet.

⁵⁴ Zur Berechnung geschätzter Randmittel siehe: SPSS Inc. Chicago: SPSS Statistical Algorithms GLM UNIVARIATE AND MULTIVARIATE S.5 ff (<http://support.spss.com/tech/default.asp>)

PISA 2000 Leseverständnis (Klassenstufe 11)

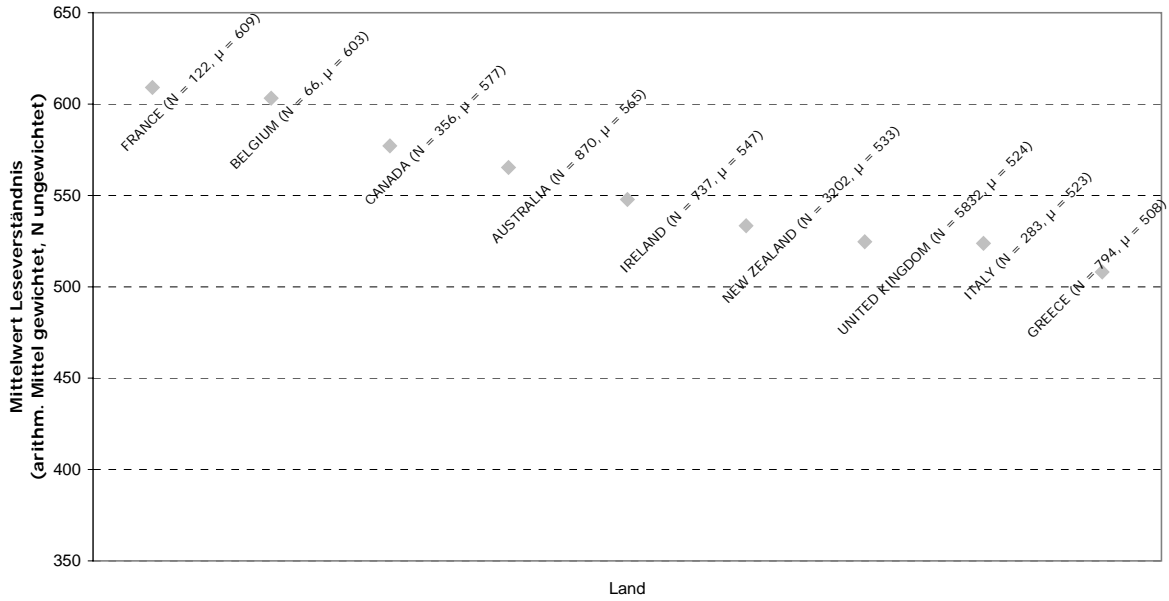


Abbildung 20 Leseverständnis Klassenstufe 11 (gewichtet)

PISA 2000 Leseverständnis (Klassenstufe 10)

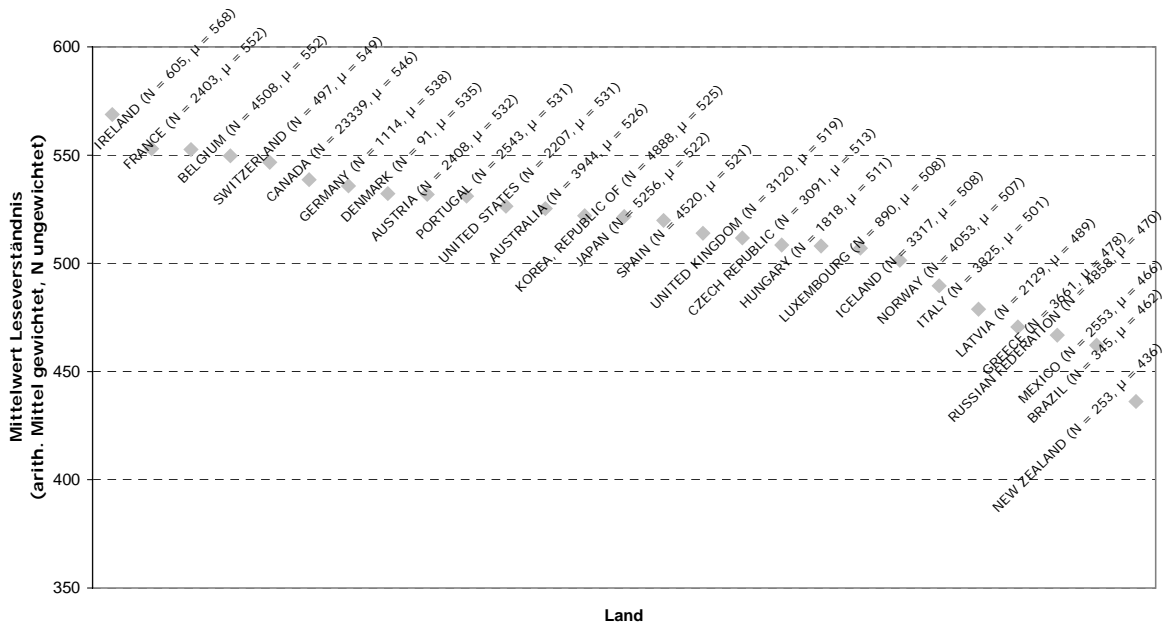


Abbildung 21 Leseverständnis Klassenstufe 10 (gewichtet)

PISA 2000 Leseverständnis (Klassenstufe 9)

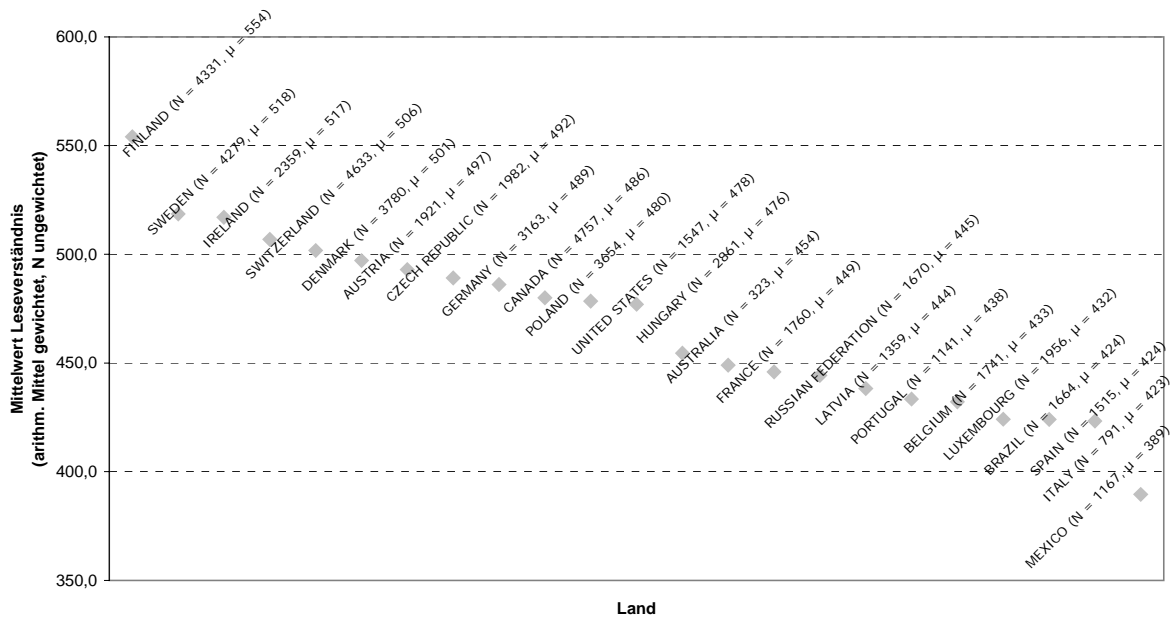


Abbildung 22 Leseverständnis Klassenstufe 9 (gewichtet)

PISA 2000 Leseverständnis (geschätzte Randmittel)

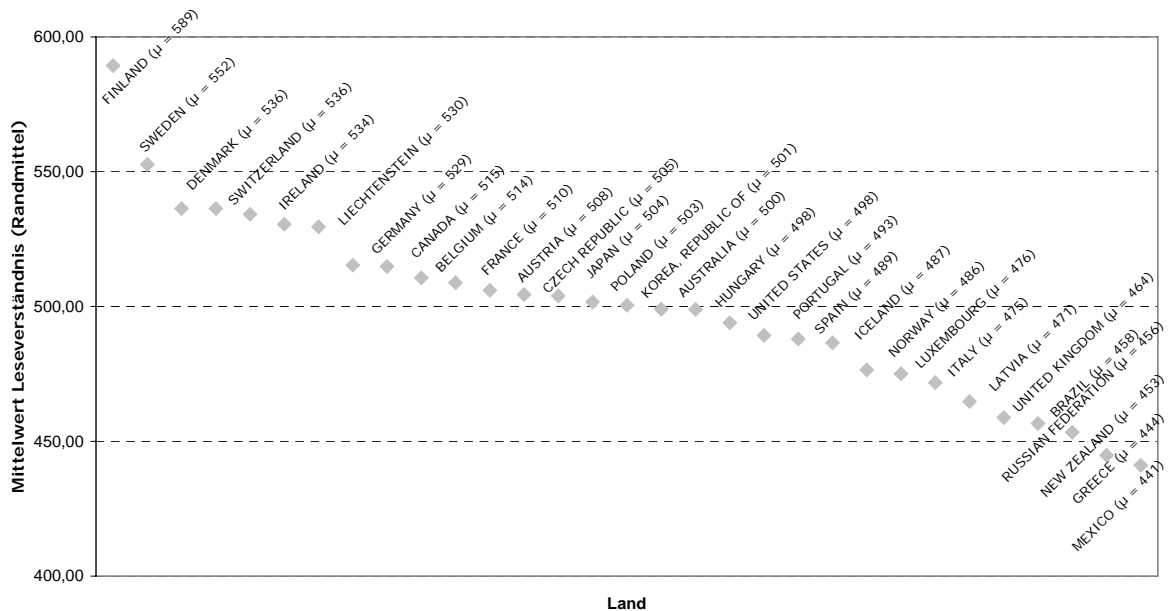


Abbildung 23 Leseverständnis (Randmittel ungewichtet)

**PISA 2000 Leseverständnis
(geschätzte Randmittel)**

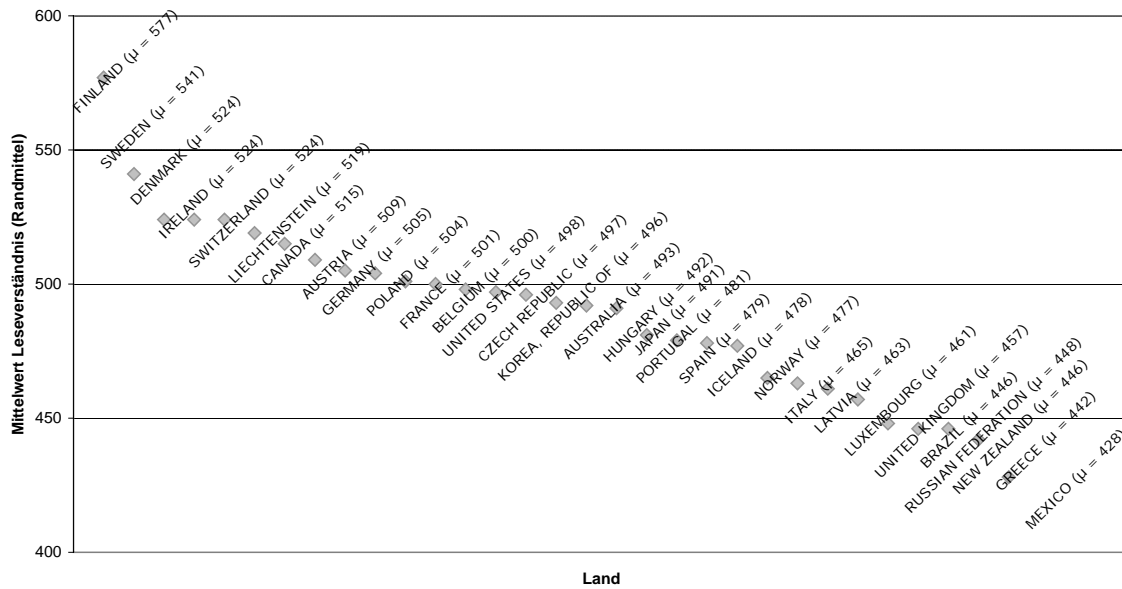


Abbildung 24 Leseverständnis (Randmittel - Datensatz gewichtet)

**PISA 2000 Science
(geschätzte Randmittel)**

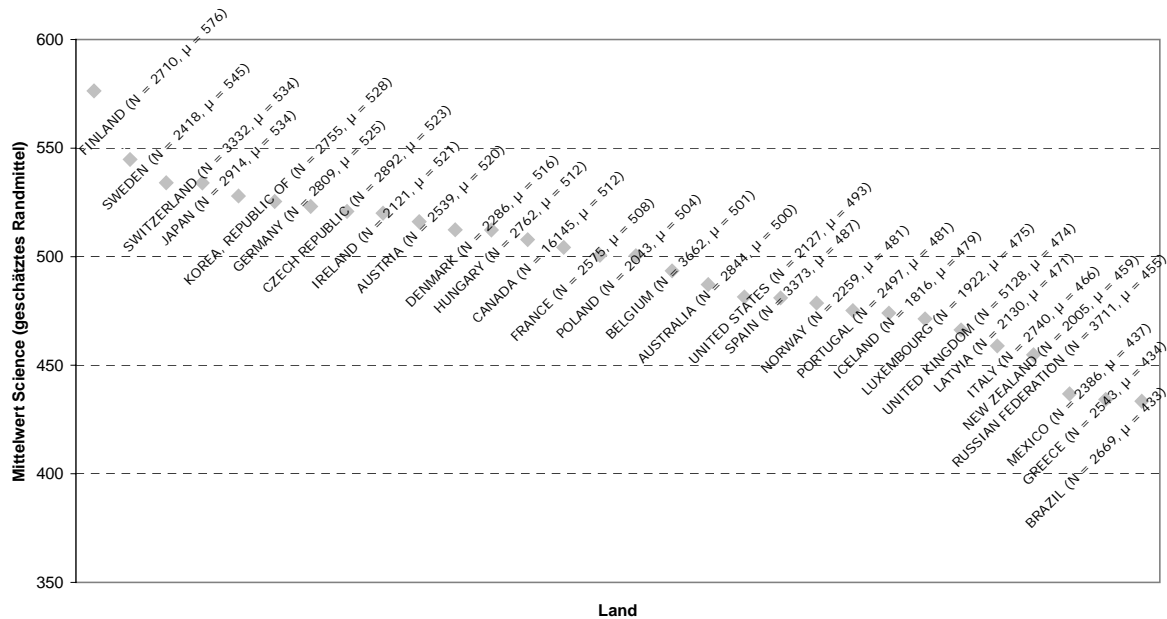


Abbildung 25 Science (Randmittel ungewichtet)

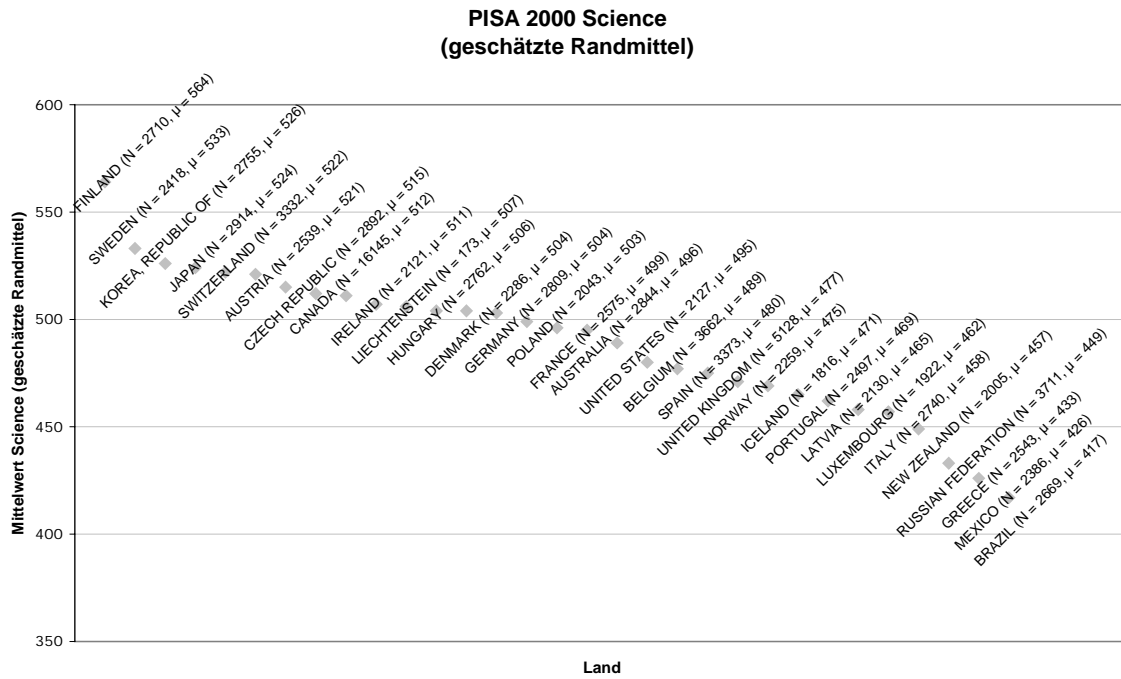


Abbildung 26 Science (Randmittel - Datensatz gewichtet)

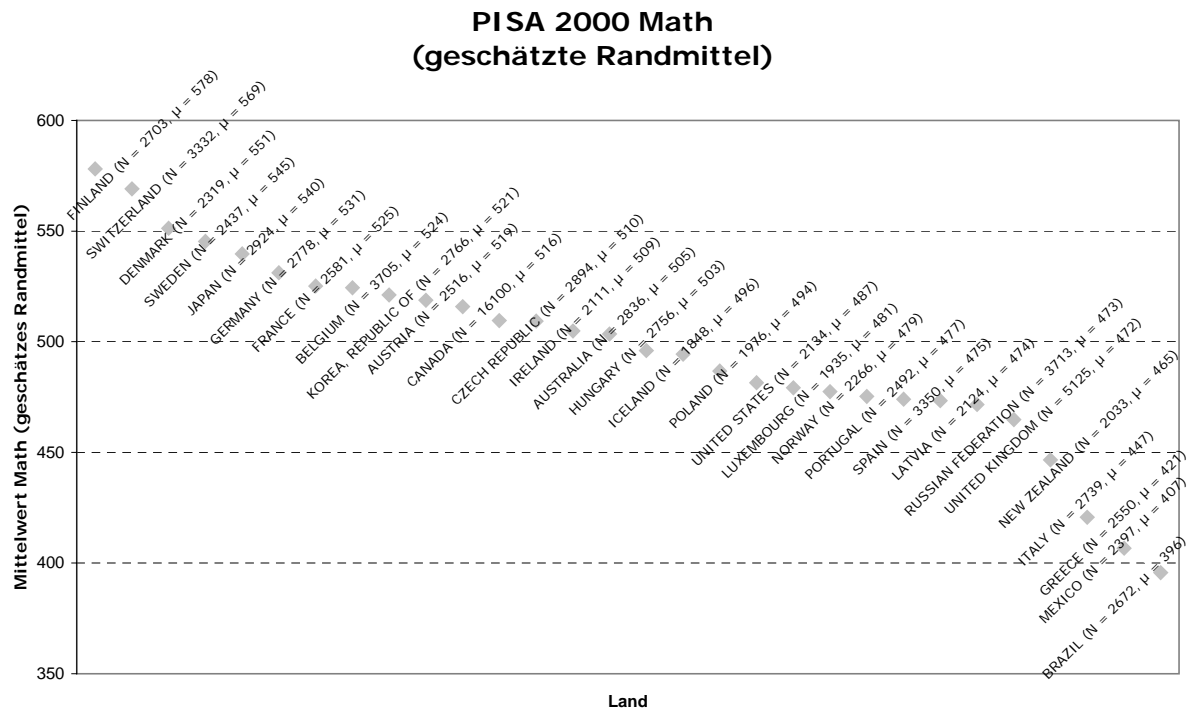


Abbildung 27 Math (Randmittel ungewichtet)

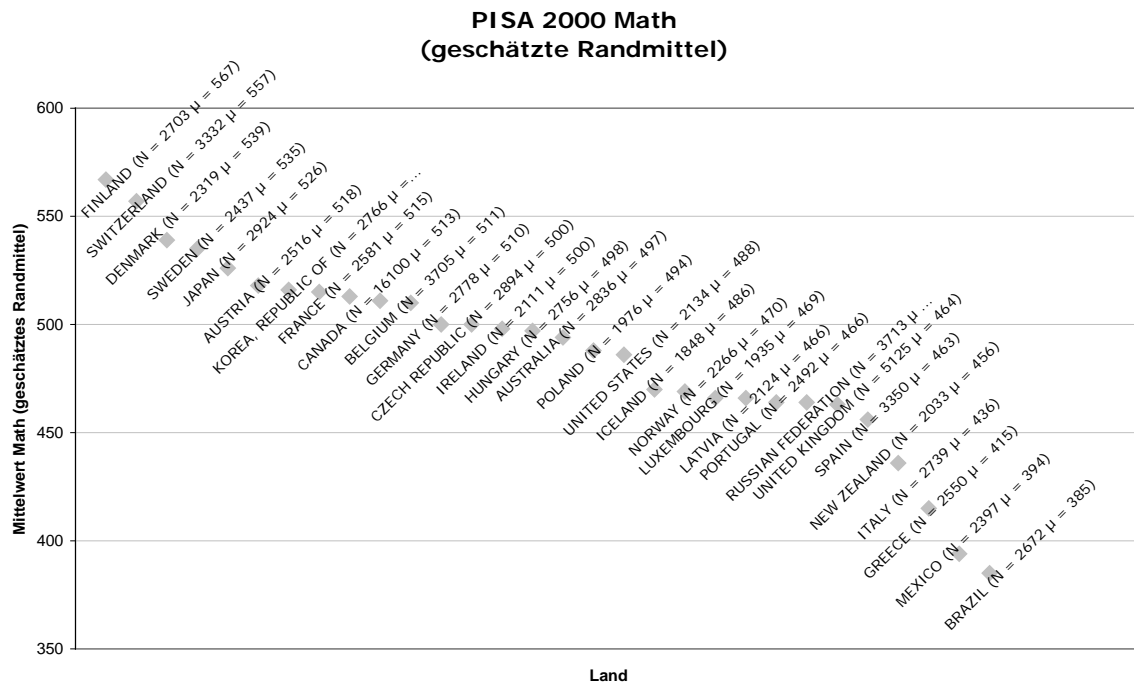


Abbildung 28 Math (Randmittel - Datensatz gewichtet)

2.4.6 Interpretation

Die Schülerinnen und Schüler im deutschen Sample erreichen über die einzelnen Darstellungen einen Testwert jeweils im oberen oder mittleren Bereich. Die Niederlande müssen wegen der niedrigen Ausschöpfungsquote auf Schulebene von nur 27 Prozent bei Mittelwertvergleichen unberücksichtigt bleiben.⁵⁵ Liechtenstein weist eine Differenz zwischen Geburtskohorte und altersgleicher Schulbevölkerung von 21% auf.⁵⁶ Dessen Testwert ist daher möglicherweise verzerrt.

Die Rangreihenfolge ändert sich nach Berücksichtigung der beiden Kontrollvariablen Schulalter und Geschlecht deutlich. Der Mittelwert der neuseeländischen und englischen Schülerinnen und Schüler reduziert sich erheblich. Unter Berücksichtigung unterschiedlicher Einschulungsmodalitäten, die sich auf das Verhältnis von Klassenstufe und tatsächlicher Besuchsdauer beziehen, ließe sich für Neuseeland und das Vereinigte Königreich die Klassenstufe um ein volles Jahr in einem Modellvergleich rein

⁵⁵ Baumert, Stanat, Demmrich: PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 37

⁵⁶ Baumert, Stanat, Demmrich: PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 36

rechnerisch reduzieren. Die geschätzten Randmittel über die korrigierte Variable zeigen für beide Länder, wie zu erwarten, höhere mittlere Werte, die jedoch insbesondere im Falle des Vereinigten Königreichs den Wert der Bundesrepublik Deutschland nicht übersteigen.

2.4.7 Zwischenbewertung

Nach Berücksichtigung der Variablen *Klassenstufe* (grade) und *Geschlecht* zeigen sich zu den ursprünglich berichteten Ergebnissen der PISA-Studie abweichende Testwerte. Ursache hierfür ist die Problematik der Vergleichbarkeit.

*Es hat sich als ausgesprochen schwierig erwiesen, Untersuchungspopulationen zu definieren, die für alle Teilnehmerstaaten eindeutig und zugleich innerhalb der Systematik der nationalen Bildungssysteme sinnvoll sind. Es ist kein Zufall, dass die IEA in ihren Schulleistungsuntersuchungen mit mehreren Populationsdefinitionen experimentiert hat. Je nach theoretischer Perspektive kann man Populationen für Schulleistungsuntersuchungen nach dem Lebensalter zu einem bestimmten Stichtag, oder nach dem Schulalter, definiert über die Zugehörigkeit zur Jahrgangsstufe zum Testzeitpunkt, festlegen. Entsprechend der leitenden Fragestellung, der untersuchten Wissensdomäne und der Abhängigkeit des Wissenserwerbs von schulischen Lerngelegenheiten wird man eher an einer Lebensalter- oder an einer Schulalterdefinition interessiert sein. Im Falle der Untersuchung des Leseverständnisses kann man bei Jugendlichen - weniger bei Grundschulern - einer Festlegung der Population über das Lebensalter den Vorzug geben, da Leseerfahrungen in dieser Altersgruppe möglicherweise stärker vom Lebensalter als von der besuchten Klassenstufe abhängen. Bei der Erhebung mathematischer Kompetenzen wird eine Populationsdefinition nach Schulalter näher liegen, da der Erwerb von mathematischem Wissen in hohem Maße von den schulischen Lerngelegenheiten abhängig ist.*⁵⁷

Die Ergebnisse zeigen, dass in der Klassenstufe ein wichtiger Faktor zur Erklärung der international variierenden Testleistungen der 15-jährigen Schülerinnen und Schüler zu sehen ist.

In Bezug auf die Messung mathematischer Kompetenzen folgen die Autoren der eigenen Analyse nicht. Im Bereich der Messung mathematischer und wissenschaftlicher

⁵⁷ Baumert, Stanat, Demmrich: PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 34

Kompetenz (Science) wurde ebenfalls eine Populationsdefinition auf der Basis des Lebensalters, nicht des Schulalters gewählt.

Dieser Befund macht die hohen Testergebnisse bei gleichzeitig geringer Streubreite der Länder Korea und Japan nachvollziehbar.

In der Darstellung sticht die Besonderheit Koreas hervor: Korea erreicht eine Spitzenleistung mit einer sehr niedrigen Leistungsstreuung. Der untere Leistungsbereich (25. Perzentil) beginnt bei einem Wert von 411; das ist etwa der Kennwert, den der untere Leistungsbereich (25. Perzentil) in Deutschland maximal erreicht. (...) Korea ist damit ein Beispiel für eine gelungene Förderung naturwissenschaftlicher Grundbildung auf einem relativ homogenen hohen Niveau.⁵⁸

Das Ausmaß der Streuung ist selbst ein wichtiger Indikator für die Fähigkeit eines Bildungssystems, die mathematische Bildung seiner Schülerpopulation zu homogenisieren.⁵⁹

In Korea und Japan befinden sich die getesteten 15-jährigen Schülerinnen und Schüler weitgehend beziehungsweise ausschließlich in der 10. Jahrgangsstufe. Eine Streuung über verschiedene Schulalter, wie in den meisten Ländern, existiert praktisch nicht. Der Befund macht ebenfalls verständlich, weshalb Schülerinnen und Schüler im Vereinigten Königreich insgesamt hohe mittlere Leistungen bei großer Streuung erreichen.

Das Vereinigte Königreich liegt in den Leistungen zwar im Mittel um etwa 20 Punkte unter Korea, doch endet hier der obere Leistungsbereich (95. Perzentil) auf einem deutlich höheren Niveau (687), im Vergleich zu Korea. Die Leistungsstreuung ist insgesamt deutlich höher; offensichtlich scheint es im Vereinigten Königreich eine relativ umfangreiche Spitzengruppe mit ausgezeichneter naturwissenschaftlicher Kompetenz zu geben.⁶⁰

⁵⁸ Prenzel, Rost, Senkbeil, Häußler, Klopp: Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 236

⁵⁹ Klieme, Neubrand, Lüdtke: Mathematische Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 176

⁶⁰ Prenzel, Rost, Senkbeil, Häußler, Klopp: Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 236ff

Der überwiegende Teil englischer Schülerinnen und Schüler befindet sich bereits in der 11. Klassenstufe, jedoch verteilt sich die Altersgruppe stärker als in Japan und Korea.

Die von Stanat und Kunter in PISA berichtete Gleichheit unter den Geschlechtern an Koreas Schulen ist ein weiterer Aspekt.

*Relativ selten kommt es allerdings in der Tat vor, dass 15-Jährige in einem Land sehr gute Ergebnisse im Lesetest erzielen und sich die Leistungen von Mädchen und Jungen kaum unterscheiden. Dies ist lediglich in Korea der Fall (...).*⁶¹

In koreanischen Schulen finden sich den Daten der PISA-Studie zufolge nur noch rund 85% der Altersgruppe der 15-Jährigen, das bedeutet 15% der Schülerinnen und Schüler haben bereits die Schule verlassen. Die frühen Schulabgänger sind zumeist Schülerinnen.⁶² Es entsteht in der verbliebenen Population ein Überhang von ca. 12 Prozentpunkten zugunsten der männlichen Schüler. Die verbliebenen Schülerinnen stellen eine Auswahl dar.

Die relativ hohen Testwerte Japans und Koreas bei geringer Streuung sind nicht nur im Falle Koreas mit einer im Verhältnis der Teilnehmerstaaten geringen Rückhaltequote des Schulsystems verbunden. Parallel neben dem öffentlichen Schulsystem finden sich in Japan und Korea private Schulen (in Japan: JUKU), die in Japan von knapp 60% und in Korea von etwa 36% der Schülerinnen und Schüler (OECD-Durchschnitt 9,4 incl. Japan und Korea) besucht werden.⁶³

Die Situation in Koreas und Japans Schulen, die im PISA-Test berücksichtigt wurden, stellt sich nach Datenlage der Studie wie folgt dar: Die öffentlichen Schulen weisen durchschnittliche Klassengrößen von 38,8 (Japan) und 37,6 (Korea) Schülern auf. Diese Größenordnung entspricht den Verhältnissen der Länder Brasilien (38,8) und Mexiko (35,0). Der berichtete Durchschnitt liegt hier inklusive Japan und Korea bei 24,6. Europäische Länder weisen durchschnittliche Klassenstärken von lediglich 19,5 (Finnland), 18,2 (Schweiz) und 17,4 (Dänemark) auf.⁶⁴ Präzise soziodemographische

⁶¹ Stanat, Kunter: Geschlechterunterschiede in Basiskompetenzen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 251

⁶² vgl. Abbildung *Verteilung Geschlecht*

⁶³ 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. 417

⁶⁴ 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. 422

Angaben der koreanischen Schulabgänger konnten in der PISA-Studie nicht gefunden werden.

Nach Datenlage wird über Schulprobleme in Japan und Korea besonders wenig in den Elternhäusern gesprochen. 10% der koreanischen und 14% der japanischen Schülerinnen und Schüler geben an, nie mit ihren Eltern über Schulprobleme zu sprechen (Länderdurchschnitt: 4,1%, Finnland 1,6% und Deutschland 1,7%). Rund 43% der koreanischen Schüler geben an, die Lehrer würden sie nie in ihrer Arbeit unterstützen.⁶⁵ Der Länderdurchschnitt beträgt hier 8,6%, Deutschland 13,2%, Schweiz 6,2%, Schweden 3,4%, Finnland 3,1% und Dänemark 2%.⁶⁶

Jeweils etwa zwei Drittel der koreanischen und japanischen Lehrer beanstanden mangelhafte Bauqualität der Schulgebäude,⁶⁷ mangelnde Heizung⁶⁸ und ungenügenden Raum⁶⁹ ihrer Schulen. Zwei Drittel der koreanischen Lehrer beanstanden darüber hinaus einen Mangel an Computern,⁷⁰ eine schlecht ausgestattete Bibliothek⁷¹ und schlechte Ausstattung mit wissenschaftlichem Material.⁷²

Die folgende Abbildung zeigt den statistischen Zusammenhang zwischen Ergänzungsunterricht (supplementary), Nachhilfeunterricht (tuition/tutoring) und Klassengröße (class size). Es lässt sich mit $r_p=0,658$ ($p<0,01$) und $r_s=0,551$ ($p<0,01$) ein statistisch signifikanter Zusammenhang der Klassengröße in Bezug auf Ergänzungsunterricht in Prozent der Nennungen und mit $r_p=0,417$ ($p<0,05$) und $r_s=0,482$ ($p<0,01$) ein statistischer Zusammenhang in Bezug auf Nachhilfeunterricht nachweisen. Die Abbildung legt einen kompensatorischen Zusammenhang nahe. Finnland, Schweden und die Schweiz weisen kleine Klassengrößen bei sehr geringem zusätzlichem Unterricht auf. Japan und Korea hingegen, gemeinsam mit Mexiko und Brasilien, weisen die zahlenmäßig umfangreichsten durchschnittlichen Klassengrößen auf, wobei Mexiko und Brasilien bei niedrigen Testleistungen einen geringen Anteil an zusätzlichem Unterricht, Japan und Korea jedoch bei hohen Testergebnissen einen hohen Anteil an privatem Zusatzunterricht aufweisen.

⁶⁵ Codierung: teachers help with work Stufe: never

⁶⁶ Berechnungen aus Rohdaten der Public-Use-Files

⁶⁷ Item sc11q01 (poor buildings) Zusammenfassung: a little, some, a lot

⁶⁸ Item sc11q02 (poor heating) Zusammenfassung: a little, some, a lot

⁶⁹ Item sc11q03 (inadequate space) Zusammenfassung: a little, some, a lot

⁷⁰ Item SC11Q05 (lack of computers) Zusammenfassung: a little, some, a lot

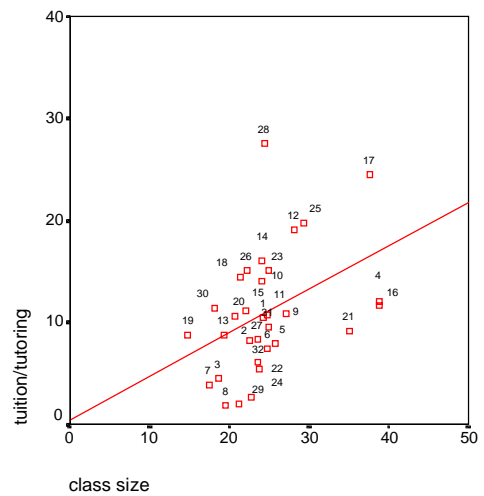
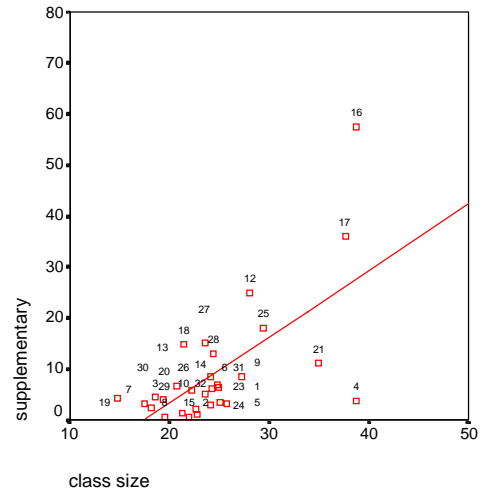
⁷¹ Item SC11Q06 (poor library) Zusammenfassung: a little, some, a lot

⁷² Item SC11Q08 (poor science equipment) Zusammenfassung: a little, some, a lot

country	case no	class size *)	supplementary *)	tuition / tutoring *)	HISEI**)
AUSTRALIA	1	24,30	6,10	10,50	52,25
AUSTRIA	2	22,60	2,20	8,20	49,72
BELGIUM	3	18,60	4,40	4,50	48,95
BRAZIL	4	38,80	3,70	12,10	43,93
CANADA	5	25,70	3,30	8,00	52,83
CZECH REPUBLIC	6	24,80	6,80	7,40	48,31
DENMARK	7	17,40	3,10	3,80	49,73
FINLAND	8	19,50	0,50	1,90	50,00
FRANCE	9	27,20	8,60	10,90	48,27
GERMANY	10	24,10	3,00	14,10	48,85
GREECE	11	24,80		10,70	47,76
HUNGARY	12	28,10	25,00	19,10	49,53
ICELAND	13	19,30	3,90	8,70	52,73
IRELAND	14	24,10	8,60	16,00	48,43
ITALY	15	22,00	0,60	11,10	47,08
JAPAN	16	38,80	57,50	11,60	50,54
KOREA, REPUBLIC OF	17	37,60	36,00	24,50	42,80
LATVIA	18	21,40	14,80	14,50	50,15
LIECHTENSTEIN	19	14,80	4,20	8,80	47,46
LUXEMBOURG	20	20,70	6,70	10,60	44,79
MEXICO	21	35,00	11,20	9,20	42,48
NETHERLANDS	22	23,80		5,40	50,85
NEW ZEALAND	23	25,00	3,50	15,10	52,20
NORWAY	24	22,70	1,10	2,70	53,91
POLAND	25	29,40	18,10	19,80	46,03
PORTUGAL	26	22,20	5,70	15,10	43,85
RUSSIAN FEDERATION	27	23,50	15,10	8,30	49,38
SPAIN	28	24,40	13,10	27,60	44,99
SWEDEN	29	21,20	1,30	2,00	50,57
SWITZERLAND	30	18,20	2,40	11,40	49,21
UNITED KINGDOM	31	24,90	6,30	9,50	51,26
UNITED STATES	32	23,50	5,10	6,10	52,40

*) Reference: PISA 2000, P.417, 422

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



Korrelationen

		supplementary	tuition/tutoring
class size	Korrelation nach Pearson	,658**	,417*
	Signifikanz (2-seitig)	,000	,018
	N	30	32

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

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

Korrelationen

		supplementary	tuition/tutoring
Spearman-Rho	class size	Korrelationskoeffizient	,551**
		Sig. (2-seitig)	,002
		N	30

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

Abbildung 29 Klassengröße – Extra Unterricht

Mit $r_p = -0,512$ ($p < 0,01$) $r_s = -0,498$ ($p < 0,01$) zeigt sich ein negativer statistischer Zusammenhang von durchschnittlichem sozioökonomischem Index der Schüler eines Landes und prozentualem Anteil privater Nachhilfe.

		tuition / tutoring *)
HISEI**)	Korrelation nach Pearson	-,512**
	Signifikanz (2-seitig)	,003
	N	32

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

			tuition / tutoring *)
Spearman-Rho	HISEI**)	Korrelationskoeffizient	-,498**
		Sig. (2-seitig)	,004
		N	32

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

Abbildung 30 Lerngelegenheiten – Soz.Ökon. Index/Korrelation

Je geringer der durchschnittliche sozioökonomische Index der 15-jährigen Schülerpopulation eines Landes ist, desto mehr außerschulische Lerngelegenheiten werden in Anspruch genommen.

Geringer sozioökonomischer Index kovariiert positiv mit hohen Klassenfrequenzen. Japan und Korea bilden mit einem hohen Anteil außerschulischer Lerngelegenheiten hier die Ausnahme. Finnland, Schweden, Dänemark und weitere westeuropäische Teilnehmerstaaten hingegen weisen sehr geringe Anteile außerschulischer Lerngelegenheiten auf. In Finnland berichten lediglich 1,9% der Schülerinnen und Schüler die Inanspruchnahme von extra Nachhilfeunterricht und nur 0,5% von weiterem Ergänzungsunterricht.

2.5 Soziale Disparität in der Lesekompetenz

Baumert und Schümer berichten *Disparitäten* im Erwerb von Lesekompetenz.⁷³ Für jedes Land wurden die Differenzen der Mittelwerte von zwei Schülergruppen gebildet: die mittleren Testleistungen der 25% Jugendlichen aus Familien mit dem höchsten Sozialstatus innerhalb des Landes und die mittleren Testleistungen der 25% Jugendlichen aus Familien mit dem niedrigsten Sozialstatus. Die Werte in der folgenden Abbildung zeigen, welchen Testleistungsvorsprung die erste Gruppe gegenüber der zweiten aufweist.

⁷³ Baumert, Schümer: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 383

*In Deutschland, gefolgt von den beiden ehemaligen sozialistischen Staaten Tschechische Republik und Ungarn sowie der Schweiz, Luxemburg und Belgien, können am Ende der Sekundarstufe I soziale Disparitäten der Lesekompetenz nachgewiesen werden, wie sie in anderen Ländern – allen voran Japan, Korea, Island oder Finnland - kaum denkbar sind.*⁷⁴

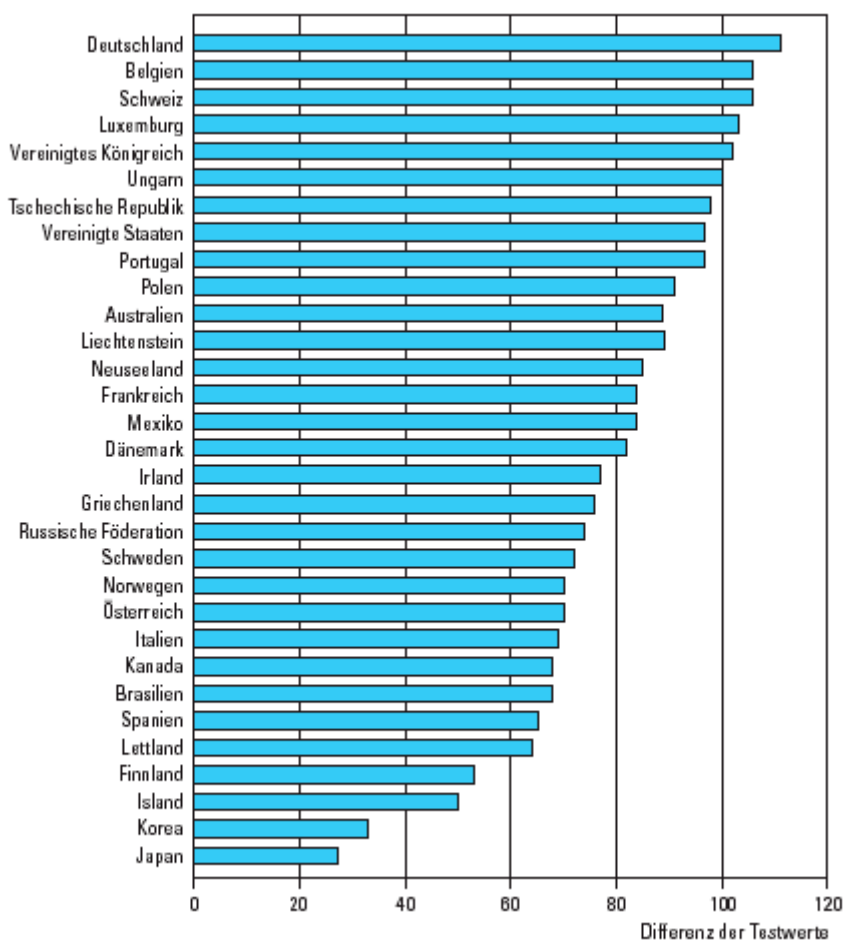


Abbildung 31 Disparität Quelle: PISA 2000

Die oben gezeigte Abbildung stellt die Teilnehmerstaaten nach der Differenz zwischen der mittleren Lesekompetenz von 15-Jährigen aus Familien des oberen und unteren Viertels der Sozialstruktur dar.

Der Befund wird für viele, auch mit der Disparitätsforschung vertraute Leser überraschend sein: Deutschland und die Schweiz sind die beiden Länder mit den größten Unterschieden in der Lesekompetenz von Jugendlichen aus höheren und

⁷⁴ Baumert, Schümer: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 389

*niedrigeren Sozialschichten. Die Differenz beträgt in Deutschland mehr als eineinhalb Kompetenzstufen oder 1,2 Standardabweichungen. Selbst die Vereinigten Staaten, die immer wieder als Beispiel für große soziale Disparitäten in den Bildungschancen angeführt werden, weisen zwar immer noch beträchtliche, aber signifikant niedrigere sozial bedingte Leistungsunterschiede auf. (...) In Korea sinken die Disparitätswerte auf 36 Punkte und in Japan auf 27 Punkte, wohlgermerkt bei gleich großer sozialer Heterogenität der Elternhäuser (...).*⁷⁵

Auffallend sind nicht nur die hohen sozialen Disparitäten in der Schweiz, sondern auch die moderaten Disparitäten in der Russischen Föderation, Mexiko, insbesondere aber die außerordentlich geringen Disparitäten Brasiliens.

2.5.1 Streuung des Schulalters

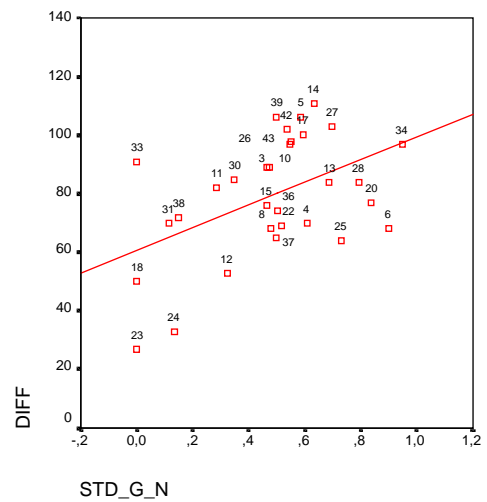
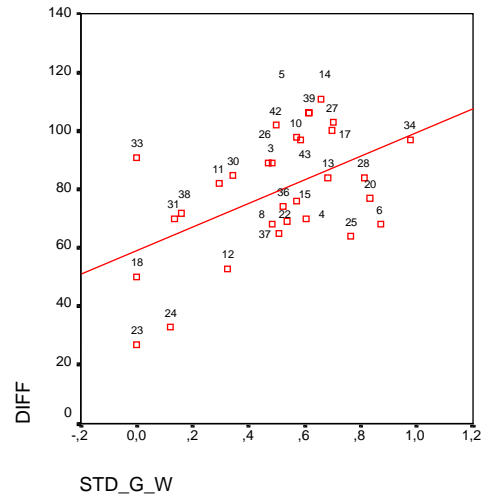
Die am Lebensalter orientierte Definition der PISA-Studie hat eine Variabilität des Schulalters und wie bereits gezeigt wurde einen Einfluss auf die Testwerte zur Folge. Es ist nun zu prüfen, in welchem Umfang diese Variabilität einen Einfluss auf die *Disparität* im Lesen (wie oben dargestellt) ausübt. Mit anderen Worten, es ist zu prüfen, ob die hohen *Disparitäten* im Lesen auf die Stichprobendefinition zurückzuführen sind. Geprüft wird der Zusammenhang von *Disparität* entsprechend den vom MPIB veröffentlichten Daten (hier als DIFF codiert)⁷⁶ und der Standardabweichung der Länderstichproben jeweils über die gewichtete (STD_G_W) und ungewichtete Variable Klassenstufe (STD_G_U).

⁷⁵ Baumert, Schümer: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 383

⁷⁶ Baumert, Schümer: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 384 ff

COUNTRY	CASE No	STD_G_w	STD_G_n	DIFF
ALBANIA	1	0,62	0,52	
ARGENTINA	2	0,74	0,69	
AUSTRALIA	3	0,48	0,47	89
AUSTRIA	4	0,60	0,61	70
BELGIUM	5	0,62	0,59	106
BRAZIL	6	0,87	0,90	68
BULGARIA	7	0,47	0,43	
CANADA	8	0,48	0,48	68
CHILE	9	0,76	0,70	
CZECH REPUBLIC	10	0,57	0,55	98
DENMARK	11	0,29	0,28	82
FINLAND	12	0,32	0,32	53
FRANCE	13	0,68	0,69	84
GERMANY	14	0,66	0,63	111
GREECE	15	0,57	0,46	76
HONG KONG	16	0,90	0,89	
HUNGARY	17	0,70	0,59	100
ICELAND	18	0,00	0,00	50
INDONESIA	19	0,93	0,91	
IRELAND	20	0,83	0,84	77
ISRAEL	21	0,33	0,36	
ITALY	22	0,54	0,52	69
JAPAN	23	0,00	0,00	27
KOREA, REPUBLIC OF	24	0,12	0,13	33
LATVIA	25	0,76	0,73	64
LIECHTENSTEIN	26	0,47	0,47	89
LUXEMBOURG	27	0,70	0,70	103
MEXICO	28	0,81	0,79	84
NETHERLANDS	29	0,62	0,60	
NEW ZEALAND	30	0,34	0,35	85
NORWAY	31	0,13	0,11	70
PERU	32	1,09	1,06	
POLAND	33	0,00	0,00	91
PORTUGAL	34	0,98	0,95	97
ROMANIA	35	0,60	0,47	
RUSSIAN FEDERATION	36	0,52	0,50	74
SPAIN	37	0,51	0,50	65
SWEDEN	38	0,16	0,15	72
SWITZERLAND	39	0,62	0,50	106
THAILAND	40	0,56	0,54	
MACEDONIA	41	0,52	0,48	
UNITED KINGDOM	42	0,50	0,53	102
UNITED STATES	43	0,59	0,55	97

STD_G_w standard deviation GRADE -weighted by W_FSTUWT
 STD_G_n standard deviation GRADE -unweighted
 DIFF "Disparität" (disparity)
 PISA 2000 database : http://pisaweb.acer.edu.au/oeecd/oeecd_pisa_data_s1.html
 data: intstud_read.txt



Korrelationen

	STD_G_W	STD_G_N
DIFF Korrelation nach Pearson	,514**	,483**
Signifikanz (2-seitig)	,003	,006
N	31	31

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

Korrelationen

	STD_G_W	STD_G_N
Spearman-Rho DIFF Korrelationskoeffizient	,390*	,379*
Sig. (2-seitig)	,030	,035
N	31	31

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 32 Disparität im Lesen – Streuung des Schulalters

Mit $r_p = 0,514$ ($p < 0,01$), $r_s = 0,39$ ($p < 0,05$) (gewichteter Datensatz) und $r_p = 0,483$ ($p < 0,01$), $r_s = 0,379$ ($p < 0,05$) lässt sich ein signifikanter Zusammenhang zwischen den Disparitäten und der jeweiligen Streuung über die Schulalteranteile nachweisen.

Dieses Ergebnis zeigt einen signifikanten Einfluss der Stichprobendefinition.

Durch die am Lebensalter orientierte Stichprobendefinition variiert das Schulalter von der 4. bis zur 13. Klassenstufe. Es soll im zweiten Schritt geprüft werden, ob ein Testkonstrukt vorliegt, das auf diese hohe Variabilität abgestimmt ist.

2.5.2 Bodeneffekte/Deckeneffekte

So genannte Decken- beziehungsweise Bodeneffekte⁷⁷ treten bei ungeeignetem Verhältnis von Schwierigkeit und Leistungsniveau einer untersuchten Population auf, insbesondere bei großer Leistungsheterogenität. Die folgenden Abbildungen zeigen die Messung von Differenzen der Mittelwerte (entsprechend der Messung von so genannten *Disparitäten*) in Abhängigkeit der Schiefe einer Verteilung.

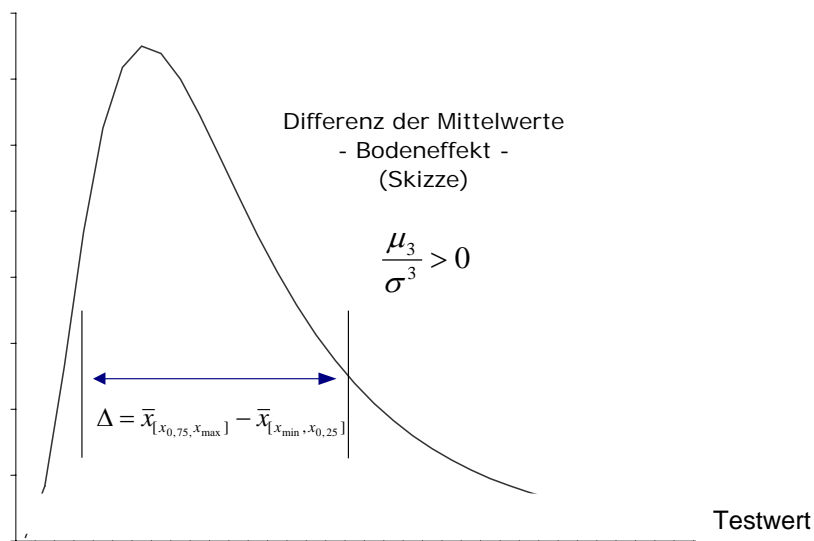


Abbildung 33 *Disparität* - schiefe Testwertverteilung (Bodeneffekt)

⁷⁷ Rost, J.: Testtheorie Testkonstruktion 1996, S.96

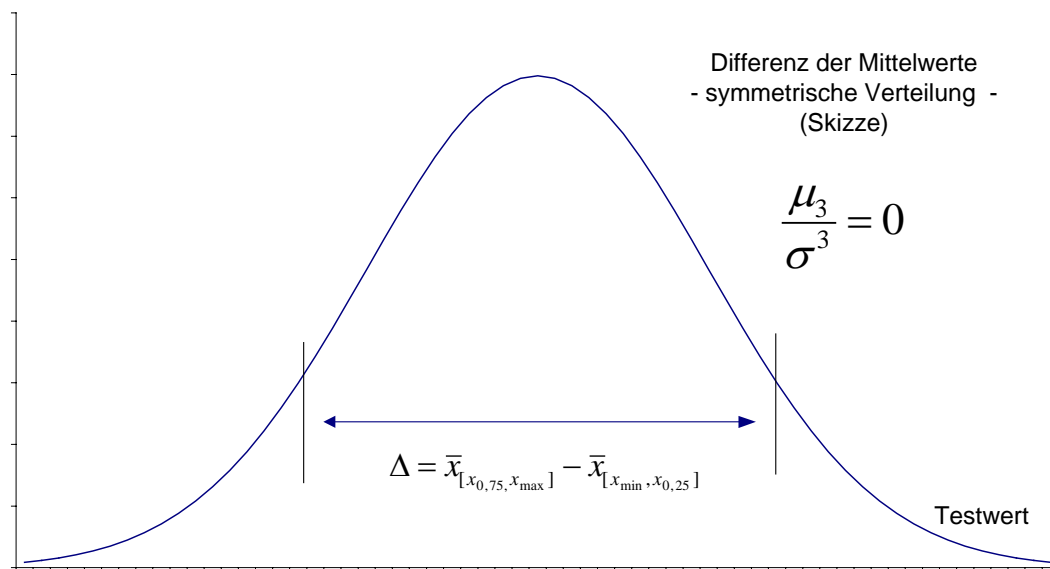


Abbildung 34 *Disparität* - symmetrische Testwertverteilung

Im Vergleich einer Verteilung mit $\frac{\mu_3}{\sigma^3} > 0$ (Bodeneffekt) gegen eine Verteilung mit $\frac{\mu_3}{\sigma^3} = 0$ (symmetrische Verteilung) wird die Ursache geringerer Differenz ($\Delta = \bar{x}_{[x_{0,75}, x_{max}]} - \bar{x}_{[x_{min}, x_{0,25}]}$) bei schiefer Verteilung deutlich. Da die Verteilung bei Vorhandensein eines Bodeneffektes eine untere Schranke aufweist, verringern sich ebenso die über die Mittelwerte $\bar{x}_{[x_{0,75}, x_{max}]}$ und $\bar{x}_{[x_{min}, x_{0,25}]}$ gemessenen Differenzen. Gleiches gilt für so genannte Deckeneffekte (obere Schranke), die hier jedoch nicht extra dargestellt werden. Decken- und Bodeneffekte können als weitere Ursache für die hohen statistischen *Disparitäten* nachgewiesen werden. Ein Deckeneffekt läge vor, wenn bei steigendem Mittelwert der Daten bis zum 1. Quartil, also das Intervall $[x_{min}, x_{0,25}]$ (hier QUANT 1 genannt), die Differenz zwischen diesem Mittelwert und dem Mittelwert des Bereichs ab dem 3. Quartil, also das Intervall $[x_{0,75}, x_{max}]$, systematisch kleiner wird sowie parallel dazu die Schiefe der Verteilung zunimmt. Die Schiefe der Verteilung wäre zunehmend linksschief (rechtssteil) ($\mu_3/\sigma^3 < 0$).⁷⁸ Für Bodeneffekte gilt entsprechend der umgekehrte Zusammenhang (siehe vorangegangene Abbildung mit $\mu_3/\sigma^3 > 0$).

⁷⁸ Notation nach Büning: Nichtparametrische Statistische Methoden 1994. S. 304

Insbesondere im Split Klassenstufe (grade) 10 lässt sich ein solcher Effekt nachweisen. Bei steigendem unterem Mittelwert im Intervall $[x_{\min}, x_{0,25}]$ (Quant1) steigt ebenfalls der mittlere Testwert eines Landes (PV1READ), jedoch bei sinkender Differenz zwischen den Mittelwerten aus den Bereichen $[x_{0,75}, x_{\max}]$ und $[x_{\min}, x_{0,25}]$ hier als DIFF bezeichnet.

Korrelationen

		QUANT1	PV1READ	DIFF
QUANT1	Korrelation nach Pearson	1	,955*	-,429*
	Signifikanz (2-seitig)	.	,000	,029
	N	26	26	26
PV1READ	Korrelation nach Pearson	,955*	1	-,164
	Signifikanz (2-seitig)	,000	.	,425
	N	26	26	26
DIFF	Korrelation nach Pearson	-,429*	-,164	1
	Signifikanz (2-seitig)	,029	,425	.
	N	26	26	26

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

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

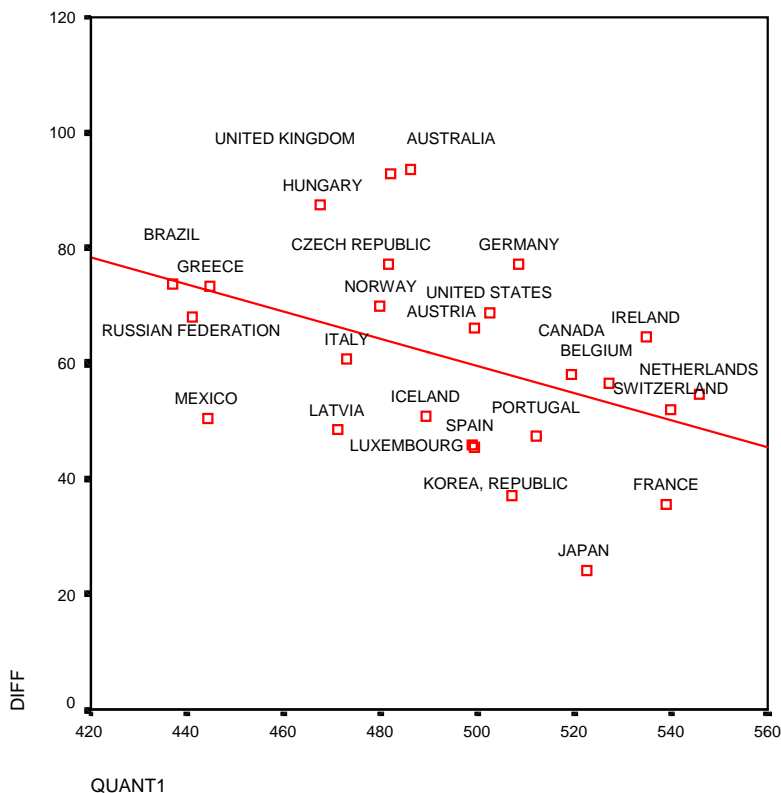


Abbildung 35 Deckeneffekt/Bodeneffekt (Abbildung I)

Der lineare Zusammenhang lässt sich mit $r_p = -0,429$ $p=0,029$ beziehungsweise $r_p = 0,955$ $p < 0,01$ nachweisen.

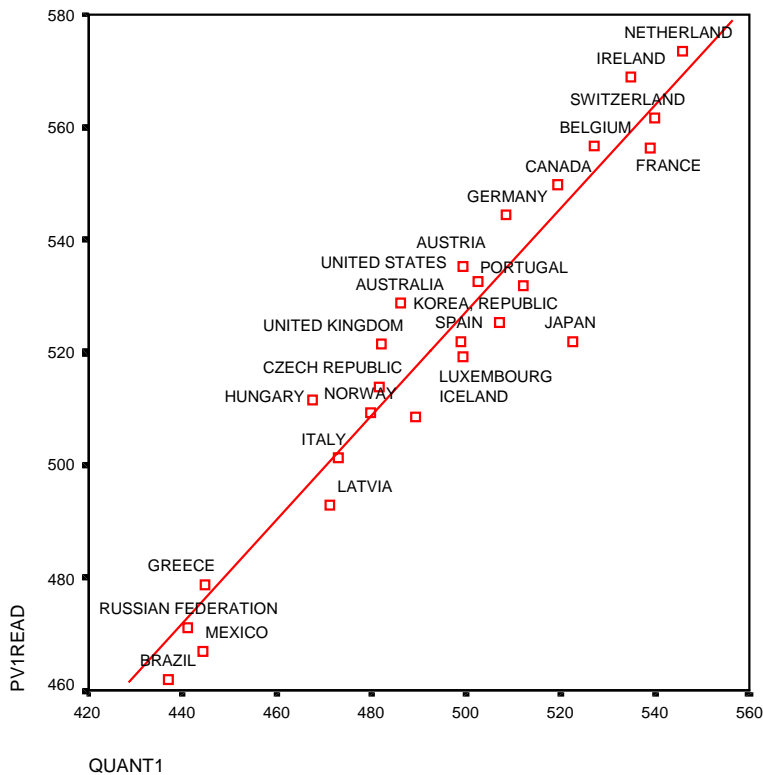


Abbildung 36 Deckeneffekt/Bodeneffekt (Abbildung II)

Für die 11. Klassenstufe lassen sich aufgrund der geringen Fallzahl keine Berechnungen sinnvoll durchführen. Die Klassenstufe 9 hingegen weist diesen Effekt nicht mehr auf, wie folgende Abbildungen zeigen. Es besteht, wie zu erwarten, ein statistisch signifikanter Zusammenhang zwischen dem unteren Mittelwert der Werte im Intervall $[x_{\min}, x_{0,25}]$ und dem mittlerem Testwert eines Landes ($r_p = 0,931$ $p < 0,01$), jedoch kein statistisch signifikanter Zusammenhang der Differenz der Mittelwerte aus den Bereichen $[x_{\min}, x_{0,25}]$ und $[x_{0,75}, x_{\max}]$.

Korrelationen

		DIF_9	QUANT1	PVREAD
DIF_9	Korrelation nach Pearson	1	-,022	,293
	Signifikanz (2-seitig)	.	,919	,155
	N	25	25	25
QUANT1	Korrelation nach Pearson	-,022	1	,931**
	Signifikanz (2-seitig)	,919	.	,000
	N	25	25	25
PVREAD	Korrelation nach Pearson	,293	,931**	1
	Signifikanz (2-seitig)	,155	,000	.
	N	25	25	25

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

Abbildung 37 Disparität/Korrelation

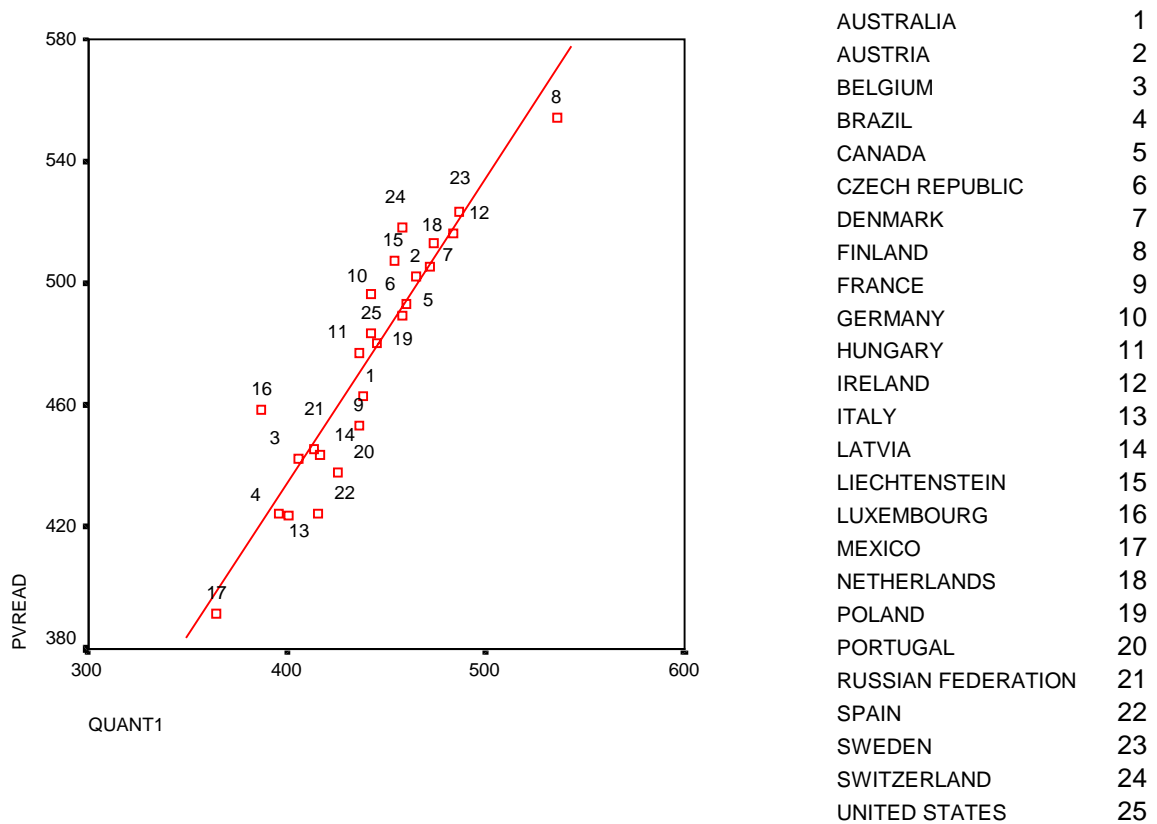


Abbildung 38 Disparität Lesetestwert Klassenstufe (GRADE) 9

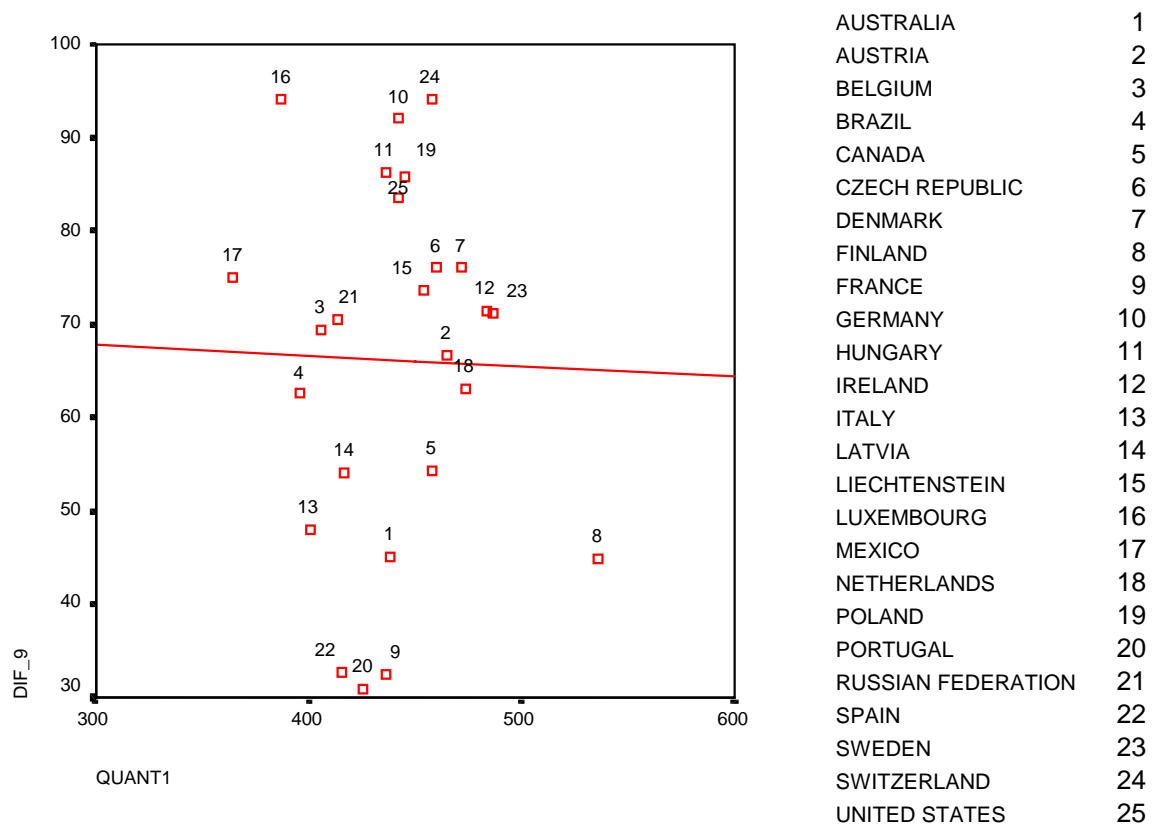
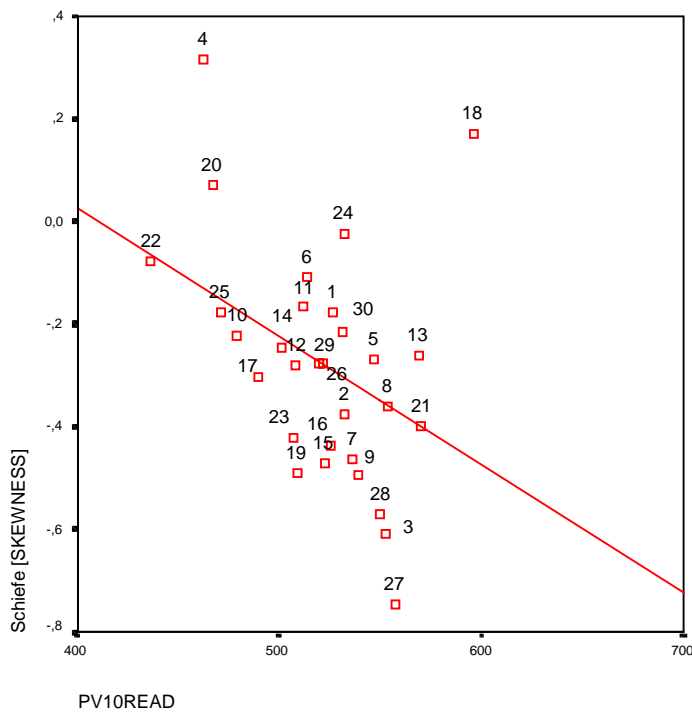


Abbildung 39 Disparität Differenz/Korrelation Klassenstufe (GRADE) 9

Prinzipiell könnte der Zusammenhang von mittlerem Testwert und Schiefe der Verteilung graphisch abgetragen werden.



- AUSTRALIA 1
- AUSTRIA 2
- BELGIUM 3
- BRAZIL 4
- CANADA 5
- CZECH REPUBLIC 6
- DENMARK 7
- FRANCE 8
- GERMANY 9
- GREECE 10
- HUNGARY 11
- ICELAND 12
- IRELAND 13
- ITALY 14
- JAPAN 15
- KOREA, REPUBLIC OF 16
- LATVIA 17
- LIECHTENSTEIN 18
- LUXEMBOURG 19
- MEXICO 20
- NETHERLANDS 21
- NEW ZEALAND 22
- NORWAY 23
- PORTUGAL 24
- RUSSIAN FEDERATION 25
- SPAIN 26
- SWEDEN 27
- SWITZERLAND 28
- UNITED KINGDOM 29
- UNITED STATES 30

Korrelationen

		SKEWNESS
PV10READ	Korrelation nach Pearson	-,386*
	Signifikanz (2-seitig)	,035
	N	30

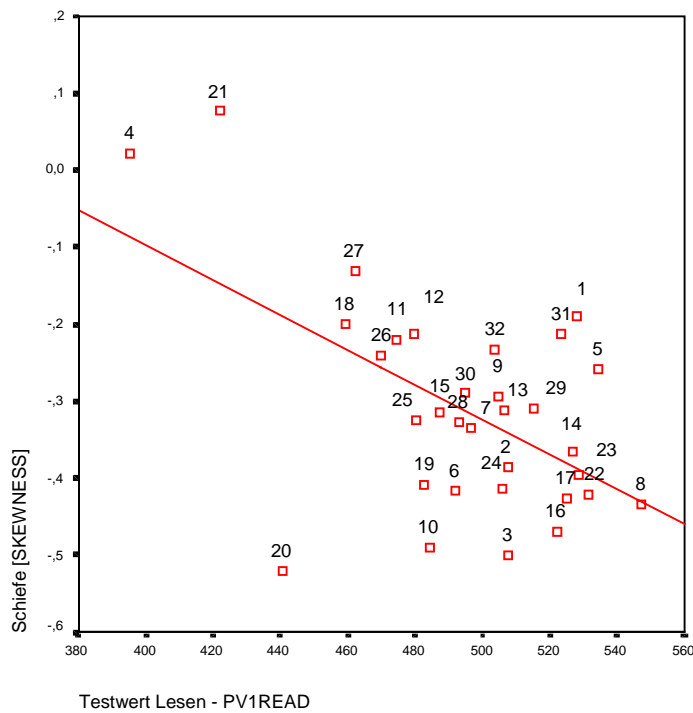
*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Korrelationen

			SKEWNESS
Spearman-Rho	PV10READ	Korrelationskoeffizient	-,422*
		Sig. (2-seitig)	,020
		N	30

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 40 Testwert – Schiefe/Korrelation (Klassenstufe 10)



- AUSTRALIA 1
- AUSTRIA 2
- BELGIUM 3
- BRAZIL 4
- CANADA 5
- CZECH REPUBLIC 6
- DENMARK 7
- FINLAND 8
- FRANCE 9
- GERMANY 10
- GREECE 11
- HUNGARY 12
- ICELAND 13
- IRELAND 14
- ITALY 15
- JAPAN 16
- KOREA, REPUBLIC OF 17
- LATVIA 18
- LIECHTENSTEIN 19
- LUXEMBOURG 20
- MEXICO 21
- NETHERLANDS 22
- NEW ZEALAND 23
- NORWAY 24
- POLAND 25
- PORTUGAL 26
- RUSSIAN FEDERATION 27
- SPAIN 28
- SWEDEN 29
- SWITZERLAND 30
- UNITED KINGDOM 31
- UNITED STATES 32

Korrelationen

		SKEWNESS
PV1READ	Korrelation nach Pearson	-,551**
	Signifikanz (2-seitig)	,001
	N	32

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

Korrelationen

			SKEWNESS
Spearman-Rho	PV1READ	Korrelationskoeffizient	-,371*
		Sig. (2-seitig)	,037
		N	32

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 41 Lesetestwert (sämtliche Klassenstufen) Schiefe der Verteilung

Mit $r_p = -0,551$ ($p < 0,01$) und $r_s = -0,371$ ($p < 0,05$) zeigt sich die Problematik einer geeigneten Testkonstruktion angesichts einer außerordentlich heterogenen Leistungsfähigkeit der untersuchten Schüler. Als Konsequenz der am Lebensalter orientierten Stichprobendefinition umfasst die Stichprobe Schülerinnen und Schüler von der vierten bis zur dreizehnten Klassenstufe. Diese werden in einem gemeinsamen

Testverfahren bewertet.⁷⁹ Unter diesen Umständen werden Decken- und Bodeneffekte (bei so genannten J-förmigen Verteilungen) kaum vermeidbar sein.⁸⁰ Lienert und Raatz sprechen in diesem Zusammenhang von einer *anormalen Verteilung infolge mangelhafter Testkonstruktion*.⁸¹ Korrespondierende Daten liefern auch die Verteilungen der PISA Teilnehmerstaaten.

2.6 Migration/durchschnittliches Schulalter

Nachdem gezeigt wurde, dass das Schulalter (Klassenstufe) im Datensatz einen statistisch signifikanten Einfluss auf die Testleistung ausübt, ist nun die Frage von Interesse, ob ein unterschiedliches Schulalter der einzelnen Segmente der Schülerpopulationen eines Landes nachweisbar ist.

Eine Reihe europäischer Staaten weist umfangreiche Migrationsanteile in ihrer Schülerpopulation auf. Der Anteil variiert je nach Definition. Die folgend gebrauchte Definition richtet sich nicht nach der Verkehrssprache in der Familie, sondern nach dem Geburtsort des Probanden (Codierung im Datensatz: *self*), der Mutter (Codierung im Datensatz: *mother*) und des Vaters (Codierung im Datensatz: *father*). Es ergeben sich insgesamt 8 verschiedene Kombinationen.⁸²

Definiert man Schülerinnen und Schüler mit Migrationshintergrund als diejenigen, deren beide Eltern außerhalb des Ziellandes geboren wurden, unabhängig vom Geburtsland des einzelnen Schülers, was der Kombination 4 und 8 entspricht (siehe Fußnote), finden sich in Luxemburg etwa 30%, Australien, Kanada, Schweiz und Neuseeland ca. 20% und in Deutschland, den USA, Frankreich, Schweden und Österreich zwischen 10% und 15% Schüler mit Migrationshintergrund im Datensatz der PISA-Testteilnehmer.

Praktisch keine Schülerinnen und Schüler mit Migrationshintergrund nach dieser Definition weisen die Datensätze der Länder Finnland, Tschechien, Island, Brasilien, Polen, Japan und Korea auf.

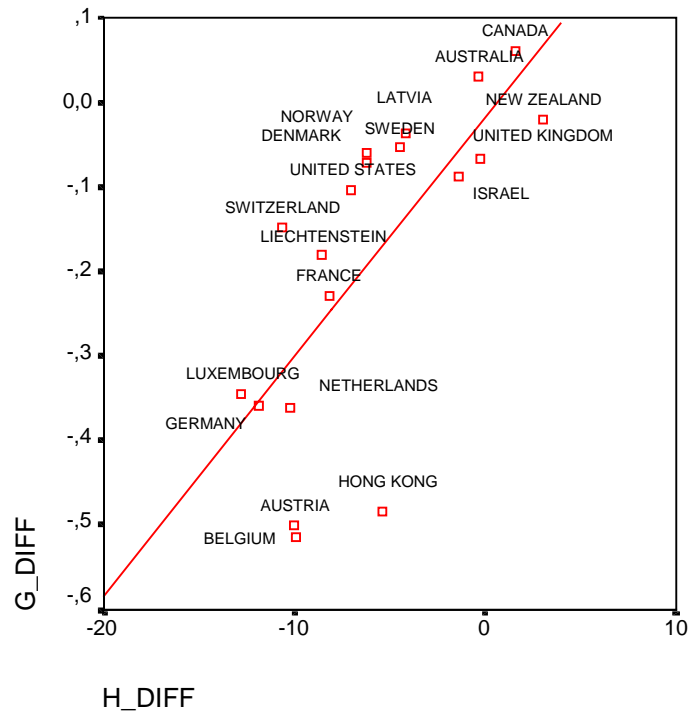
⁷⁹ Der Test umfasste zwei 60-minütige Sitzungen für die Bereiche Reading, Mathematical und Scientific Literacy (PISA-Manual. S. 9)

⁸⁰ Merkens hatte so genannte J-förmige Verteilungen bei lernzielorientierten Tests angesprochen. Merkens: Einführung in die Statistik für Pädagogen 1975. S. 68 ff

⁸¹ Lienert, Raatz: Testaufbau und Testanalyse, 1998, S. 155 ff

⁸² Die Gruppenbildung weist $2^3 = 8$ Kombinationen auf: Geburtsort (2 Ausprägungen: Inland (I)/Ausland (A)) sowie Person (3 Ausprägungen: Proband/Mutter/Vater). Folgende Gruppen werden definiert: **1** 'self (I) - mother (I) - father (I)' **2** 'self (I) - mother (I) - father (A)' **3** 'self (I) - mother (A) - father (I)' **4** 'self (I) - mother (A) - father (A)' **5** 'self (A) - mother (I) - father (I)' **6** 'self (A) - mother (I) - father (A)' **7** 'self (A) - mother (A) - father (I)' **8** 'self (A) - mother (A) - father (A)' (I: Inland, A: Ausland). Deckungsanteil: 94,1% (internationaler Datensatz) 94,3% (nationaler Datensatz) an gesamter Stichprobe.

Migration lässt sich nach sozioökonomischer Differenz zum Zielland unterscheiden. Trägt man die sozioökonomische Differenz der beiden Schülergruppen (Codierung: H_DIFF)⁸³ gegen die jeweilige Schulalterdifferenz (Codierung: G_DIFF)⁸⁴ je Teilnehmerstaat für diejenigen Länder ab, die mit 5% Anteil einen nennenswerten Migrationsanteil aufweisen, zeigt sich über die aggregierten Werte ein hochgradiger linearer Zusammenhang der Differenzen (Berechnung über den erweiterten Datensatz).



Korrelationen

		G_DIFF
H_DIFF	Korrelation nach Pearson	,715**
	Signifikanz (2-seitig)	,001
	N	19

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

Korrelationen

		G_DIFF
Spearman-Rho	H_DIFF Korrelationskoeffizient	,781**
	Sign. (2-seitig)	,000
	N	19

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

Abbildung 42 Sozioökonomische Differenz – Schulalter Differenz/Korrelation

Zu erkennen ist der mit $r_p = 0,715$ beziehungsweise $r_s = 0,781$ ⁸⁵ hochgradige statistische Zusammenhang. Je höher also die sozioökonomische Differenz der Schülerinnen und

⁸³ Codierung orientiert sich am Anfangsbuchstaben des Variablenlabels (**H**ISEI), $H_diff = \bar{x}_{HISEI_{mitMig}} - \bar{x}_{HISEI_{ohneMig}}$

⁸⁴ Codierung orientiert sich am Anfangsbuchstaben des Variablenlabels (**G**RADE), $G_diff = \bar{x}_{GRADE_{mitMig}} - \bar{x}_{GRADE_{ohneMig}}$

⁸⁵ Berechnung des Koeffizienten berücksichtigt hier nicht die variierenden N_i der einzelnen Länder

Schüler mit Migrationshintergrund zu den Schülern des Ziellandes, desto höher deren Rückstand in Bezug auf das Schulalter.

Auffällig ist weiterhin, dass Schweden und Norwegen, die ebenfalls Aufnahmeländer für Flüchtlinge aus Krisenregionen sind und ebenfalls eine mehr oder weniger offene Arbeitszuwanderung zu verzeichnen haben, offensichtlich zu einer besseren Balance zwischen Umgangssprache in der Familie und Beherrschung der Verkehrssprache gelangen. (...) Auffallend ist die insgesamt günstigere Situation in Norwegen und Schweden. Auch wenn die Zuwandererfamilien an ihrer Herkunftssprache festhalten, gelingt es ihnen, einen höheren Sozialstatus als in Deutschland zu erreichen.⁸⁶

Schweden weist geringere Differenzen zwischen dem sozioökonomischen Niveau der Schüler mit und ohne Migrationshintergrund auf als beispielsweise die Schweiz, Österreich, Luxemburg oder die Bundesrepublik Deutschland.

Der vorliegende Befund steht ebenfalls im Gegensatz zu den berichteten Ergebnissen von Schnabel und Schwippert. Die Autoren hatten in Bezug auf das Alter der Schülerinnen und Schüler mit und ohne Migrationshintergrund in Deutschland keinen statistisch signifikant nachweisbaren Unterschied zwischen beiden Gruppen feststellen können.

Wie Tabelle VI.6 zeigt, lassen sich Vermutungen über ein höheres Durchschnittsalter der Schüler aus Familien nicht-deutscher Herkunft nicht belegen. Der Altersabstand zwischen den Jugendlichen aus deutschsprachigen Familien zu denjenigen aus partiell deutschsprachigen Familien ist mit durchschnittlich rund zwei Monaten nicht zufallskritisch abzusichern. Auch der Unterschied zwischen den Jugendlichen deutschsprachiger und nicht-deutschsprachiger Familien ist mit rund einem halben Jahr zwar etwas deutlicher ausgeprägt, aber ebenfalls statistisch nicht signifikant.⁸⁷

Nun wäre die Frage nach der Definition des Begriffes Migration zu stellen, die sich aus der Verkehrssprache ableitet sowie starker oder schwacher Abgrenzung von Vergleichspopulationen (deutschsprachig – partiell deutschsprachig). Besondere

⁸⁶ Baumert, Schümer: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 394 ff

⁸⁷ Schnabel, Schwippert: Einflüsse sozialer und ethnischer Herkunft beim Übergang in die Sekundarstufe II und den Beruf. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 288

Aufmerksamkeit sollte das eingesetzte statistische Verfahren und -angesichts der berichteten Unterschiede von im Durchschnitt eines halben Jahres- der Zusammenhang von Signifikanz, Effektgröße und Stichprobenumfang erfahren.

2.6.1 Chi-Quadrat-Test/Klassenstufe

Die Rohdaten der PISA-Studie bieten nun die Möglichkeit, Unterschiede in Bezug auf das Lebens- und Schulalter zu prüfen. Die PISA-Schülerpopulation bestand zwar lediglich aus der Altersgruppe der 15-Jährigen, jedoch wurde die Klassenstufe variiert. Es ist im Umkehrschluss anzunehmen, dass diejenigen Schülerinnen und Schüler, die im Alter von 15 Jahren niedrigere Klassenstufen besuchen, in ihren Klassen zu den älteren Schülern gehören. Es ist also der Nachweis von Interesse, ob Schüler mit Migrationshintergrund ein signifikant geringeres Schulalter aufweisen.

Der Chi-Quadrat-Unabhängigkeitstest ist angezeigt. Migration wird anhand des Geburtslandes der Eltern unabhängig von der Verkehrssprache in der Familie wie eingangs definiert. Modell: Die Ausgangshypothese H_0 geht von der Unabhängigkeit der beiden Zufallsvariablen *Migrationshintergrund* und *Klassenstufe* aus.

$$H_0 : \pi_{ij} = \pi_{i.} \pi_{.j} \text{ für alle } i = 1, \dots, k \text{ und alle } j = 1, \dots, l$$

$$H_1 : \pi_{ij} \neq \pi_{i.} \pi_{.j} \text{ für mindestens ein Paar } (i,j).$$

Getestet wird über die Prüfgröße

$$\chi^2_{emp} : \sum_{i=1}^k \sum_{j=1}^l \frac{(n_{ij} - \tilde{n}_{ij})^2}{\tilde{n}_{ij}} .$$

Die Frage nach der unter Gültigkeit von H_0 erwarteten Zellbesetzung prüfen die *standardisierten Residuen* mit $u_{ij} \sim N(0;1)$ nach folgendem Modell:

$$u_{ij} = \frac{R_{ij}}{\sqrt{\tilde{n}_{ij}}}, \text{ wobei } R_{ij} = n_{ij} - \tilde{n}_{ij} .$$

Klassenstufe * Migration 2 Gruppen (4,8) Kreuztabelle

		Migration 2 Gruppen (4,8)		Gesamt
		,00	1,00	
Klassenstufe 6	% von Migration 2 Gruppen (4,8)	,0%	,1%	,0%
	Korrigierte Residuen	-3,7	3,7	
7	% von Migration 2 Gruppen (4,8)	,9%	4,5%	1,3%
	Korrigierte Residuen	-19,5	19,5	
8	% von Migration 2 Gruppen (4,8)	12,3%	29,3%	14,4%
	Korrigierte Residuen	-29,8	29,8	
9	% von Migration 2 Gruppen (4,8)	64,9%	54,5%	63,6%
	Korrigierte Residuen	13,3	-13,3	
10	% von Migration 2 Gruppen (4,8)	21,9%	11,6%	20,6%
	Korrigierte Residuen	15,7	-15,7	
11	% von Migration 2 Gruppen (4,8)	,0%	,0%	,0%
	Korrigierte Residuen	1,1	-1,1	
Gesamt		100,0%	100,0%	100,0%

Chi-Quadrat-Tests

	Wert	df	Asymptotische Signifikanz (2-seitig)
Chi-Quadrat nach Pearson	1408,461 ^a	5	,000
Likelihood-Quotient	1168,728	5	,000
Zusammenhang linear-mit-linear	1137,396	1	,000
Anzahl der gültigen Fälle	34533		

a. 2 Zellen (16,7%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 1,00.

Abbildung 43 Chi-Quadrat-Test/Klassenstufe⁸⁸

Die Hypothese der Unabhängigkeit wird auf dem Niveau $p < 0.01$ verworfen.

Es stellt sich die Frage nach den Ursachen. Der Vergleich der Populationen in Bezug auf Klassenwiederholung zeigt, dass ein Teil dieser Schulalter-Differenz zur Vergleichsgruppe durch überproportionale Klassenwiederholung der Schülerinnen und Schüler mit Migrationshintergrund zu erklären ist. Diese weisen mit etwa 28 Prozent eine höhere Klassenwiederholung auf als die Vergleichsgruppe (nach Definition) mit etwa 16 Prozent.⁸⁹

Das mittlere Schulalter der beiden Vergleichsgruppen im Ergänzungstest derjenigen Schülerinnen und Schüler, die keine Klassenwiederholung berichten, zeigt ebenfalls eine statistisch signifikante Differenz von etwa einem viertel Schuljahr (Klassenstufe). Selbst unter Berücksichtigung von weiteren Faktoren als Kovariaten bleibt diese Größenordnung statistisch nachweisbar. Dieser Zusammenhang lässt sich für eine Reihe von Teilnehmerstaaten in unterschiedlicher Ausprägung nachweisen.⁹⁰

⁸⁸ Codierung Migration: Mit Migrationshintergrund: 1,00 [entsprechend Definition (Gruppe 4,8)]. Ohne Migrationshintergrund 0,00 [entsprechend Definition (Gruppe 1,2,3,5,6,7)].

⁸⁹ Der nationale Datensatz (PISA-OVE-15J) (gewichtet) weist insgesamt 73,14% Schülerinnen und Schüler auf, die keine Klassenwiederholung berichten. Einmal: 16,73%, Zweimal: 0,29%, Unallocated: 1,02%, missing: 1,30%, system: 7,51%, Gültige Prozente ohne Wiederholung: 81,12%, einmal: 18,55%, ein zweites Mal: 0,33%

⁹⁰ Es ergeben sich für die Länder Australien, Canada, Irland und Neuseeland signifikant höhere und für die Länder Österreich, Belgien, Brasilien, Tschechien, Deutschland, Griechenland, Ungarn und UK signifikant niedrigere Schulalter für Schülerinnen und Schüler mit Migrationshintergrund entsprechend Definition.

2.7 Vergleich IGLU/PISA

Prenzel, Geiser, Langeheine und Lobemeier stellen Vergleiche zwischen den einzelnen internationalen Studien an mit der Feststellung, dass einigen Teilnehmerstaaten eine Kompetenzentwicklung vom Ende der Grundschulzeit über die Sekundarstufe gelingt.⁹¹

*Im ersten IGLU-Bericht (Prenzel et al., 2003) war unter anderem hervorgehoben worden, dass Staaten wie die USA oder Deutschland im Grundschulbereich noch relativ gut abschneiden, dann jedoch bei den Sekundarstufenvergleichen deutliche Schwächen erkennen lassen. Demgegenüber lassen sich für andere Staaten Entwicklungen beobachten, die auf eine qualitätsvolle und beschleunigte Kompetenzentwicklung vom Ende der Grundschulzeit über die Sekundarstufe (etwa bis zum Ende der 9. Jahrgangsstufe oder bis zum 15. Lebensjahr) schließen lassen. Anhaltspunkt dafür sind die deutlichen Veränderungen in den Abständen zum deutschen Mittelwert bei IGLU-E und bei PISA. So ändern Staaten wie Kanada, Neuseeland, England oder Japan ihre relative Position zu Deutschland in einer Größenordnung von 40 Punkten oder mehr.*⁹²

Die Ausführungen zur Problematik unterschiedlicher Stichprobendefinitionen haben deutlich gemacht, dass als Ursache der Lokationsverschiebungen der Teilnehmerstaaten auf den Ratingskalen die Schwerpunktsetzung zugunsten des Schulalters beziehungsweise des Lebensalters zu sehen ist. Die von den Autoren angesprochenen Staaten Neuseeland, England (Vereinigtes Königreich) und Japan weisen Schülerpopulationen mit hohem Schulalter auf. Diese Schülerinnen und Schüler werden daher in einer am Lebensalter orientierten Stichprobe höhere Werte erreichen, in einer am Schulalter orientierten Studie hingegen aufgrund des geringen Grenzertrags früher Einschulung niedrigere. Deutlich wird dies an den sehr unterschiedlichen Testwerten Neuseelands, das im PISA-Lesetest mit 529 Punkten den dritten Platz erreichte, in IGLU mit 529 Punkten den 13. Platz. Zu bemerken ist hier, dass Neuseeland in der IGLU-Studie seine Stichprobe in der 5. Klassenstufe zog, während in den meisten übrigen

⁹¹ Prenzel, Geiser, Langeheine, Lobemeier: Naturwissenschaftliche Kompetenz am Ende der Grundschulzeit: Vergleiche zwischen einigen Ländern der Bundesrepublik Deutschland. In: Bos, Lankes, Prenzel, Schwippert, Valtin, Walther (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II) 2004. S. 104ff

⁹² Prenzel, Geiser, Langeheine, Lobemeier: Naturwissenschaftliche Kompetenz am Ende der Grundschulzeit: Vergleiche zwischen einigen Ländern der Bundesrepublik Deutschland. In: Bos, Lankes, Prenzel, Schwippert, Valtin, Walther (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II) 2004. S. 104ff

Teilnehmerstaaten die Schülerinnen und Schüler der 4. Klassenstufe bewertet wurden. Die Frage stellt sich demnach, ob die einzelnen Studien verglichen werden können.

*Für diese Staaten und Länder können die in den beiden Studien erzielten Ergebnisse zueinander in Beziehung gesetzt werden.*⁹³

Werden die Ergebnisse der einzelnen Studien im Rahmen einer Korrelationsanalyse in Beziehung gesetzt, so liegt mit $r_p = 0,193/p = 0,389$ beziehungsweise $r_s = 0,173/p = 0,44$ kein statistisch signifikanter Zusammenhang vor, d.h. die Ausgangshypothese $H_0 : \rho = 0$ kann nicht verworfen werden.⁹⁴

⁹³Walther, Geiser, Langeheine, Lobemeier: Mathematische Kompetenzen am Ende der vierten Jahrgangsstufe in einigen Ländern der Bundesrepublik Deutschland. In: Bos, Lankes, Prenzel, Schwippert, Valtin, Walther (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II) 2004. S. 128

⁹⁴ Korrelationsrechnung über die einzelnen Teilnehmerstaaten ohne Berücksichtigung der unterschiedlichen N_i .

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

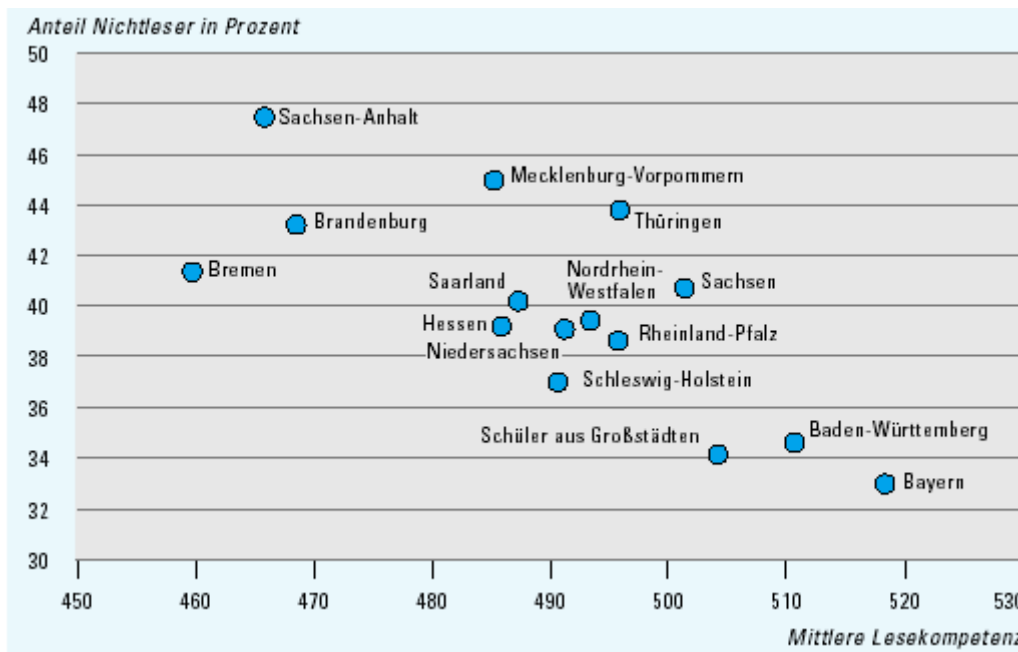


Abbildung 45 Anteil Nichtleser und mittlere *Lesekompetenz* (Quelle: PISA)

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

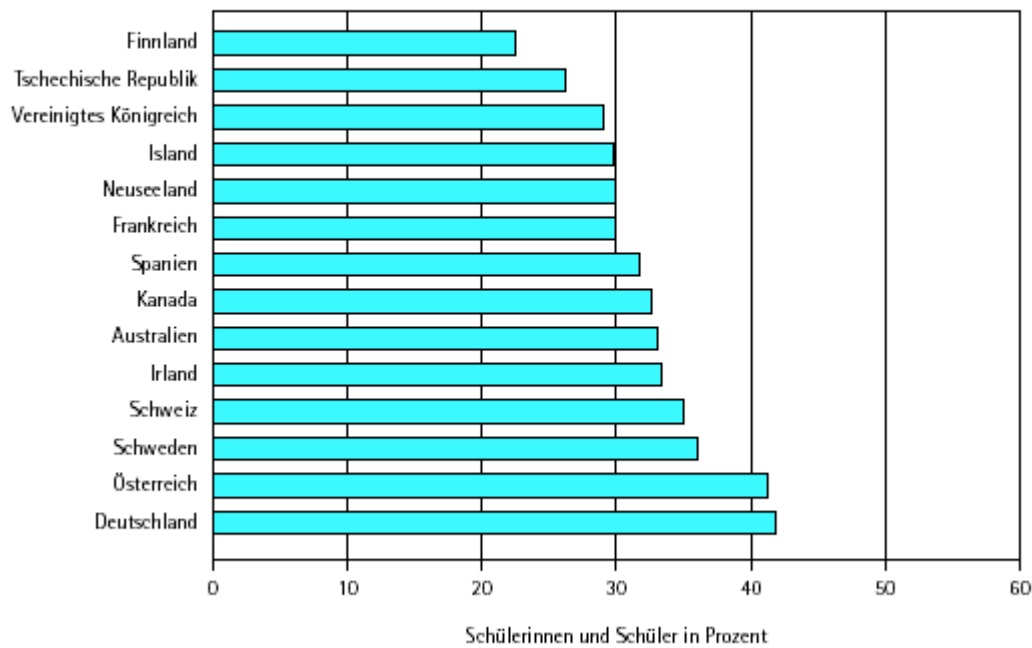


Abbildung 46 Nichtleser im internationalen Vergleich (Quelle: PISA)

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

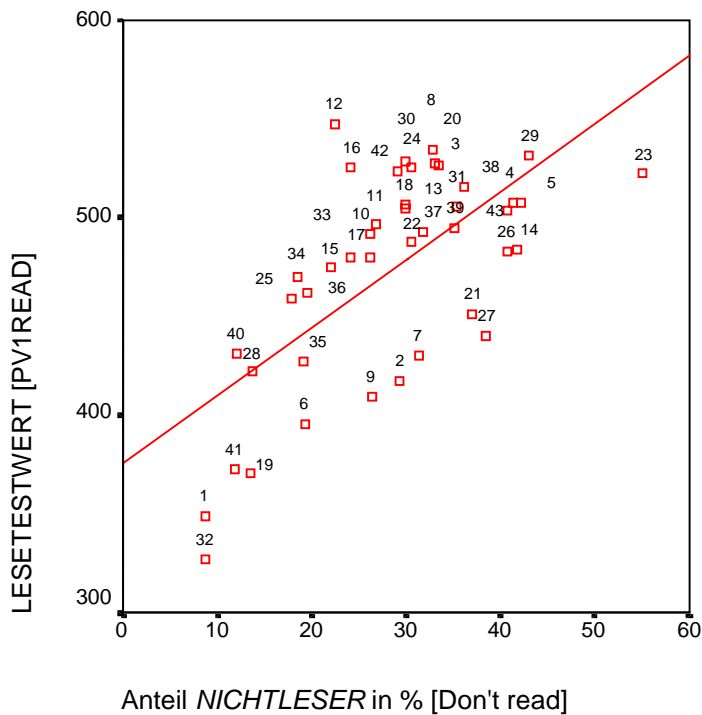
CASE NO.	Read each day - Q34					GRADE	HISEI	PV1READ	
	Don't read	30 min or less	31- 60 min	1-2 hours	More than 2 hours				
	percent								
ALBANIA	1	8,63	19,10	30,63	27,63	14,02	9,71	45,55	349,02
ARGENTINA	2	29,28	30,81	21,50	11,37	7,03	9,59	43,35	417,41
AUSTRALIA	3	33,10	30,47	20,54	11,80	4,00	10,10	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	9,73	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	4,00	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	2,80	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	5,95	10,00	42,80	525,03
LATVIA	25	17,86	25,61	29,59	19,66	7,20	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	7,90	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	2,90	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).

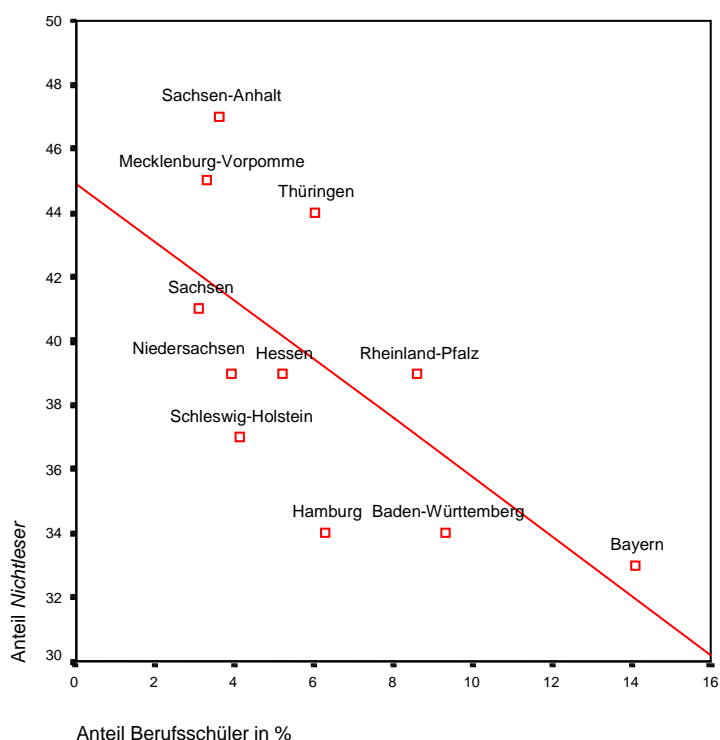
Abbildung 48 *Nichtleser* - Testwert Lesen International/Korrelation

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

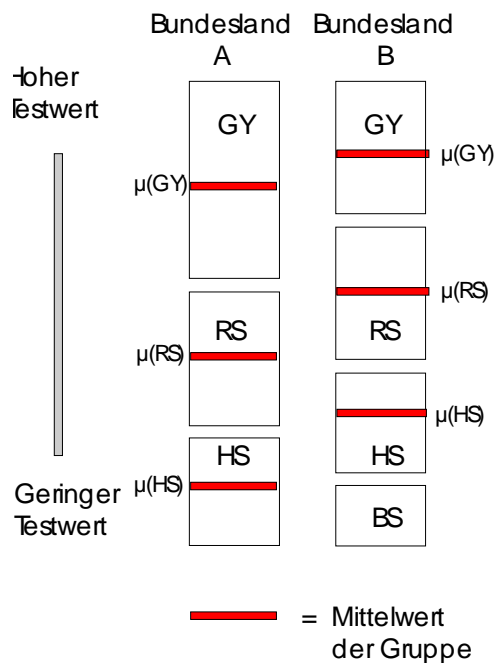
Es zeigt sich ein negativer, linearer Zusammenhang zwischen Anteil der *Nichtleser* und prozentualem Anteil der Berufsschüler des Jahrgangs eines Bundeslandes: Je mehr Berufsschüler ein Bundesland aufweist, desto weniger *Nichtleser* finden sich in den

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

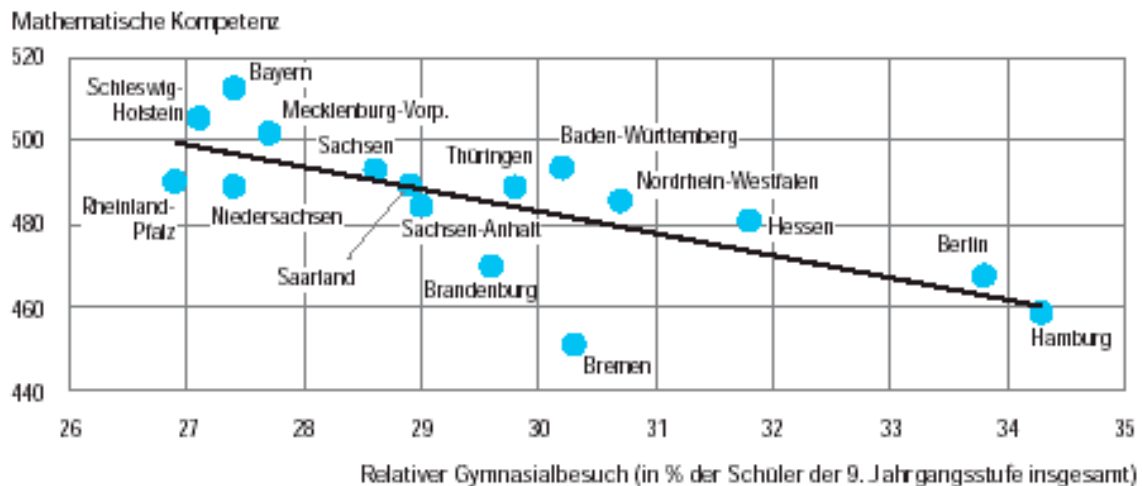


Abbildung 52 Schüler in Gymnasien – *Mathematische Kompetenz* (Quelle: PISA)

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

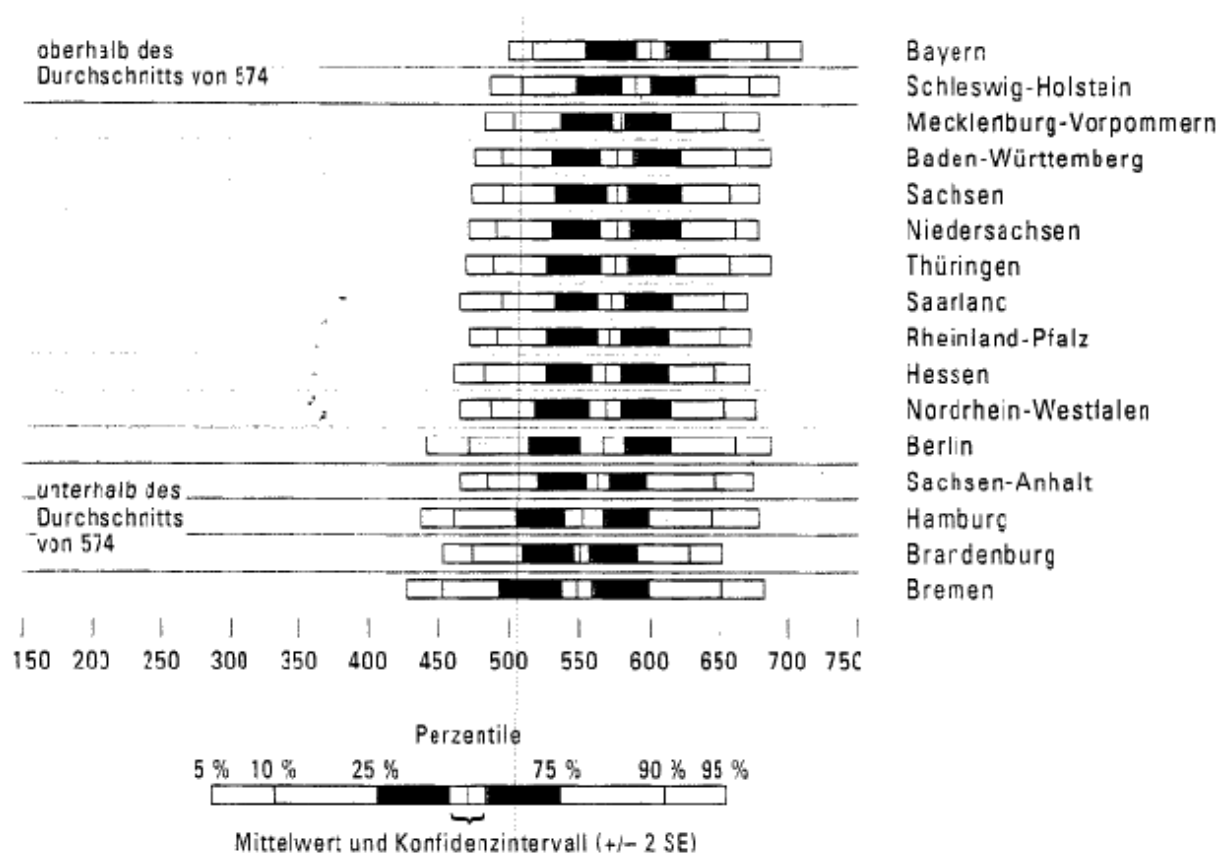
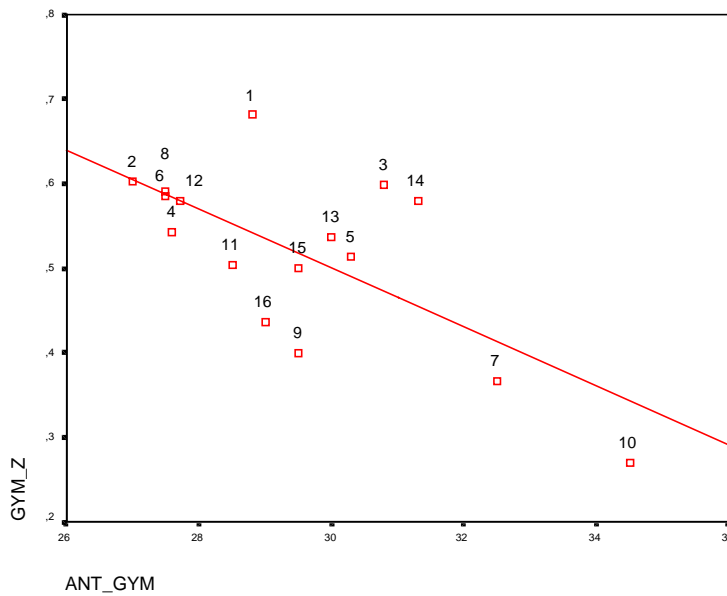


Abbildung 53 Perzentilbänder Testwert Lesen Gymnasien (Quelle: PISA)

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

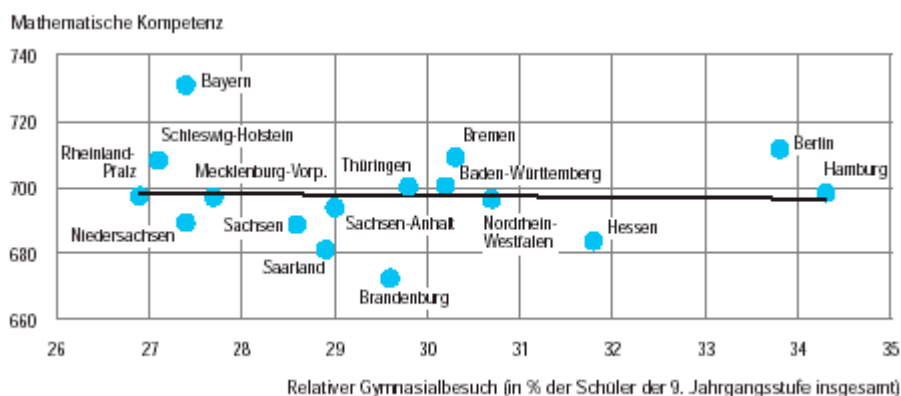


Abbildung 55 Zusammenhang zwischen relativem Gymnasialbesuch und Leistungsniveau der oberen 5 % der Gymnasiasten der 9. Jahrgangsstufe in Mathematik auf Länderebene (Mittelwerte) (Quelle: PISA 2000 Zusammenfassung zentraler Befunde, S.68)

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

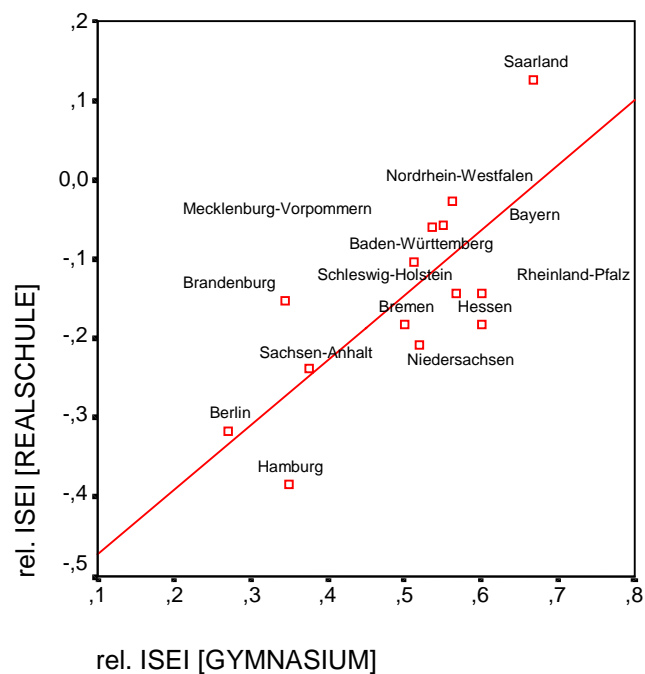
Die folgende Abbildung zeigt den Zusammenhang von durchschnittlichem (je Bundesland) z-standardisierten sozioökonomischen Index der Schülerschaft aus Gymnasien und Realschulen auf der Datenbasis der 31.751 15-Jährigen der 9. Klassenstufe (PISA 2000-E). Der positive lineare Zusammenhang der sozioökonomischen 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															Fach-Berufs-akademien	Gesamt	
		Hauptschule	Schularten mit mehreren Bildungsgängen	Realschulen	Gymnasien	Integrierte Gesamtschulen	Freie Waldorfschulen	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	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%	,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	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%	,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	Anzahl	4600	0	3400	6100	6300	100	600	2100	3600	1300	0	2000	0	700	200	300	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%	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	1440
		% von Bundesland	,0%	,0%	,0%	27,8%	55,6%	2,1%	,7%	6,9%	6,9%	,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	-3,5	-1,0	-1,0
	HB Bremen	Anzahl	1200	0	1000	1200	600	10	300	600	300	0	0	800	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%	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	5,1	-7,7	-7,9	-2,3	-2,3
	HH Hamburg	Anzahl	4300	900	1800	5200	7300	100	200	2300	3000	2300	0	2400	0	200	400	300	30700
		% von Bundesland	14,0%	2,9%	5,9%	16,9%	23,8%	,3%	,7%	7,5%	9,8%	,0%	,0%	7,8%	,7%	,9%	1,3%	1,0%	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	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%	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-Vorpommern	Anzahl	100	100	300	300	100	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%	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	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%	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	46,0	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	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%	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	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%	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	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%	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	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%	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	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%	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	13030
		% von Bundesland	29,9%	,0%	16,9%	13,8%	4,6%	,2%	1,5%	8,4%	15,3%	,8%	,0%	5,4%	,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	960
		% von Bundesland	,0%	62,5%	,0%	31,3%	2,1%	,0%	,0%	4,2%	,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	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%	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

	Schwache Leser rec	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Internat. Lesen PV 1	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	
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

Schwache Leser rec		N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Internat. Lesen PV 1	nicht identifiziert	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	
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

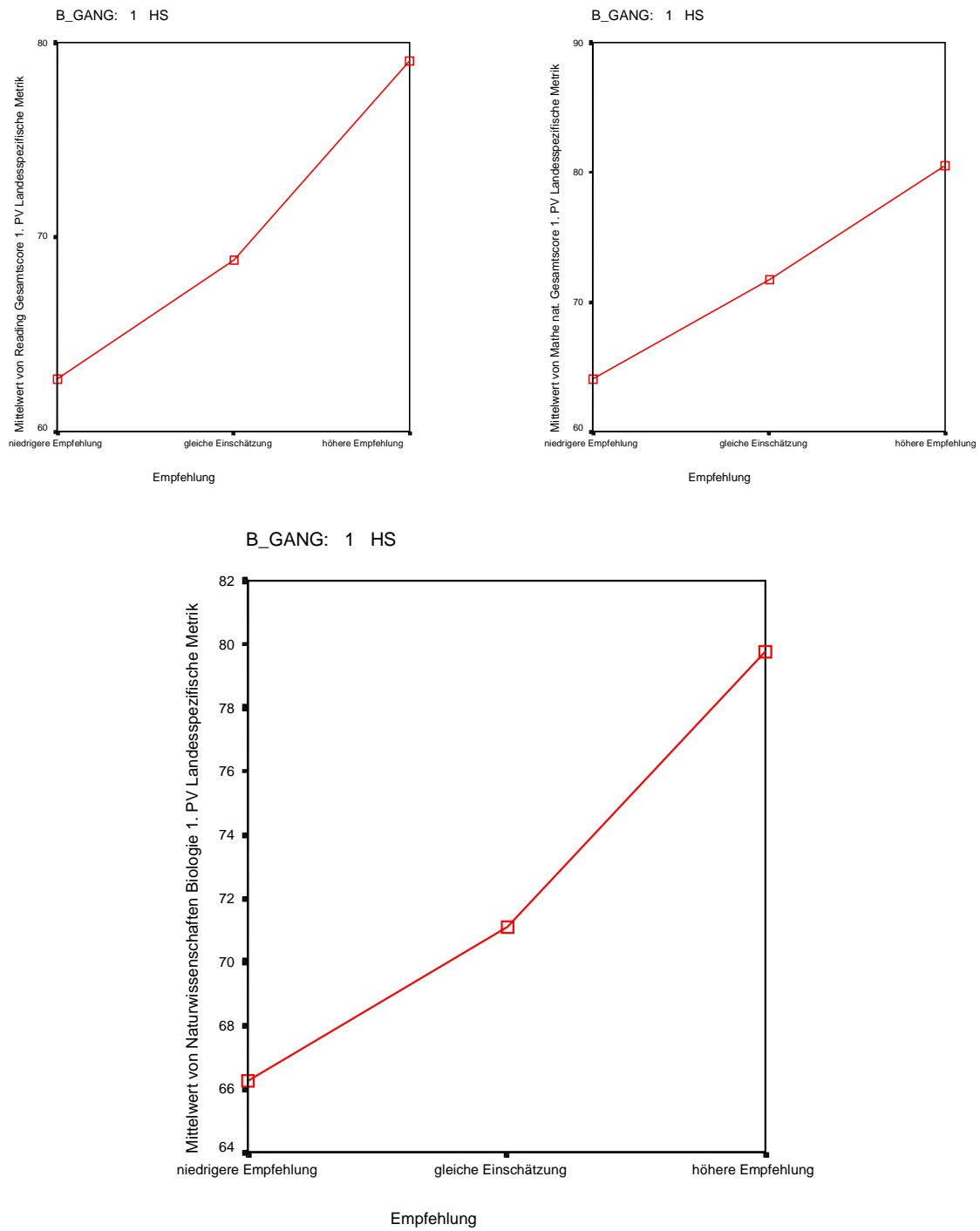


Abbildung 60 Leseschwäche Diagnose(III) Hauptschule

3.5.2.2 Realschule

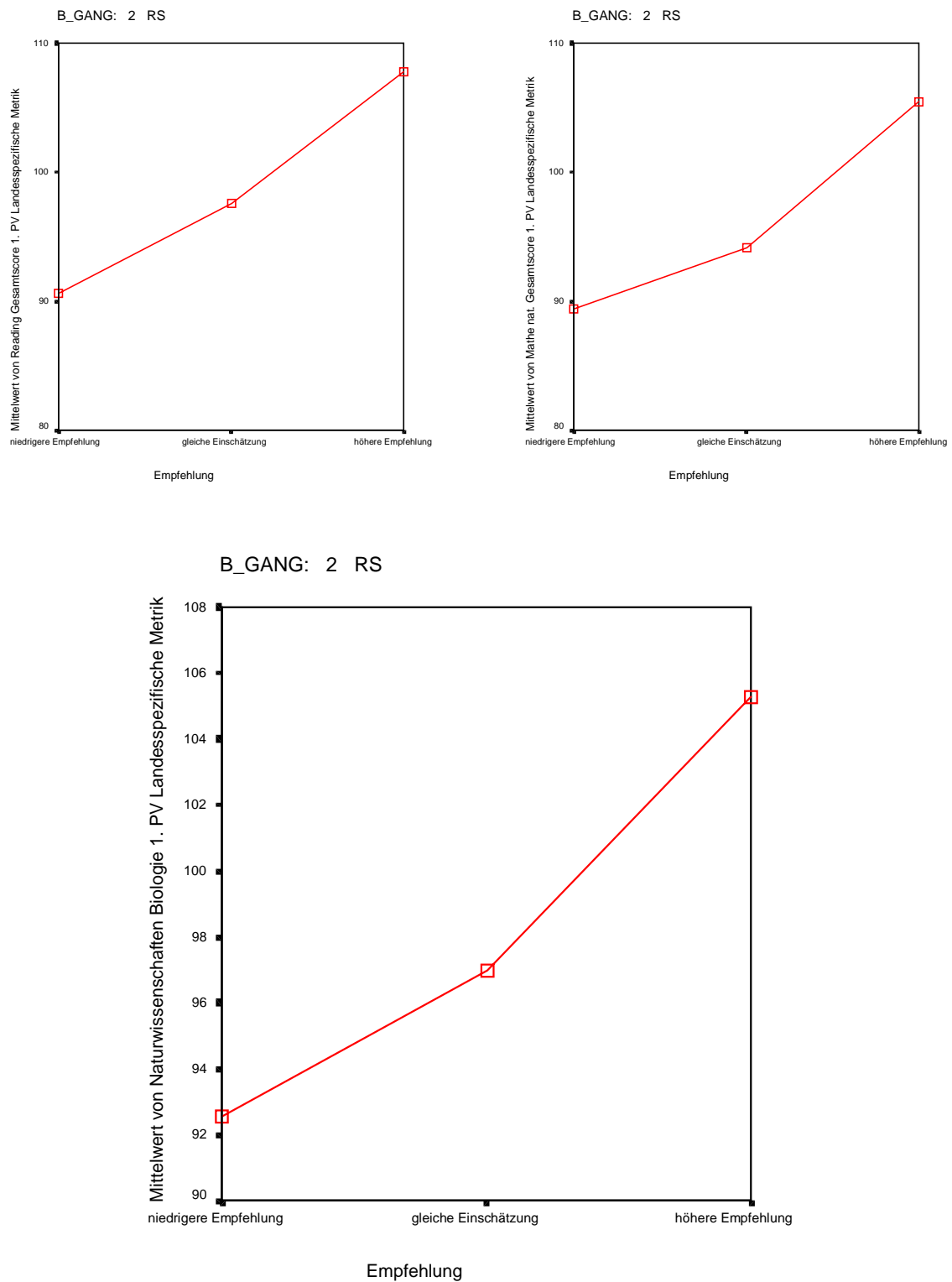


Abbildung 61 Leseschwäche Diagnose(III) Realschule

3.5.2.3 Gymnasium

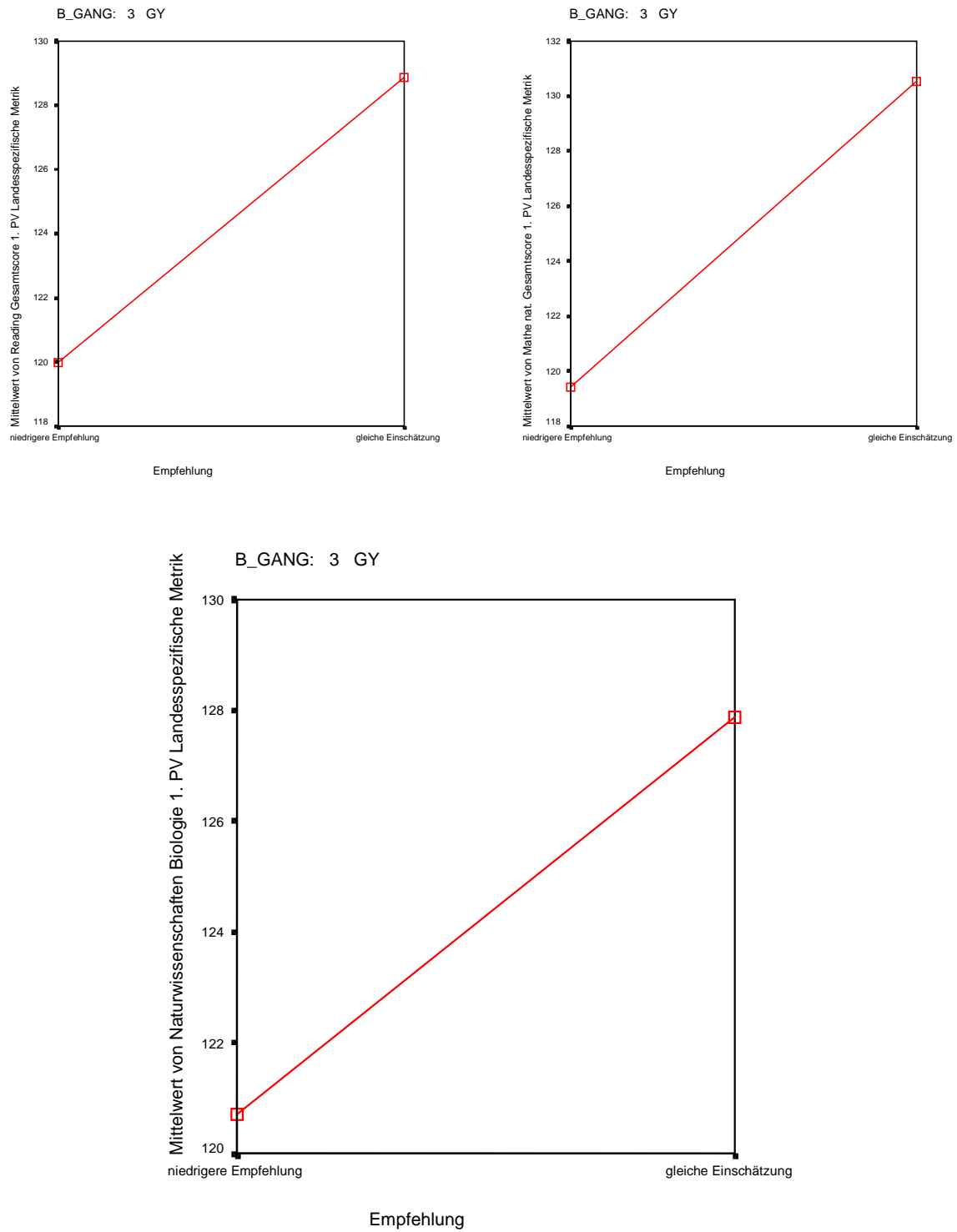


Abbildung 62 Leseschwäche Diagnose(III) Gymnasium

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

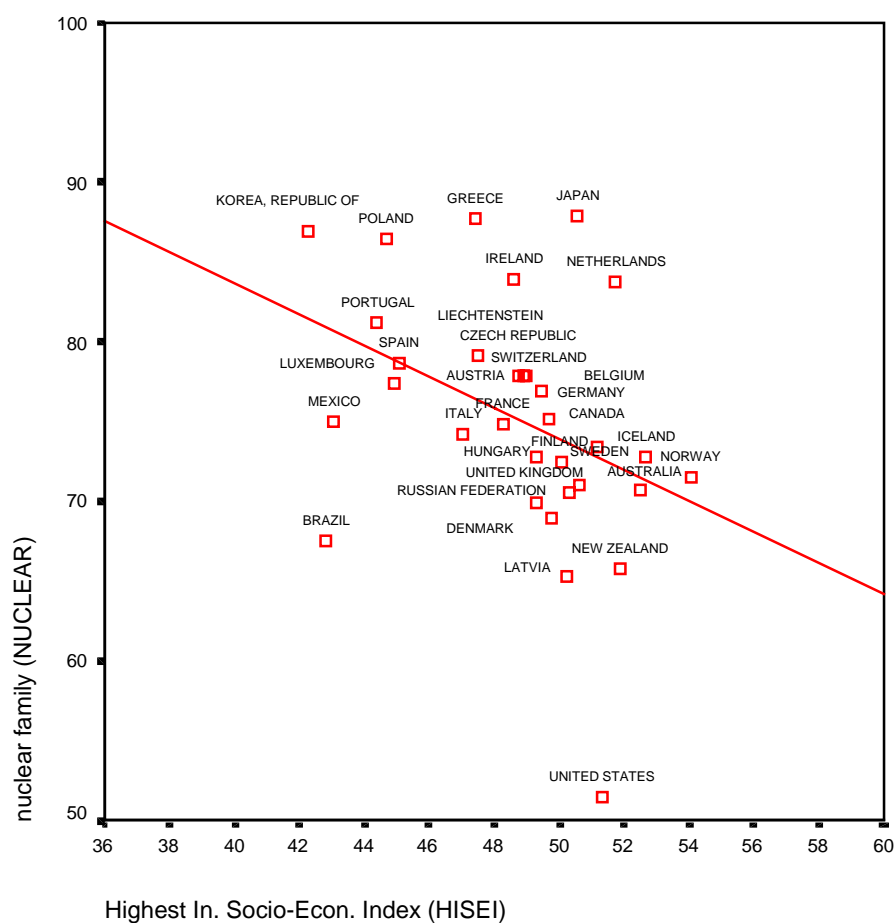


Abbildung 63 Anteil Kernfamilie in Abhängigkeit zum sozioökonomischen Index

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.

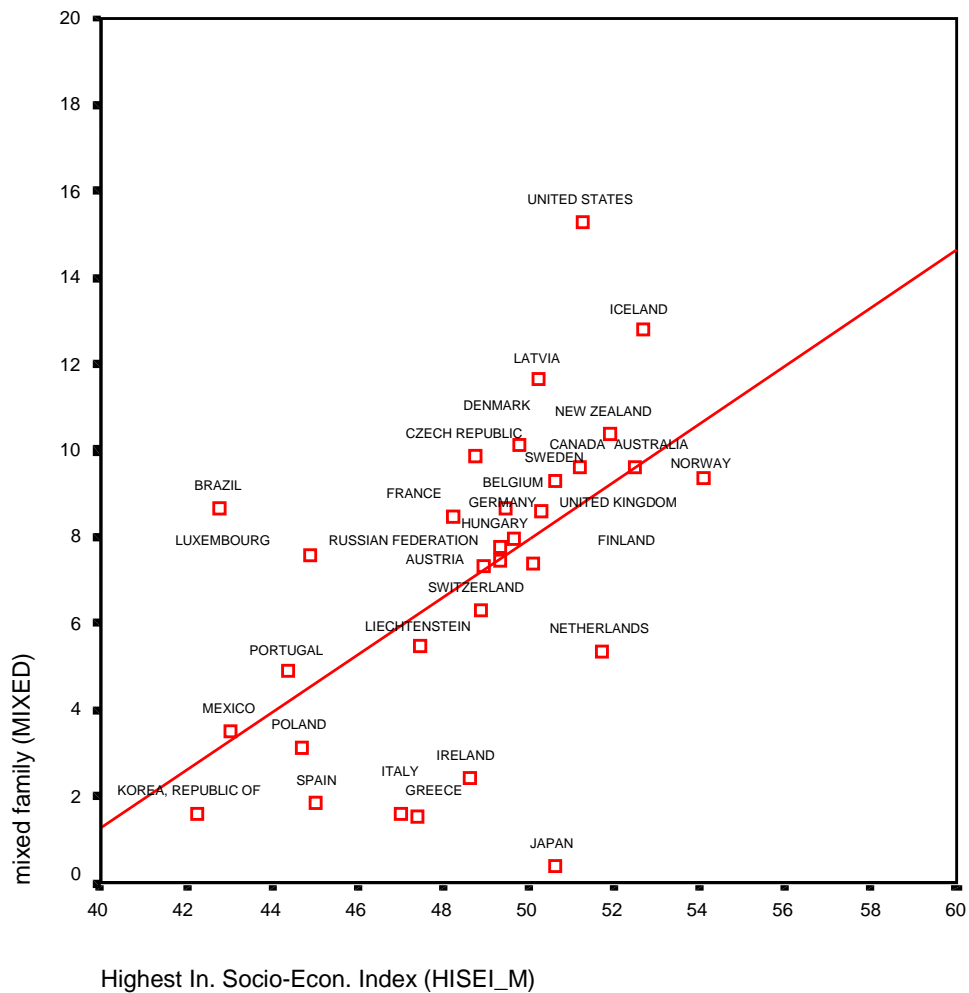
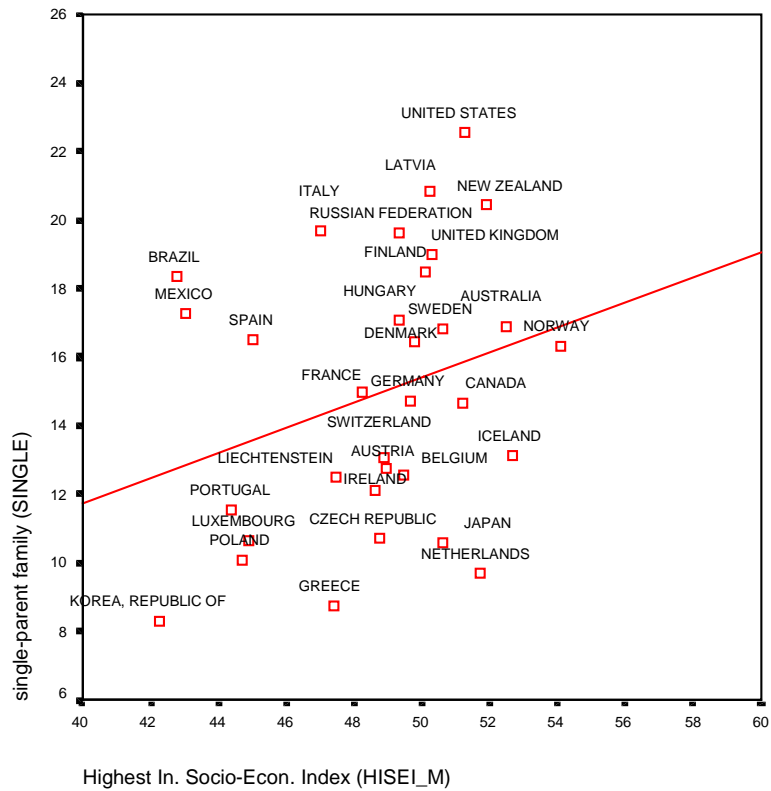


Abbildung 65 Anteil Neuer Partnerschaften (mixed family) in Abhängigkeit zum sozioökonomischen Index



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

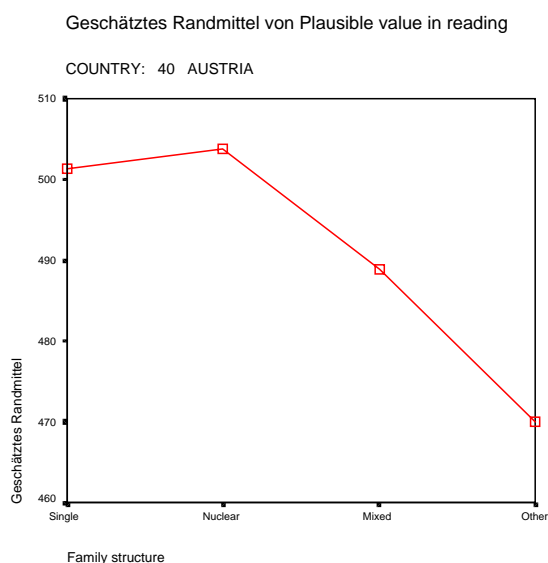
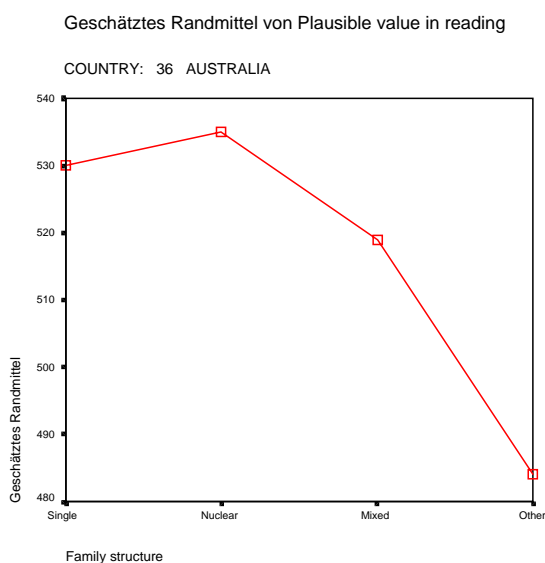
			HISEI_M
Spearman-Rho	SINGLE	Korrelationskoeffizient	,304
		Sig. (2-seitig)	,090
		N	32
NUCLEAR		Korrelationskoeffizient	-,474*
		Sig. (2-seitig)	,006
		N	32
MIXED		Korrelationskoeffizient	,622*
		Sig. (2-seitig)	,000
		N	32
OTHER		Korrelationskoeffizient	-,235
		Sig. (2-seitig)	,195
		N	32

** . 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

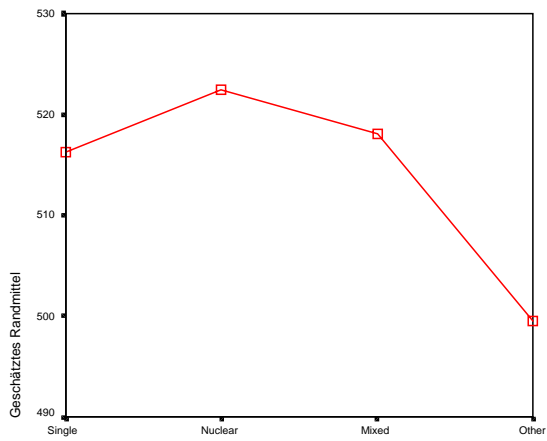


¹³¹ Schümer, Weiß, Steinert, Baumert, Tillmann, Meier: Lebens- und Lernbedingungen von Jugendlichen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S. 480

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

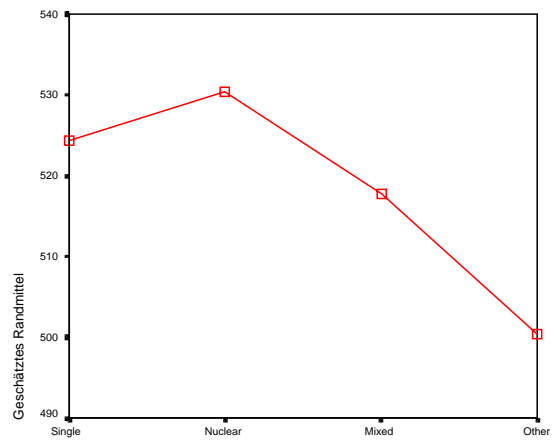
COUNTRY: 56 BELGIUM



Family structure

Geschätztes Randmittel von Plausible value in reading

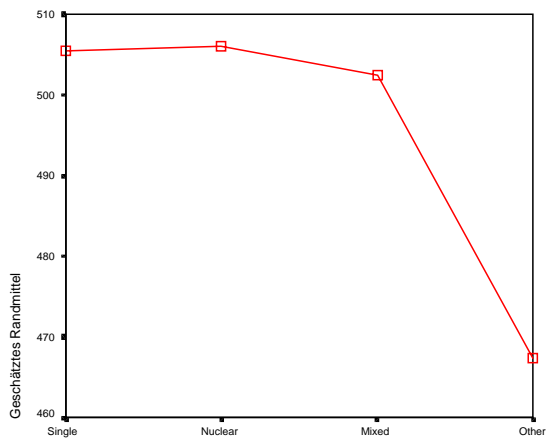
COUNTRY: 124 CANADA



Family structure

Geschätztes Randmittel von Plausible value in reading

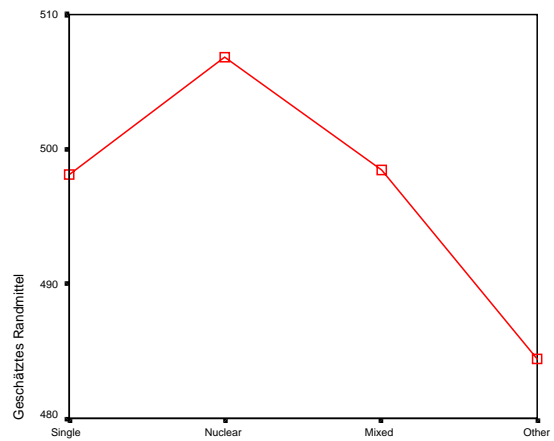
COUNTRY: 203 CZECH REPUBLIC



Family structure

Geschätztes Randmittel von Plausible value in reading

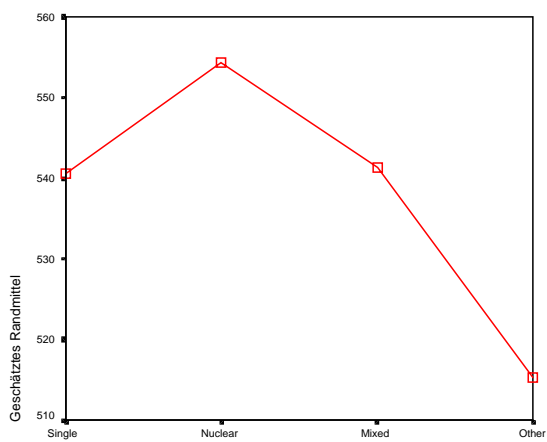
COUNTRY: 208 DENMARK



Family structure

Geschätztes Randmittel von Plausible value in reading

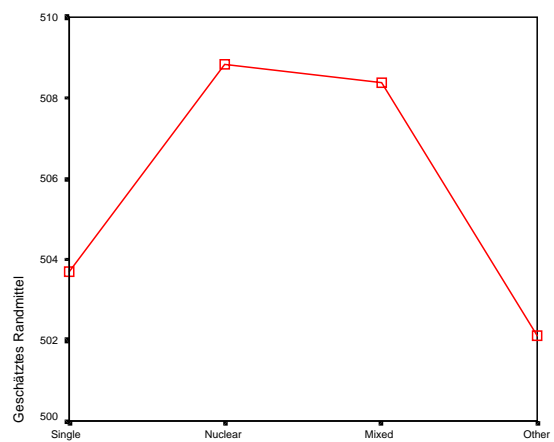
COUNTRY: 246 FINLAND



Family structure

Geschätztes Randmittel von Plausible value in reading

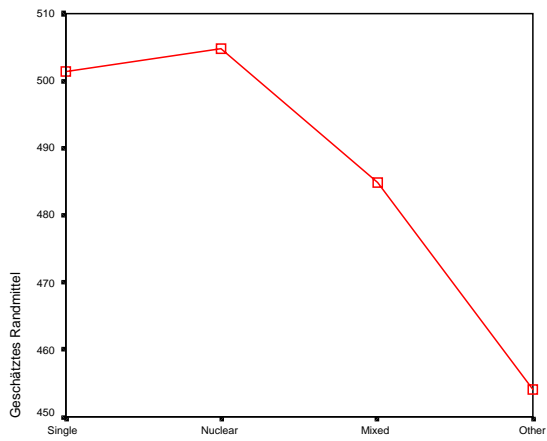
COUNTRY: 250 FRANCE



Family structure

Geschätztes Randmittel von Plausible value in reading

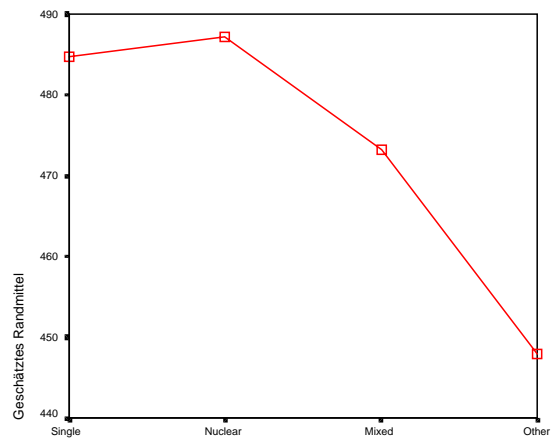
COUNTRY: 276 GERMANY



Family structure

Geschätztes Randmittel von Plausible value in reading

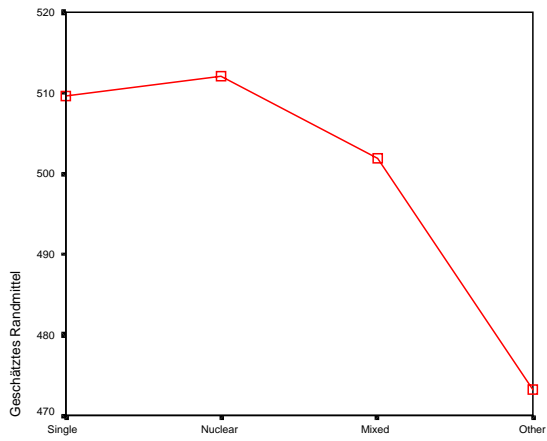
COUNTRY: 348 HUNGARY



Family structure

Geschätztes Randmittel von Plausible value in reading

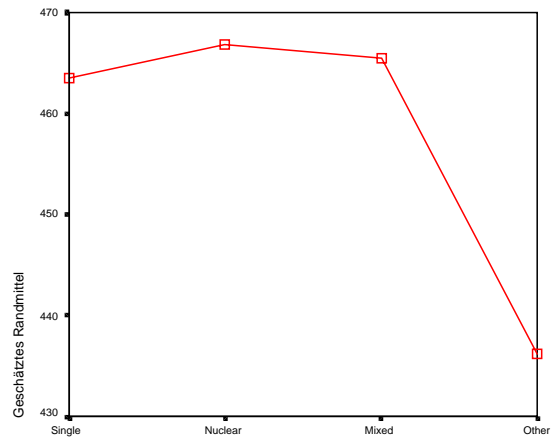
COUNTRY: 352 ICELAND



Family structure

Geschätztes Randmittel von Plausible value in reading

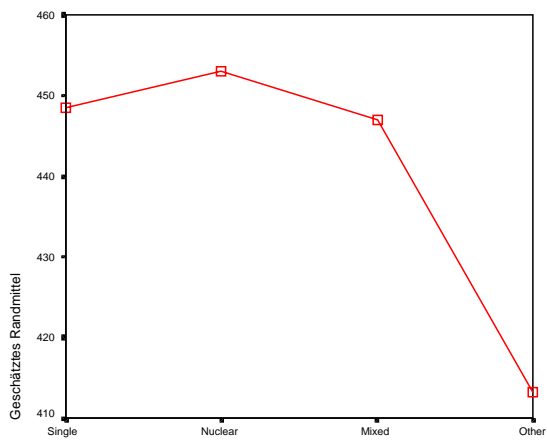
COUNTRY: 428 LATVIA



Family structure

Geschätztes Randmittel von Plausible value in reading

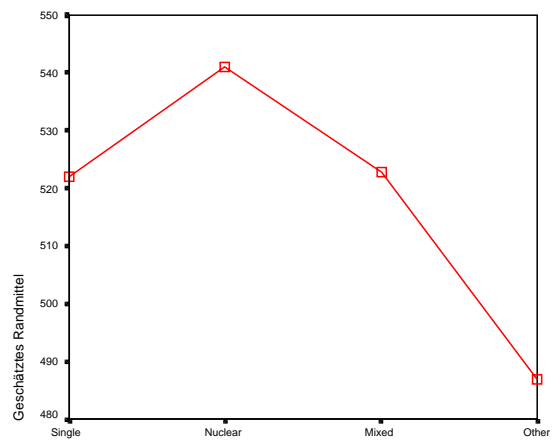
COUNTRY: 442 LUXEMBOURG



Family structure

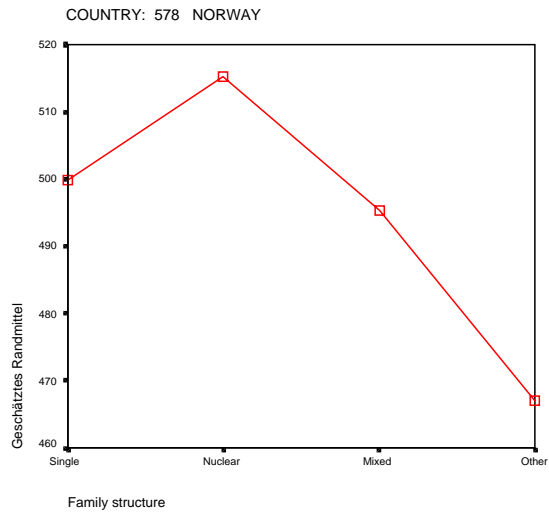
Geschätztes Randmittel von Plausible value in reading

COUNTRY: 554 NEW ZEALAND

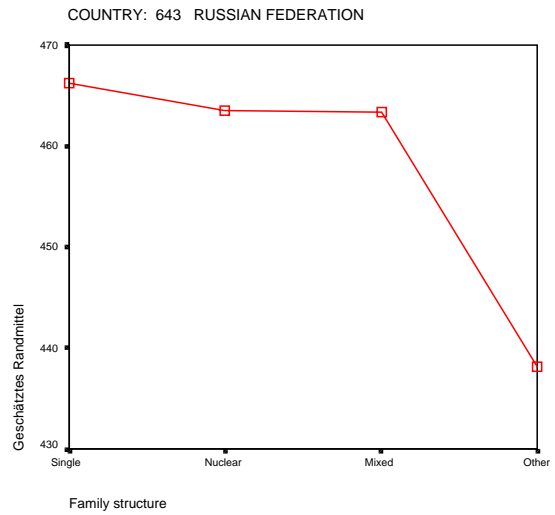


Family structure

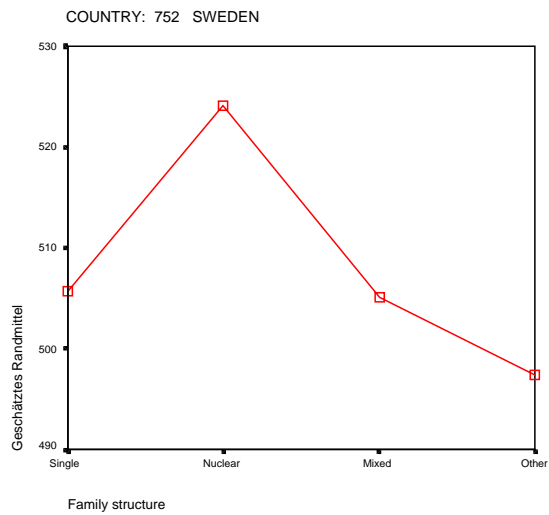
Geschätztes Randmittel von Plausible value in reading



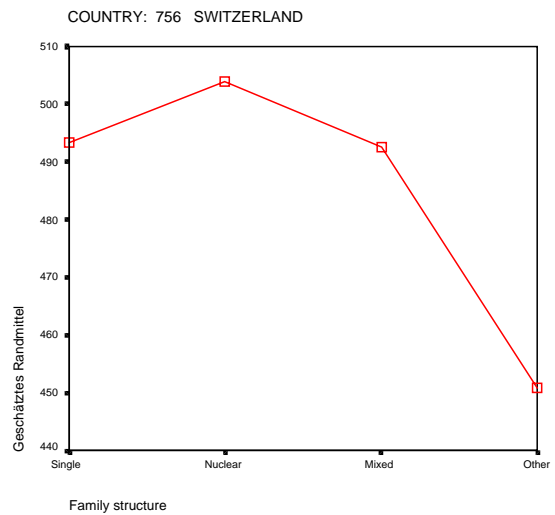
Geschätztes Randmittel von Plausible value in reading



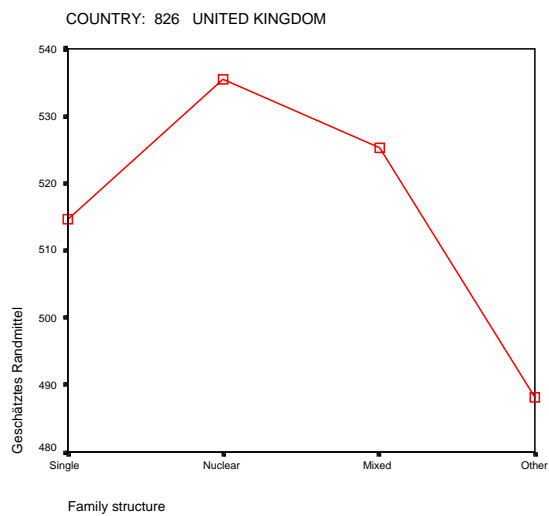
Geschätztes Randmittel von Plausible value in reading



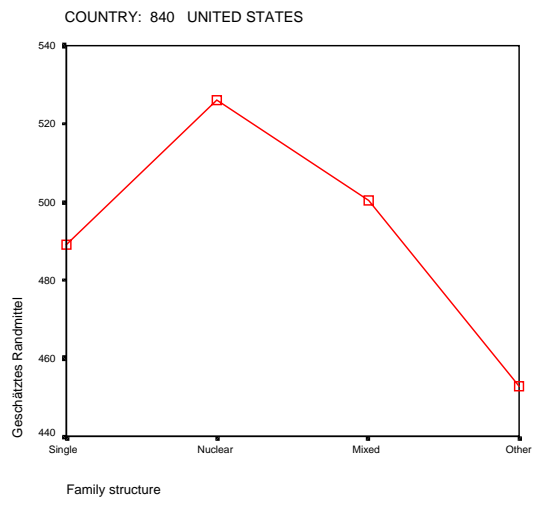
Geschätztes Randmittel von Plausible value in reading



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.

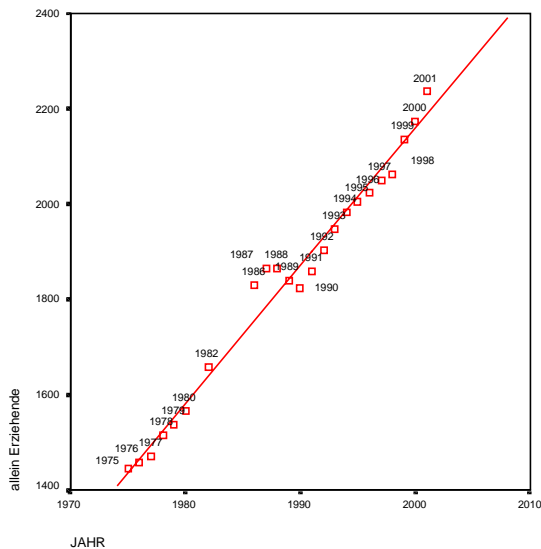
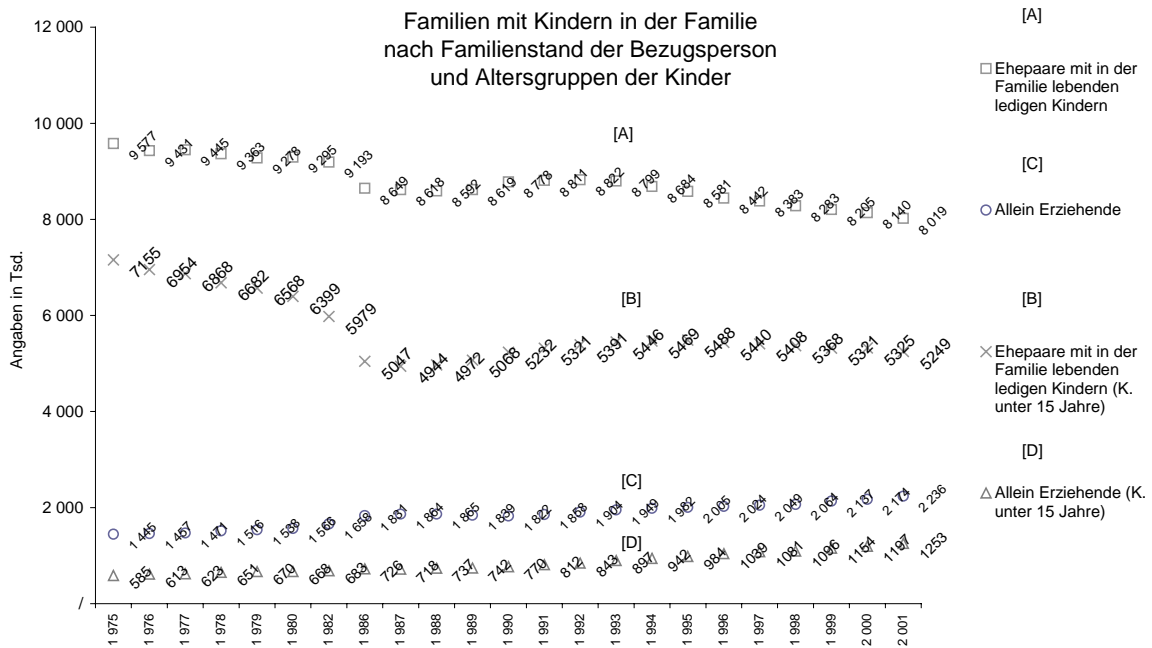
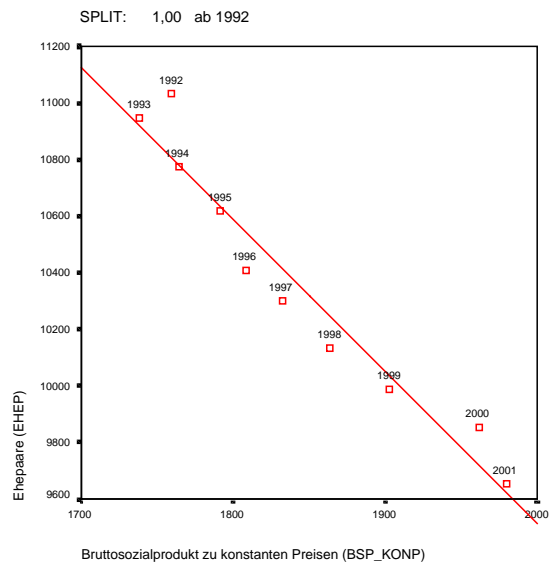
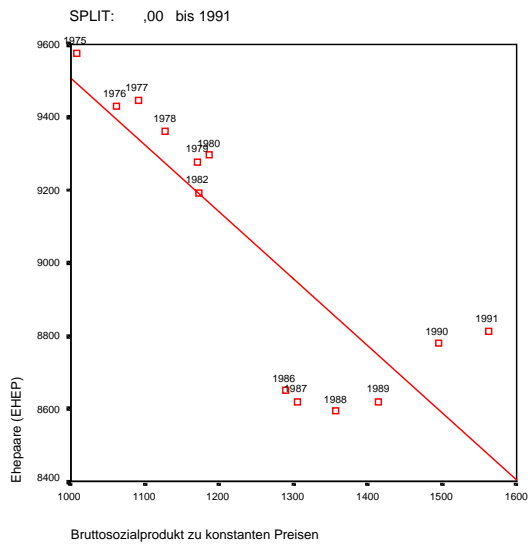
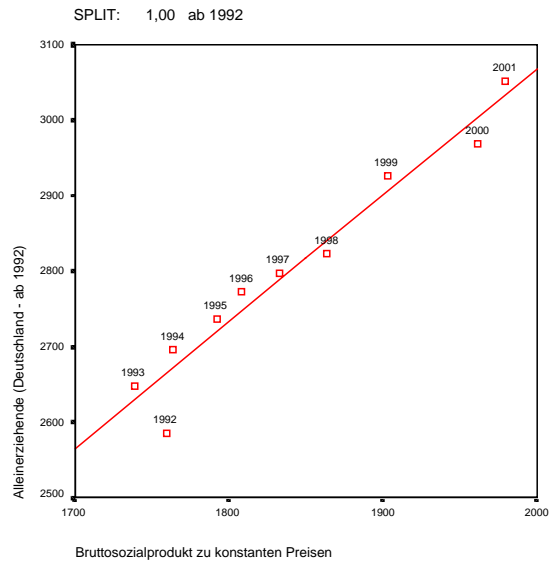
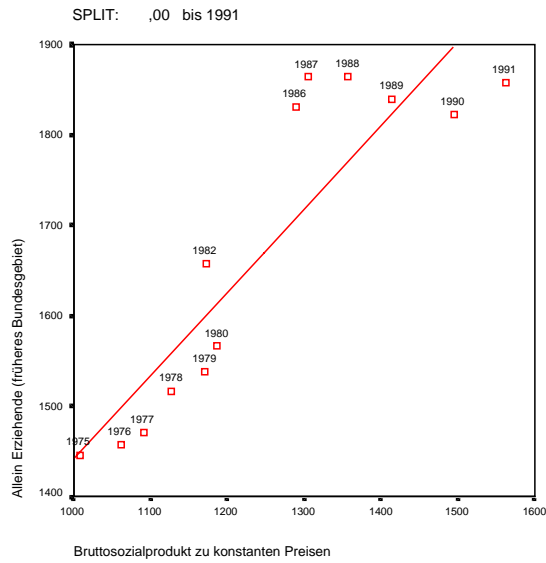


Abbildung 68 Familie längsschnittanalytische Betrachtung

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.

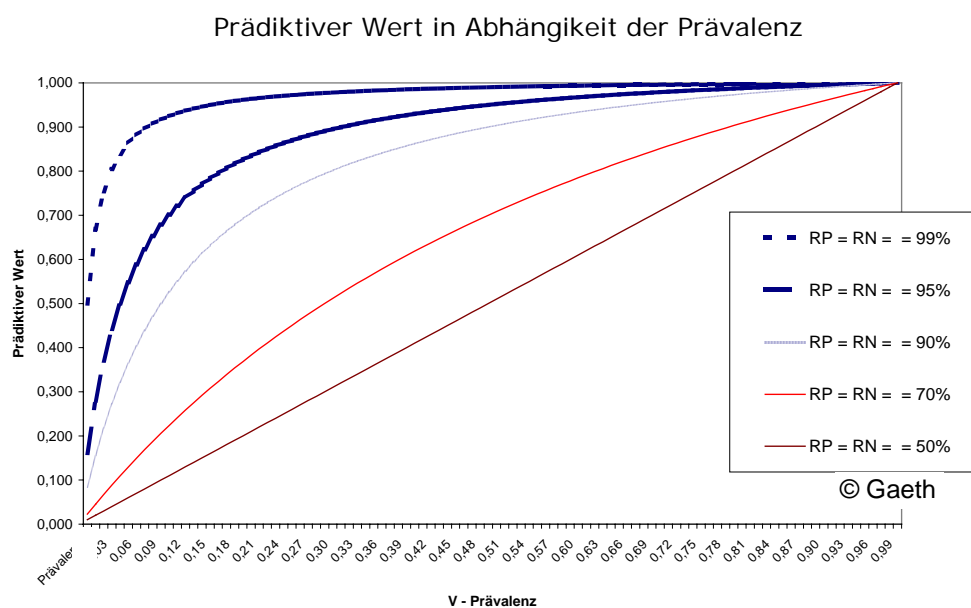


Abbildung 70 Prädiktiver Wert in Abhängigkeit von der Prävalenz bei verschiedenen Sensitivitäten und Spezifitäten

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

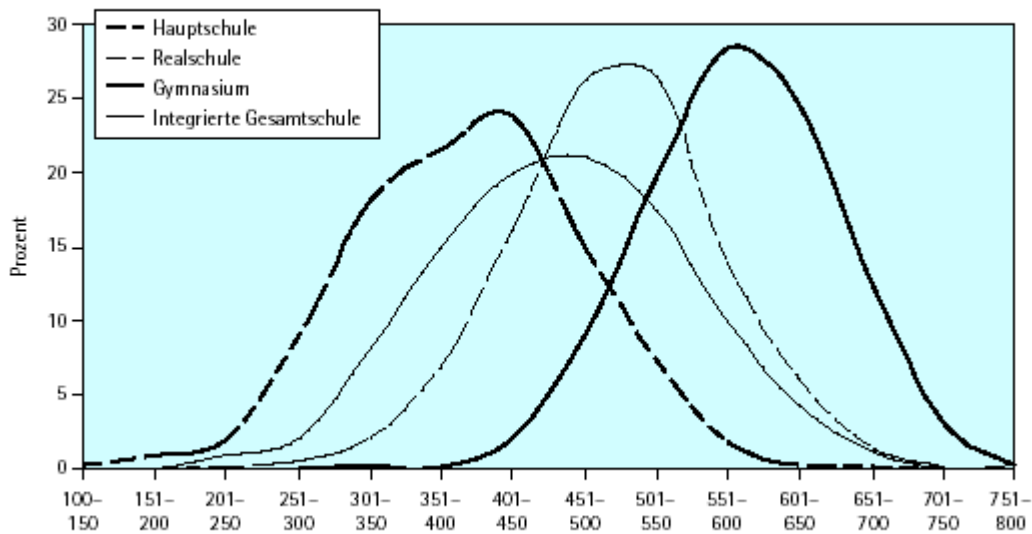


Abbildung 71 Bildungsgang und Testleistung (Quelle: PISA 2000)

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%	94,2%	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	3,072	5,419
	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.

4 Zusammenfassung der Ergebnisse

4.1 International

Die vorliegende Untersuchung konnte den Zusammenhang von Testleistung und zwischen den Teilnehmerstaaten systematisch variierendem Schulalter mit einer Reihe statistischer Verfahren nachweisen. Die Auswahl der Verfahren orientierte sich dabei an den Vergleichsuntersuchungen TIMSS, PISA und IGLU. Die geringen Testwerte der Schülerinnen und Schüler im deutschen Sample sind als Ergebnis des Testdesigns, das heißt einer Schwerpunktsetzung zugunsten des Lebensalters zu bewerten.

Es hat sich gezeigt, dass weitere Variablen in der Auswertung berücksichtigt werden müssen: insbesondere das Schulalter, die unterschiedliche Rückhaltequote (Retentivität) der Schulsysteme, der Integrationsaspekt, die Bildungsexpansion, der Anteil zusätzlicher Beschulung (insbesondere in Japan und Korea) sowie die Verteilung der Variablen *Geschlecht*. Unter Berücksichtigung der Variablen *Klassenstufe* und *Geschlecht* erreichen bundesdeutsche Schülerinnen und Schüler Testleistungen im oberen Leistungsbereich der Vergleichsländer.

Die Studien TIMSS, PISA und IGLU weisen unterschiedliche Populationsdefinitionen auf, in der Regel eine Schwerpunktsetzung zugunsten einer Lebensalterdefinition, einer Schulalterdefinition oder einer Kombination aus beidem. Aus diesem Grund ist der direkte Vergleich der Testwerte einzelner Studien als längsschnittanalytischer Leistungsvergleich beziehungsweise als Messung von Leistungsfortschritten ungeeignet.

Der Technical Report verweist auf einen Mangel präziser Kenntnisse der zu untersuchenden Populationen. Hier wurde unter anderem Österreich angeführt. Vor diesem Hintergrund sind die Testwerte aber auch die Gewichtungszurteilung zu beurteilen.

Ein weiteres Problem stellt die variierende Rückhaltequote (Retentivität) der Schulsysteme dar. Ein positiver Zusammenhang zwischen Rückhaltequote (Retentivität) und Testleistung konnte nicht bestätigt werden. Besonders betroffen sind hier die Testleistungen Koreas.

Die Stichprobendefinition nach Lebensalter ist als wesentliche Ursache für *Disparitäten* anzusehen. Zudem sind so genannte Decken- und Bodeneffekte nachzuweisen. Es hat sich gezeigt, dass es bislang nicht gelungen ist, über eine Leistungsbandbreite von Schülern der 4. bis 13. Klassenstufe ein gleichermaßen geeignetes Testverfahren zu entwickeln. Die Schwerpunktsetzung zugunsten einer am Lebensalter orientierten

Stichprobendefinition angesichts erheblich unterschiedlich strukturierter Schulsysteme der Teilnehmerstaaten verstärkt die gemessene Leistungsheterogenität derart, dass Vergleiche nicht mehr zielführend erscheinen.

Die Definition der Stichprobe nach Lebensalter schafft für diejenigen Länder ohne strenge Koppelung von Lebens- und Schulalter, das sind mehr als zwei Drittel der Teilnehmerstaaten, Untersuchungspopulationen, die nicht mit schulischen Verhältnissen übereinstimmen. Man könnte unter Bezug auf schulische Verhältnisse von fiktiven Untersuchungspopulationen sprechen. Die Testergebnisse sind lediglich eine Folge der unterschiedlichen Gewichtung einzelner Schulalteranteile. Wenn diese über die verschiedenen Studien (PISA2000, PISA2003) relativ konstant bleiben, ändern sich auch die Testwerte nicht. Diese Konstanz als Reliabilität zu interpretieren ist nicht zielführend.

4.2 National (Bundesrepublik)

Die differierenden mittleren Testergebnisse der einzelnen Bundesländer erweisen sich im wesentlichen als Ergebnis unterschiedlicher Bildungsexpansion. Es hat sich gezeigt, dass die berichteten Mittelwerte mit fortschreitender Bildungsexpansion negativ korreliert sind und sich daraus statistische Effekte ergeben (Stage Migration).

Der Vergleich der in PISA gezogenen Schülerpopulation mit Daten des Statistischen Bundesamtes weist Diskrepanzen sowohl in Bezug auf das Segment der Schülerinnen und Schüler mit Migrationshintergrund als auch in Bezug auf das Segment der Schüler im Dualen System der Berufsschulen auf. Es ist mit hoher Wahrscheinlichkeit davon auszugehen, dass in der Stichprobe derjenigen Bundesländer mit einem Anteil 15-jähriger Schüler in Berufsschulen, diese Gruppe nicht adäquat repräsentiert ist. Hieraus ergeben sich einerseits zum internationalen Vergleich gegensätzliche Zusammenhänge (Beispiel *Nichtleser*), andererseits überschätzte mittlere Testwerte. Ein direkter Vergleich dieser mittleren Werte ist daher ebenfalls nicht sinnvoll.

Daten des Statistischen Bundesamtes bezüglich der strukturellen Zusammensetzung der Bevölkerung zeigen einen kontinuierlichen anteiligen Rückgang der Kernfamilie an der Gesamtbevölkerung in moderaten Grenzen. Dieser Sachverhalt ist über die gesamten Teilnehmerstaaten auch querschnittsanalytisch nachzuweisen. Aus dieser mit steigendem sozioökonomischem Niveau kovariierenden soziodemographischen Entwicklung werden Konsequenzen bezüglich der Schülerleistungen zu erwarten sein.

Ein weiterer Aspekt des nationalen Ergänzungstests betraf die Erkennung leseschwacher Schüler durch Lehrkräfte. Hier konnte am Beispiel verschiedener systematischer Untersuchungen gezeigt werden, dass Lehrkräfte testleistungsstarke und weniger testleistungsstarke Schülerinnen und Schüler identifizieren können. Weiterhin wurde gezeigt, dass für die Bewertung eines Testverfahrens die Gleichsetzung von Testergebnis und Testeigenschaften nicht sinnvoll ist.

Es hat sich gezeigt, dass für den Nachweis der Koppelung von sozioökonomischem Hintergrund der Eltern und Testleistung des Schülers der Schulformvergleich ungeeignet ist. Die Entkoppelung von Schulform und Schulabschluss ist ein wesentliches Element des deutschen Schulsystems. Der an dessen Stelle in der vorliegenden Untersuchung eingesetzte Vergleich der geplanten Schulabschlüsse entspricht dem der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS).

5 Abschließende Bewertung

5.1 International

Die Generierung international vergleichbarer Stichproben angesichts strukturell unterschiedlicher Schulsysteme ist ein nach wie vor ungelöstes Problem. Insbesondere ist das Verhältnis von Schul- zu Lebensalter ein wesentlicher Faktor, der den direkten Vergleich unterschiedlicher Systeme zu einer Frage der Perspektive macht. Die Ergebnisse von Leistungstests variieren in dem Maße, in dem der einen oder anderen Variablen mehr Bedeutung beigemessen wird.

Weitere Probleme betreffen die praktische Umsetzung einer weltweit angelegten Untersuchung.¹⁴⁵ Teilnehmerstaaten in Asien mit hohen Testwerten folgten nicht den Vorgaben der Stichprobenziehung. Ein europäischer Teilnehmerstaat weitete das Zeitfenster der Erhebung aus. Die in dieser Zeitspanne getesteten Schulen wurden nachträglich wegen auffällig hoher Performanz disqualifiziert. Es ist nicht auszuschließen, dass der besondere Leistungszuwachs dieses Teilnehmerstaats in der PISA 2003 Studie auf eine unterschiedliche Handhabung dieses Problems in der Folgestudie zurückzuführen ist. In einem weiteren Teilnehmerstaat wurden sämtliche Schulen eines Landesteils ausgeschlossen.

Die Auswahl der Schüler war bereits in der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS) Gegenstand der Diskussion. Dieser Sachverhalt gilt

¹⁴⁵ Vgl. Adams (ACER), Wu: PISA 2000 Technical Report. S. 182 ff

ebenfalls für die in PISA gezogenen Stichproben der Teilnehmerstaaten. Die Testergebnisse sind gleichermaßen vor diesem Hintergrund zu bewerten. Stichprobenspezifische Probleme hatte bereits Collani angesprochen.¹⁴⁶

Es hat sich gezeigt, wie sehr die Bildung von Aggregaten letztlich Zusammenhänge aufzeigen, die lediglich statistische Effekte darstellen, etwa Stage Migration als Ergebnis einer über die Bundesländer unterschiedlichen Bildungsexpansion.

Die statistische ‚Genauigkeit‘, beispielshalber in Form sehr kleiner Konfidenzintervalle, ein Resultat des Stichprobenumfangs von etwa 180.000 Schülern, mindert inhaltliche Präzision. In dem Maße, in dem der zu erklärende Gegenstand präziser definiert wird, reduzieren sich die Vergleichsgruppen. Unscharfe Definition des Untersuchungsgegenstands ist gewissermaßen der Preis großer Stichproben.¹⁴⁷

Es konnte erwartungsgemäß ein starker Effekt schulischer, nomineller Lernzeit auf die Testleistungen von Schülern nachgewiesen werden. Eine vergleichende Untersuchung, die Aussagen über Schulsysteme treffen will, muss daher zumindest die Verweildauer in dessen Institutionen berücksichtigen. Die Frage liegt deshalb nahe: Kann man unter dieser Voraussetzung überhaupt Schülerleistungen in verschiedenen Ländern vergleichen und sind dann gefundene Unterschiede nicht trivial?¹⁴⁸

Die Vergleiche über einzelne Klassenstufen sowie die kovarianzanalytischen Berechnungen zeigen einen geringen Grenzertrag besonders früher Einschulung in England oder Neuseeland. Die englischen Schülerinnen und Schüler der IGLU-Studie erreichen trotz höheren Schulalters mit $\mu=553$ ein Niveau, das nur knapp über der deutschen Vergleichsgruppe liegt ($\mu=539$). Dass in der internationalen Grundschulstudie (IGLU) die Mehrzahl der Schülerinnen und Schüler der Klassenstufe 4, in England, Neuseeland und Schottland jedoch die Klassenstufe 5 getestet wurde, ist zu bemerken. Die englischen, australischen und neuseeländischen Schülerinnen und Schüler erreichen in PISA unter Berücksichtigung des hohen Schulalters lediglich ein niedriges Niveau.

¹⁴⁶ Collani, E.v.: OECD PISA – An Example of Stochastic Illiteracy?, Economic Quality Control Vol 16 (2001), No.2, S. 240

¹⁴⁷ Wößmann berichtet vom Poolen zweier TIMSS-Datenstätze zu einem Datensatz von $N = 447.089$. Wößmann, L.: How Central Exams Affect Educational Achievement: International Evidence from TIMSS and TIMSS-Repeat S. 19 Paper prepared for the conference Taking Account of Accountability: Assessing Politics and Policy John F. Kennedy School of Government Harvard University June 10 – 11, 2002

¹⁴⁸ Baumert, Bos, Watermann: Mathematische und naturwissenschaftliche Grundbildung im internationalen Vergleich. In: Baumert, Bos, Lehmann (Hrsg.): TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie 2000, Band 1. S. 192

Nachdem in vorliegender Arbeit ein Einfluss des zwischen den Teilnehmerstaaten variierenden Schulalters auf die Höhe der Testergebnisse nachgewiesen und damit die Einschätzungen von Mullis, Martin, Beaton, Gonzalez, Kelly und Smith¹⁴⁹ bestätigt werden konnten, stellt sich die Frage nach den Ursachen der hohen finnischen Testwerte angesichts später Einschulung und hoher Rückhaltequote (Retentivität).

Das finnische Schulsystem wird in der fachwissenschaftlichen Diskussion sowohl als hochgradig integriert als auch im Gegensatz dazu als hochgradig selektiv dargestellt, wobei die Darstellung der hohen Testwerte als Folge eines integrierten Systems mit später Einschulung und hoher Rückhaltequote nicht kompatibel zu den vorliegenden Daten wäre.^{150 151} Eine Stratifikation des Schulsystems hätte Konsequenzen für die Stichprobenziehung und infolgedessen für die Testergebnisse. Diese würde für das im Test eingesetzte Stichprobenverfahren erhebliche, kaum zu lösende Probleme aufwerfen, insofern die Schichten nicht explizit ausgewiesen sind. Eine genauere Prüfung dieser Frage bleibt weiteren Untersuchungen vorbehalten. Es liegt auch in Hinblick auf vorliegende Untersuchungsergebnisse nahe, den je nach Teilnehmerstaat erheblich unterschiedlichen Freiheitsgraden der Stichprobenziehung insbesondere der Anzahl an Stratifikationsvariablen einen Einfluss auf die Testergebnisse einzuräumen.

¹⁴⁹ Martin, M.O., Beaton, A.E., Gonzalez, E.J., Kelly, D.L., Smith, T.A.: Mathematics Achievement in the Primary School Years (TIMSS), Boston College 1997, S. 30

¹⁵⁰ Linnakylä, Pirjo, Arbeitsgruppe „Internationale Vergleichsstudie“: Berlin Juni 2003 Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. S. 40

¹⁵¹ von Freyemann: PISA in Finnland. In: MUT 09/2003. S. 74 ff

5.2 National (Bundesrepublik)

Der Vergleich innerhalb Deutschlands war nicht Schwerpunkt der hier durchgeführten Untersuchung. Die Testleistungen der Bundesländer ergeben sich aus der unterschiedlichen Zusammensetzung der Stichproben sowie unterschiedlicher Bildungsexpansion. Hier kommen statistische Effekte zum Tragen.¹⁵²

Der Vergleich zwischen Gesamtschule und gegliedertem Schulsystem zeigt etwas höhere Korrelationen von sozialer Herkunft und Testleistung im gegliederten System. Insgesamt haben sich die in PISA berichteten hohen *Disparitäten* jedoch lediglich als statistischer Effekt erwiesen.

Es ist die Eigenschaft einer Vielzahl von Bildungssystemen, u.a. auch des deutschen, in Anerkennung unterschiedlicher Entwicklungsniveaus, Schul- und Lebensalter in engen Grenzen individuell zu differenzieren. Diese Flexibilität schafft zwar eine größere Streuung der Testergebnisse bei einer am Lebensalter orientierten Perspektive, jedoch eine verhältnismäßig geringe Streuung bei einer am Schulalter orientierten Perspektive, wie die Autoren der IGLU-Studie, Bos, Valtin, Lankes, Schwippert, Voss, Badel und Plaßmeier bestätigen.

*Für Deutschland insgesamt gilt: Die Streuung der Leistungswerte ist am Ende der vierten Jahrgangsstufe klein. Nur wenige andere Staaten erreichen eine geringere Streuung und übergeben somit eine in ihren Leseleistungen insgesamt homogenere Schülerschaft an nachfolgende Klassen.*¹⁵³

¹⁵² Fertig und Wright beschreiben statistische Effekte, die im Rahmen von Auswertungen von PISA-Daten auftreten und lediglich verschiedenen Aggregatniveaus zuzuschreiben sind (Aggregation Bias).

As the level of aggregation at which class size is measured increases, the effect changes from being small, positive and statistically significant (elasticity=0.04) to being small, negative and statistically significant (elasticity =-0.07).

Fertig, Wright: School Quality, Educational Attainment and Aggregation Bias, IZA DP No. 994, Januar 2004

¹⁵³ Bos, Valtin, Lankes, Schwippert, Voss, Badel, Plaßmeier: Lesekompetenzen am Ende der vierten Jahrgangsstufe in einigen Ländern der Bundesrepublik Deutschland im nationalen Vergleich. In: Bos, Lankes, Prenzel, Schwippert, Valtin, Walther (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II) 2004. S. 85

5.3 Weiterführende Perspektiven

5.3.1 Testkonstruktion

Es wäre sicherlich von Nutzen, die in PISA eingesetzten Aufgaben zu betrachten, sobald diese der Öffentlichkeit zugänglich werden. Die individuellen Testwerte der Schüler in Mathematik und Naturwissenschaften, kovariieren mit dem Lesetestwert mit $r_p = 0,839$ und $r_p = 0,872$, wodurch sich die Frage der Abgrenzbarkeit stellt.¹⁵⁴

Die bislang veröffentlichten Testaufgaben werfen meßtheoretische Fragen auf: Neben dem hohen Sprachanteil des Bereichs *Scientific Literacy*¹⁵⁵ wären insbesondere die Präzision der Fragestellungen, Kongruenz zu wissenschaftlichen Erklärungen,¹⁵⁶ Eindeutigkeit des Explanandums und eine Reihe weiterer Probleme anzusprechen. Messener hatte in diesem Zusammenhang nach dem Gegenstand der Untersuchung gefragt und auf die Praxis der Testentwicklung hingewiesen.¹⁵⁷ Hagemeyer hatte diese Problematik bereits in Bezug auf die TIMSS angesprochen.¹⁵⁸ Eine breite wissenschaftliche Öffentlichkeit mit Zugang zu den Testaufgaben scheint die dringend notwendige Voraussetzung valider Testkonstruktion zu sein.¹⁵⁹

¹⁵⁴ vgl. dazu auch die Diskussion um die in der TIMSS eingesetzten Items von Hagemeyer. Hagemeyer, V.: Was wurde bei TIMSS erhoben? Die Deutsche Schule, 91.Jg.1999, S.160-177

¹⁵⁵ Baumert, Klieme, Lehrke, Savelsbergh hatten in Bezug auf das Argument der Sprachlastigkeit der TIMSS (Hagemeyer) hingegen die hohe Mathematiklastigkeit des Physikunterrichts angesprochen. (Jürgen Baumert, Eckhard Klieme, Manfred Lehrke & Elwin Savelsbergh: Konzeption und Aussagekraft der TIMSS-Leistungstests. Zur Diskussion um TIMSS-Aufgaben aus der Mittelstufenphysik S. 19ff)

¹⁵⁶ Als Beispiel aus der PISA-Studie: S129Q01 (DAYLIGHT/TAGESLICHT) Frage: *Welche Aussage erklärt, warum es auf der Erde Tageslicht und Dunkelheit gibt?* Antwort A (–als *vollständig gelöst* bewertete Antwort- *Die Erde rotiert um ihre Achse*) ist notwendige (jedoch nicht hinreichende) Voraussetzung für den Wechsel von Tag und Nacht. Antwort D (–als *nicht gelöst* bewertete Antwort- *Die Erde dreht sich um die Sonne*) ist ebenfalls ein Teil der Voraussetzung, da nicht die Erdrotation allein ursächlich für den Wechsel von Tag und Nacht ist, sondern das Verhältnis von Erdrotation (um ihre Achse) und der Rotation der Erde um die Sonne. Nikolas Knake (---.dip.t-dialin.net) hat folgende Interpretation für die Erklärung, *warum es auf der Erde Tageslicht und Dunkelheit gibt*, die weder Erdrotation noch Rotation der Erde um die Sonne notwendig macht (und der Fragestellung möglicherweise näher kommt): *Tageslicht gibt es nämlich weil die Sonne elektromagnetische Strahlen eines bestimmten Frequenzbereiches emittiert, die wir Menschen als Tageslicht bezeichnen* (http://www.skh.de/phorum/read.php?f=8&i=70&t=70#reply_70).

¹⁵⁷ Messener, R.: PISA und Allgemeinbildung In: Zeitschrift für Pädagogik 3/2003, S. 402

¹⁵⁸ Hagemeyer: Was wurde bei TIMSS erhoben? Die Deutsche Schule, 1999

¹⁵⁹ Baumert, Klieme, Lehrke, Savelsbergh: Konzeption und Aussagekraft der TIMSS-Leistungstests. MPIB 1999

5.3.2 Curriculare Validität

Curriculare Validität wäre ein weiterer Gegenstand von wissenschaftlichem Interesse. In der Bundesrepublik Deutschland beziehen sich schulische Curricula vorwiegend auf Klassen-, jedoch kaum auf Altersstufen. Curriculare Validität im Rahmen einer Altersstichprobe von 15-Jährigen würde sich in der Bundesrepublik Deutschland unter Bezug auf die PISA-Daten gleichermaßen auf 1,2% Schüler der Klasse 7, 13,2% Schüler der Klasse 8, 63,3% Schüler der Klasse 9, 22,3% Schüler der Klasse 10 und 0,1% Schüler der Klasse 11 beziehen. Dieses Problem stellt sich in ähnlicher Weise für mehr als zwei Drittel der teilnehmenden Staaten und verweist auf den explizit normativen Charakter der Studien.¹⁶⁰

5.4 Resümee

PISA wirft grundlegende Fragen auf, die über den wissenschaftlichen Rahmen hinaus gehen. Ergebnisse international vergleichender Bildungsstudien sollten nicht ungeprüft übernommen werden. Eine sachgerechte Behandlung der Thematik wäre zu wünschen.

¹⁶⁰ Baumert, Stanat, Demmrich: PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. S.19

6 Anhang Tabellen

PISA 2000 Leseverständnis (Datensatz gewichtet)

COUNTRY	pv1read8	pv1read9	pv1read10	pv1read11
ALBANIA*)	243,17	312,18	367,93	368,70
ARGENTINA*)	296,06	336,53	458,32	429,94
AUSTRALIA	256,75	454,51	526,41	565,34
AUSTRIA	392,05	497,12	532,22	603,51
BELGIUM	346,99	433,42	552,63	603,23
BRAZIL	368,94	424,17	462,13	
BULGARIA*)	339,02	433,83	443,54	462,13
CANADA	435,38	486,08	546,73	577,11
CHILE*)	292,41	389,68	441,86	481,21
CZECH REPUBLIC	385,45	492,90	513,93	
DENMARK	426,52	501,75	535,74	
FINLAND	492,94	554,03		
FRANCE	397,40	449,02	552,93	609,16
GERMANY	391,62	489,08	538,68	612,88
GREECE	349,84	391,23	478,84	508,10
HONG KONG*)	479,90	479,77	533,33	558,90
HUNGARY	391,61	476,99	511,73	
ICELAND			508,03	
INDONESIA*)	310,12	345,59	398,49	421,15
IRELAND	407,32	517,08	568,76	547,81
ISRAEL*)	247,90	396,86	454,89	479,76
ITALY	348,49	423,29	501,36	523,82
JAPAN			522,03	
KOREA, REPUBLIC OF		502,17	525,28	518,57
LATVIA	386,65	444,24	489,67	535,47
LIECHTENSTEIN	422,93	496,40	595,80	
LUXEMBOURG	385,68	432,13	508,50	
MEXICO	346,83	389,54	466,83	451,12
NETHERLANDS	430,73	503,65	570,02	675,44
NEW ZEALAND			436,20	533,41
NORWAY	326,16	407,02	507,06	561,44
PERU*)	259,02	298,75	374,54	387,29
POLAND		480,01		
PORTUGAL	373,12	438,24	531,85	609,90
ROMANIA*)	355,25	447,37	447,63	
RUSSIAN FEDERATION	365,31	445,87	470,64	557,15
SPAIN	366,69	424,09	521,69	702,99
SWEDEN	398,89	518,55	557,23	
SWITZERLAND	428,05	506,86	549,71	665,18
THAILAND*)	382,31	421,88	447,07	482,64
MACEDONIA*)	273,23	367,19	399,71	
UNITED KINGDOM		437,13	519,81	524,67
UNITED STATES	352,06	478,49	531,02	558,07

*) Der Datensatz enthält Teilnehmerstaaten, die in der ursprünglichen Studie nicht aufgeführt waren.

**) Download vom 9-4-2005 http://pisaweb.acer.edu.au/oeecd/oeecd_pisa_data_s1.html

**) Zum Teil sehr geringe Zellbesetzungen

****) Lesetestwert berechnet entsprechend Empfehlung vom MPIB über pv1read

pv1read8 Klassenstufe 8
pv1read9 Klassenstufe 9
pv1read10 Klassenstufe 10
pv1read11 Klassenstufe 11

Tabelle 1 PISA 2000 Leseverständnis (Datensatz gewichtet)

PISA 2000 Leseverständnis (Datensatz ungewichtet)

COUNTRY	pv1read8	pv1read9	pv1read10	pv1read11
ALBANIA*)	243,17	310,72	367,32	375,96
ARGENTINA*)	306,73	354,79	457,69	433,74
AUSTRALIA	250,05	443,61	525,08	566,06
AUSTRIA	381,61	494,17	515,31	555,36
BELGIUM	348,61	440,16	554,59	606,04
BRAZIL	362,23	430,47	465,94	
BULGARIA*)	359,26	425,30	436,30	466,85
CANADA	412,40	473,54	538,03	587,77
CHILE*)	295,89	391,45	442,84	479,20
CZECH REPUBLIC	386,24	494,97	514,48	
DENMARK	428,34	503,16	532,65	
FINLAND	498,76	554,88		
FRANCE	397,80	449,19	553,47	609,72
GERMANY	412,66	499,86	553,76	613,73
GREECE	349,03	389,86	470,70	500,54
HONG KONG*)	481,43	480,15	533,91	559,05
HUNGARY	392,72	472,90	506,53	
ICELAND			508,60	
INDONESIA*)	312,50	348,86	386,04	402,50
IRELAND	408,15	518,08	568,86	548,44
ISRAEL*)	247,90	422,14	464,82	481,70
ITALY	293,05	421,96	502,46	525,01
JAPAN			525,25	
KOREA, REPUBLIC OF		499,17	521,00	523,30
LATVIA	390,62	447,92	484,98	519,65
LIECHTENSTEIN	423,52	496,14	596,59	
LUXEMBOURG	389,77	434,83	509,61	
MEXICO	347,59	389,13	469,30	458,80
NETHERLANDS	437,27	515,98	578,26	678,94
NEW ZEALAND			436,13	532,03
NORWAY	327,23	409,29	506,98	548,52
PERU*)	261,69	302,46	380,68	395,28
POLAND		468,26		
PORTUGAL	375,57	439,68	531,64	613,00
ROMANIA*)	375,12	455,58	459,04	
RUSSIAN FEDERATION	371,60	445,77	468,97	556,11
SPAIN	365,12	422,53	521,90	702,99
SWEDEN	400,84	518,93	518,12	
SWITZERLAND	428,44	505,50	546,34	668,69
THAILAND*)	383,81	424,16	462,85	516,51
MACEDONIA*)	290,40	366,14	398,74	
UNITED KINGDOM		514,73	513,55	527,11
UNITED STATES	342,49	464,88	522,80	535,28

*) Der Datensatz enthält Teilnehmerstaaten, die in der ursprünglichen Studie nicht aufgeführt waren.

**) Download vom 9-4-2005 http://pisaweb.acer.edu.au/oeecd/oeecd_pisa_data_s1.html

***) Zum Teil sehr geringe Zellbesetzungen

****) Lesetestwert berechnet entsprechend Empfehlung vom MPIB über pv1read

pv1read8	Klassenstufe 8
pv1read9	Klassenstufe 9
pv1read10	Klassenstufe 10
pv1read11	Klassenstufe 11

Tabelle 2 PISA 2000 Leseverständnis (Datensatz ungewichtet)

Country three-digit ISO code * Grade - Q2 Kreuztabelle

Anzahl		Grade - Q2										Gesamt
		4	5	6	7	8	9	10	11	12	13	
Country	ALBANIA	0	0	0	2	25	1211	3475	245	0	0	4958
three-digit	ARGENTINA	0	0	0	47	311	705	2883	37	0	0	3983
ISO code	AUSTRALIA	0	0	0	0	3	323	3944	870	0	0	5140
	AUSTRIA	0	0	1	15	217	1921	2408	3	0	0	4565
	BELGIUM	0	0	0	14	279	1741	4508	66	1	0	6609
	BRAZIL	0	0	0	1149	1662	1664	345	0	0	0	4820
	BULGARIA	0	0	0	18	117	3828	672	14	0	0	4649
	CANADA	0	0	4	57	506	4757	23339	356	6	1	29026
	CHILE	0	0	0	80	338	1357	3098	11	0	0	4884
	CZECH REPUBLIC	0	0	3	16	91	1982	3091	0	0	0	5183
	DENMARK	0	0	0	0	249	3780	91	0	0	0	4120
	FINLAND	0	0	0	10	523	4331	0	0	0	0	4864
	FRANCE	0	0	0	16	347	1760	2403	122	0	0	4648
	GERMANY	0	0	1	58	658	3163	1114	3	0	0	4997
	GREECE	0	0	0	4	35	92	3661	794	0	0	4586
	HONG KONG	0	0	0	135	280	524	2695	767	4	0	4405
	HUNGARY	0	0	12	49	100	2861	1818	0	0	0	4840
	ICELAND	0	0	0	0	0	0	3317	0	0	0	3317
	INDONESIA	0	0	0	237	941	3516	2030	627	17	0	7368
	IRELAND	0	0	0	1	126	2359	605	737	1	0	3829
	ISRAEL	0	0	0	0	1	399	3899	176	1	0	4476
	ITALY	0	0	0	24	29	791	3825	283	0	0	4952
	JAPAN	0	0	0	0	0	0	5256	0	0	0	5256
	KOREA, REPUBLIC OF	0	0	0	0	0	59	4888	29	0	0	4976
	LATVIA	1	5	2	60	270	1359	2129	20	0	0	3846
	LIECHTENSTEIN	0	0	0	3	50	244	11	0	0	0	308
	LUXEMBOURG	0	0	0	58	578	1956	890	0	0	0	3482
	MEXICO	0	0	0	121	447	1167	2553	9	0	0	4297
	NETHERLANDS	0	0	0	7	120	1125	1215	3	0	0	2470
	NEW ZEALAND	0	0	0	0	0	0	253	3202	175	3	3633
	NORWAY	0	0	0	0	2	29	4053	15	0	0	4099
	PERU	0	0	0	323	605	1064	2008	391	0	0	4391
	POLAND	0	0	0	0	0	3654	0	0	0	0	3654
	PORTUGAL	0	9	33	229	535	1141	2543	14	0	0	4504
	ROMANIA	0	0	0	22	345	3815	647	0	0	0	4829
	RUSSIAN FEDERATION	0	0	1	2	118	1670	4858	39	0	0	6688
	SPAIN	0	0	0	1	123	1515	4520	1	0	0	6160
	SWEDEN	0	0	0	1	86	4279	10	0	0	0	4376
	SWITZERLAND	0	0	0	41	842	4633	497	3	1	0	6017
	THAILAND	0	0	0	13	81	3279	1921	46	0	0	5340
	MACEDONIA	0	0	0	13	37	3150	1266	0	0	0	4466
	UNITED KINGDOM	0	0	0	0	0	4	3120	5832	349	0	9305
	UNITED STATES	0	0	0	4	71	1547	2207	16	0	0	3845
Gesamt		1	14	57	2830	11148	78755	118066	14731	555	4	226161

Tabelle 3 Verteilung Klassenstufe PISA 2000 (Public Use File Download 9-4-2005)

PISA 2003 Leseverständnis (Datensatz gewichtet)

COUNTRY	pv1read8	pv1read9	pv1read10	pv1read11
Australia	345,10	467,85	524,18	557,53
Austria	383,31	489,84	505,33	
Belgium	343,09	439,19	547,75	606,96
Brazil	353,50	431,01	471,74	486,41
Canada	424,85	486,44	541,03	569,03
Czech Republic	370,37	472,34	508,24	
Denmark	430,13	494,95	553,44	559,83
Finland	502,87	549,11		
France	382,80	437,89	539,01	592,82
Germany	403,68	494,24	556,98	636,40
Greece	382,63	393,44	477,57	490,07
Hong Kong (China)	468,07	498,72	527,66	596,43
Hungary	409,65	478,80	506,20	634,82
Iceland			491,73	
Indonesia	337,52	370,01	413,62	423,96
Ireland	406,33	503,30	563,96	529,50
Italy	343,20	409,53	487,98	501,61
Japan			499,04	
Korea		524,82	534,80	601,04
Latvia	448,44	498,64	537,97	548,56
Liechtenstein	465,86	536,20	601,71	
Luxembourg	426,50	461,43	537,27	605,75
Macao (China)	469,34	512,39	540,45	559,79
Mexico	330,65	383,81	439,04	479,73
Netherlands	429,63	486,17	546,61	627,28
New Zealand	421,21	446,40	525,24	583,63
Norway		408,94	500,00	535,27
Poland	365,47	501,74	579,49	
Portugal	374,93	429,18	518,94	584,64
Russian Federation	361,90	422,74	453,61	496,26
Slovakia	330,44	463,31	475,69	605,45
Spain	347,74	422,61	510,64	515,82
Sweden	390,03	514,74	570,62	
Switzerland	433,23	506,73	533,90	597,57
Thailand	345,71	397,98	437,30	509,90
Tunisia	331,49	370,57	438,25	465,25
Turkey	332,97	442,86	447,36	449,52
United Kingdom		590,82	494,71	512,02
United States	381,28	468,57	509,35	525,84
Uruguay	337,20	376,77	473,80	502,50
Yugoslavia		410,21	456,55	

**) Download vom 9-4-2005 http://pisaweb.acer.edu.au/oeed_2003/oeed_pisa_data_s1.html

***) Zum Teil sehr geringe Zellbesetzungen

****) Lesetestwert berechnet entsprechend Empfehlung vom MPIB über pv1read

pv1read8	Klassenstufe 8
pv1read9	Klassenstufe 9
pv1read10	Klassenstufe 10
pv1read11	Klassenstufe 11

Tabelle 4 PISA 2003 Leseverständnis (Datensatz gewichtet)

PISA 2003 Leseverständnis (Datensatz ungewichtet)

COUNTRY	pv1read8	pv1read9	pv1read10	pv1read11
Australia	346,99	472,52	522,35	553,90
Austria	386,99	491,25	511,81	
Belgium	362,67	443,85	550,01	605,76
Brazil	358,52	428,18	467,74	483,41
Canada	412,56	468,74	527,90	554,94
Czech Republic	395,01	490,59	521,80	
Denmark	431,89	494,51	553,48	559,83
Finland	501,71	547,14		
France	386,93	438,37	539,59	593,75
Germany	410,18	499,47	559,53	641,81
Greece	386,15	393,14	473,60	488,40
Hong Kong (China)	474,92	501,15	530,31	595,47
Hungary	418,56	479,05	505,63	634,82
Iceland			491,71	
Indonesia	337,39	374,88	408,12	428,54
Ireland	407,12	505,05	564,13	532,62
Italy	367,05	435,65	512,91	520,54
Japan			498,27	
Korea		527,55	533,46	604,25
Latvia	450,89	500,17	541,08	543,21
Liechtenstein	466,53	536,56	601,64	
Luxembourg	426,69	462,36	537,22	606,04
Macao (China)	468,96	511,52	544,85	566,91
Mexico	334,52	384,28	438,60	477,48
Netherlands	424,33	489,00	550,67	626,14
New Zealand	399,96	447,55	526,82	583,29
Norway		405,60	500,09	535,27
Poland	362,69	500,47	579,67	
Portugal	376,10	430,03	518,47	586,59
Russian Federation	366,83	425,22	457,73	506,49
Slovakia	329,76	467,54	482,45	605,41
Spain	362,94	425,43	515,91	515,82
Sweden	390,22	515,25	572,54	
Switzerland	432,16	496,71	544,22	599,07
Thailand	345,99	400,24	444,50	518,74
Tunisia	331,43	370,58	438,81	466,04
Turkey	333,55	440,27	445,14	450,76
United Kingdom		590,82	496,78	510,52
United States	373,67	466,80	507,98	523,27
Uruguay	334,00	376,92	476,76	518,62
Yugoslavia		409,39	456,56	

**) Download <http://pisaweb.acer.edu.au/oeed>

***) Zum Teil sehr geringe Zellbesetzungen

****) Lesetestwert berechnet entsprechend Empfehlung vom MPIB über pv1read

pv1read8 Klassenstufe 8
pv1read9 Klassenstufe 9
pv1read10 Klassenstufe 10
pv1read11 Klassenstufe 11

Tabelle 5 PISA 2003 Leseverständnis (Datensatz ungewichtet)

Country ID * Grade Q1a Kreuztabelle

Anzahl		Grade Q1a						Gesamt
		7	8	9	10	11	12	
Country ID	Australia	1	16	1064	8982	2480	8	12551
	Austria	8	166	2034	2383	0	0	4591
	Belgium	24	333	2548	5797	66	0	8768
	Brazil	527	973	2062	867	23	0	4452
	Canada	62	397	4343	22080	330	2	27214
	Czech Republic	7	122	2785	3406	0	0	6320
	Denmark	3	377	3731	103	2	0	4216
	Finland	14	715	5067	0	0	0	5796
	France	7	213	1392	2586	100	2	4300
	Germany	73	647	2706	1088	6	0	4520
	Greece	12	121	342	3460	692	0	4627
	Hong Kong (China)	211	439	1132	2692	4	0	4478
	Hungary	64	330	3022	1348	1	0	4765
	Iceland	0	0	0	3350	0	0	3350
	Indonesia	248	1188	4948	4124	238	15	10761
	Ireland	1	109	2362	654	754	0	3880
	Italy	11	63	1775	9562	228	0	11639
	Japan	0	0	0	4707	0	0	4707
	Korea	0	0	73	5366	5	0	5444
	Latvia	55	723	3459	297	5	0	4539
	Liechtenstein	2	67	237	26	0	0	332
	Luxembourg	0	583	2195	1141	4	0	3923
	Macao (China)	191	355	426	275	3	0	1250
	Mexico	466	1350	4833	22696	104	2	29451
	Netherlands	4	176	1762	2029	19	1	3991
	New Zealand	0	2	300	4019	181	1	4503
	Norway	0	0	25	4026	13	0	4064
	Poland	31	135	4196	21	0	0	4383
	Portugal	208	508	963	2900	29	0	4608
	Russian Federation	17	129	1657	4101	70	0	5974
	Slovakia	31	75	2672	4530	38	0	7346
	Spain	4	259	2576	7951	1	0	10791
	Sweden	2	114	4420	88	0	0	4624
	Switzerland	59	1359	5910	1067	12	0	8407
	Thailand	4	44	2222	2864	102	0	5236
	Tunisia	721	1033	1181	1645	141	0	4721
	Turkey	27	92	191	2863	1670	12	4855
	United Kingdom	0	0	1	1495	6384	1655	9535
	United States	8	95	1667	3339	345	0	5454
	Uruguay	517	667	1294	3035	322	0	5835
	Yugoslavia	0	0	4301	104	0	0	4405
	Gesamt	3620	13975	87874	153067	14372	1698	274606

Tabelle 6 PISA 2003 Verteilung Klassenstufe

PISA 2003 Mathematiktest (Datensatz gewichtet)

COUNTRY	pv1math8	pv1math9	pv1math10	pv1math11
Australia	360,39	464,51	521,60	560,14
Austria	405,32	502,92	518,71	
Belgium	366,51	457,35	571,05	638,98
Brazil	303,06	382,84	423,33	443,46
Canada	431,84	493,20	545,91	583,90
Czech Republic	397,05	501,92	534,95	
Denmark	450,65	518,51	556,94	575,67
Finland	498,93	551,14		
France	399,13	454,03	553,40	612,69
Germany	418,33	505,30	566,68	650,98
Greece	355,63	378,34	449,51	462,72
Hong Kong (China)	490,90	538,51	574,09	629,67
Hungary	402,69	486,12	520,42	661,50
Iceland			514,71	
Indonesia	314,97	348,18	395,26	410,46
Ireland	405,39	492,54	545,12	516,07
Italy	329,43	406,76	477,95	486,67
Japan			533,64	
Korea		534,04	541,70	594,49
Latvia	430,31	491,12	536,06	562,74
Liechtenstein	464,42	546,84	630,87	
Luxembourg	442,23	474,74	553,79	703,24
Macao (China)	488,28	546,95	581,42	569,97
Mexico	325,13	369,78	420,47	457,87
Netherlands	440,08	506,96	574,84	671,25
New Zealand	334,25	455,42	526,94	581,01
Norway		441,99	495,25	557,25
Poland	380,03	494,53	578,89	
Portugal	373,77	417,05	504,93	589,75
Russian Federation	380,80	450,16	480,44	526,00
Slovakia	372,17	489,63	506,81	640,84
Spain	351,42	429,23	513,52	417,31
Sweden	403,42	509,57	565,46	
Switzerland	448,47	535,07	569,86	644,71
Thailand	348,92	395,13	434,23	521,97
Tunisia	317,76	348,19	419,99	443,96
Turkey	320,51	418,04	428,01	432,97
United Kingdom		505,09	496,62	512,98
United States	385,14	458,65	497,34	508,74
Uruguay	330,80	367,94	456,92	489,24
Yugoslavia		435,09	495,04	

**) Download vom 9-4-2005 http://pisaweb.acer.edu.au/oced_2003/oced_pisa_data_s1.html

***) Zum Teil sehr geringe Zellbesetzungen

****) Testwert berechnet entsprechend Empfehlung vom MPIB

pv1math8 Klassenstufe 8
pv1math9 Klassenstufe 9
pv1math10 Klassenstufe 10
pv1math11 Klassenstufe 11

Tabelle 7 PISA 2003 Mathematiktest (Datensatz gewichtet)

PISA 2003 Mathematiktest (Datensatz ungew.)

COUNTRY	pv1math8	pv1math9	pv1math10	pv1math11
Australia	349,08	469,77	519,58	556,01
Austria	408,40	504,97	525,64	
Belgium	384,67	462,34	572,86	640,41
Brazil	307,80	380,37	419,78	435,00
Canada	421,82	473,61	533,44	573,89
Czech Republic	413,30	520,15	551,55	
Denmark	450,49	518,37	559,02	575,67
Finland	496,79	549,43		
France	403,23	454,34	553,79	613,94
Germany	424,12	510,14	568,70	652,04
Greece	356,61	374,97	446,28	461,81
Hong Kong (China)	499,00	542,18	577,38	630,85
Hungary	414,21	486,40	519,84	661,50
Iceland			514,67	
Indonesia	314,57	350,63	388,58	417,28
Ireland	405,95	494,08	546,27	519,71
Italy	375,75	435,20	508,01	508,62
Japan			532,98	
Korea		537,00	540,06	595,47
Latvia	436,88	493,59	539,50	556,80
Liechtenstein	465,58	547,17	630,69	
Luxembourg	442,44	475,57	553,49	703,78
Macao (China)	487,19	548,15	585,23	554,35
Mexico	325,18	367,33	420,61	456,03
Netherlands	435,50	510,42	578,82	669,19
New Zealand	351,84	458,19	528,67	581,75
Norway		439,71	496,07	557,25
Poland	377,81	493,44	578,17	
Portugal	375,99	420,02	504,76	594,52
Russian Federation	380,07	451,61	484,09	534,72
Slovakia	365,18	493,23	513,36	640,07
Spain	371,03	434,06	519,04	417,31
Sweden	404,17	510,05	564,62	
Switzerland	450,40	524,52	575,47	644,97
Thailand	349,00	396,82	440,61	533,21
Tunisia	317,50	348,13	420,76	444,08
Turkey	313,44	422,91	427,03	435,45
United Kingdom		505,09	498,45	514,24
United States	381,85	456,72	495,34	505,34
Uruguay	328,23	369,40	460,76	490,10
Yugoslavia		434,53	495,82	

**) Download vom 9-4-2005 http://pisaweb.acer.edu.au/oeed_2003/oeed_pisa_data_s1.html

***) Zum Teil sehr geringe Zellbesetzungen

****) Testwert berechnet entsprechend Empfehlung vom MPIB

pv1math8 Klassenstufe 8
pv1math9 Klassenstufe 9
pv1math10 Klassenstufe 10
pv1math11 Klassenstufe 11

Tabelle 8 PISA 2003 Mathematiktest (Datensatz ungewichtet)

Differenz der mittleren Testwerte *** (PISA 2003 - PISA 2000 Lesetest)

COUNTRY	PISA 2000	PISA 2000	PISA 2003	PISA 2003	Differenz *)	Differenz **)
	(gewichtet)	(ungewichtet)	(gewichtet)	(ungewichtet)	(gewichtet)	(ungewichtet)
ALBANIA	349,02	353,21				
ARGENTINA	417,41	425,15				
AUSTRALIA	527,79	526,23	525,67	524,14	-2,1	-2,1
AUSTRIA	507,53	497,72	490,91	497,36	-16,6	-0,4
BELGIUM	507,31	515,05	506,99	511,15	-0,3	-3,9
BRAZIL	395,20	381,66	403,51	407,51	8,3	25,9
BULGARIA	430,54	424,28				
CANADA	534,49	524,21	527,65	516,09	-6,8	-8,1
CHILE	409,64	415,55				
CZECH REPUBLIC	491,94	497,51	488,04	505,40	-3,9	7,9
DENMARK	496,56	497,92	491,21	490,26	-5,4	-7,7
FINLAND	546,99	548,52	542,90	541,12	-4,1	-7,4
FRANCE	504,38	502,81	496,07	500,13	-8,3	-2,7
GERMANY	483,99	497,68	491,70	497,53	7,7	-0,1
GREECE	474,46	472,13	471,58	467,20	-2,9	-4,9
HONG KONG	525,15	525,82	509,21	513,60	-15,9	-12,2
HUNGARY	479,83	481,35	481,87	480,70	2,0	-0,6
ICELAND	506,58	507,15	491,73	491,71	-14,9	-15,4
INDONESIA	370,76	357,26	380,62	383,54	9,9	26,3
IRELAND	526,68	527,74	515,82	517,55	-10,9	-10,2
ISRAEL	450,96	460,75				
ITALY	487,36	488,87	474,94	500,26	-12,4	11,4
JAPAN	522,03	525,25	499,04	498,27	-23,0	-27,0
KOREA, REPUBLIC OF	525,03	520,75	534,71	533,44	9,7	12,7
LATVIA	459,14	461,82	490,88	492,73	31,7	30,9
LIECHTENSTEIN	482,58	482,65	526,63	527,21	44,1	44,6
LUXEMBOURG	440,40	443,67	478,58	478,98	38,2	35,3
Macao (China)			498,26	493,84		
MEXICO	421,74	428,33	399,53	422,00	-22,2	-6,3
NETHERLANDS	531,64	542,45	513,96	518,02	-17,7	-24,4
NEW ZEALAND	528,60	527,29	521,99	523,73	-6,6	-3,6
NORWAY	505,71	505,37	499,68	499,62	-6,0	-5,8
PERU	327,23	334,90				
POLAND	480,01	468,26	496,48	495,16	16,5	26,9
PORTUGAL	469,89	476,53	477,76	476,15	7,9	-0,4
ROMANIA	427,19	449,70				
RUSSIAN FEDERATION	461,91	461,67	442,30	446,89	-19,6	-14,8
Slovakia			469,24	475,22		
SPAIN	492,84	493,60	481,68	490,64	-11,2	-3,0
SWEDEN	515,39	515,86	514,32	513,18	-1,1	-2,7
SWITZERLAND	494,80	496,02	498,61	491,54	3,8	-4,5
THAILAND	431,18	438,02	419,91	426,29	-11,3	-11,7
Tunisia			374,44	375,08		
Turkey			441,68	443,97		
MACEDONIA	372,96	373,61				
UNITED KINGDOM	523,54	523,98	506,81	512,05	-16,7	-11,9
UNITED STATES	503,71	495,97	494,87	493,77	-8,8	-2,2
Uruguay			434,98	424,49		
Yugoslavia			411,33	410,51		

*) Differenz der gewichteten Testmittelwerte eines Teilnehmerstaates 2003-2000

**) Differenz der ungewichteten Testmittelwerte eines Teilnehmerstaates 2003-2000

***) Berechnungen über Public-Use-Files (pv1read) entsprechend den Empfehlungen des MPIB

Tabelle 9 PISA2000/PISA2003 Lesetest: Vergleiche der gewichteten und ungewichteten Ergebnisse

Lange Reihen

Tab. L 204: Familien mit Kindern in der Familie nach Familienstand der Bezugsperson und Altersgruppen der Kinder*)
1 000

Zeitpunkt	Ins-gesamt	Ehepaare mit in der Familie lebenden ledigen Kindern	Allein Erziehende										
			Männer					Frauen					
			zu-sammen	zu-sammen	davon			zu-sammen	zu-sammen	davon			
					ledig	ver-heiratet getrennt lebend	ver-witwet			ge-schieden	ledig	ver-heiratet getrennt lebend	ver-witwet
Insgesamt													
Früheres Bundesgebiet													
6.6. 1961 1).....	10 921	8 871	2 047	204	2	13	165	26	1 843	181	117	1 265	281
27.5. 1970 1).....	10 938	9 376	1 562	231	10	61	122	39	1 331	130	102	811	287
Mai 1975	11 022	9 577	1 445	189	/	26	114	43	1 256	115	64	724	353
Mai 1976	10 889	9 431	1 457	205	7	26	118	55	1 252	116	69	695	371
April 1977	10 915	9 445	1 471	203	6	25	116	57	1 268	121	65	689	392
April 1978	10 879	9 363	1 516	211	10	33	115	53	1 305	125	95	691	394
April 1979	10 816	9 278	1 538	228	11	40	118	59	1 310	128	107	674	402
April 1980	10 861	9 295	1 566	249	11	46	126	66	1 317	129	124	660	404
April 1982	10 852	9 193	1 658	270	14	58	128	71	1 388	148	146	646	448
April 1986	10 480	8 649	1 831	283	25	45	123	90	1 548	207	159	645	537
März 1987	10 482	8 618	1 864	292	25	45	123	99	1 571	219	161	637	555
April 1988	10 457	8 592	1 865	301	29	43	126	103	1 565	236	158	611	559
April 1989	10 458	8 619	1 839	284	30	41	114	99	1 556	239	147	608	562
April 1990	10 600	8 778	1 822	291	31	44	112	104	1 532	256	157	560	558
April 1991	10 668	8 811	1 858	306	34	48	112	112	1 552	275	163	545	569
Mai 1992	10 726	8 822	1 904	305	36	47	113	109	1 599	297	176	543	583
April 1993	10 748	8 799	1 949	327	46	52	115	113	1 623	322	185	531	585
April 1994	10 665	8 684	1 982	332	52	52	111	116	1 650	324	198	529	600
April 1995	10 586	8 581	2 005	358	62	52	115	128	1 647	331	202	501	614
April 1996	10 466	8 442	2 024	387	82	62	108	135	1 637	342	217	464	614
April 1997	10 431	8 383	2 049	403	90	62	106	145	1 645	356	222	441	626
April 1998	10 347	8 283	2 064	405	85	62	109	149	1 659	360	236	423	639
April 1999	10 341	8 205	2 137	400	93	59	104	144	1 737	382	240	434	682
Mai 2000	10 314	8 140	2 174	410	99	60	104	147	1 764	395	251	417	701
April 2001	10 255	8 019	2 236	438	113	61	104	160	1 798	412	254	407	725
Deutschland													
April 1991	13 639	11 098	2 540	394	60	52	136	146	2 146	479	179	652	835
Mai 1992	13 618	11 032	2 585	393	64	51	138	140	2 192	507	198	653	835
April 1993	13 596	10 948	2 647	420	76	59	143	142	2 228	541	214	643	831
April 1994	13 469	10 773	2 696	417	71	61	138	147	2 279	549	234	640	856
April 1995	13 355	10 619	2 736	451	84	61	143	162	2 286	560	250	610	866
April 1996	13 181	10 408	2 772	493	112	76	133	171	2 279	576	268	574	862
April 1997	13 096	10 299	2 797	509	124	74	128	182	2 288	593	277	548	871
April 1998	12 959	10 135	2 824	514	120	75	130	189	2 311	607	295	529	879
April 1999	12 913	9 987	2 926	512	131	73	126	183	2 414	645	303	538	929
Mai 2000	12 823	9 855	2 968	533	142	76	127	188	2 435	670	308	519	939
April 2001	12 706	9 655	3 051	574	165	78	128	203	2 477	695	315	503	964

*) 1961 und 1970 Ergebnis der Volkszählung; sonst Ergebnis des Mikrozensus (1975 aus der EG-Arbeitskräftestichprobe).- 1961 wohnberechtigte Bevölkerung, ab 1970 Bevölkerung am Familienwohnsitz.- Als allein Erziehende zählen auch Väter und Mütter mit volljährigen Kindern.- Einschl. der allein Erziehenden, die Partner/in einer nichtehelichen Lebensgemeinschaft sind.

nommen, daher Abweichungen zu den übrigen Tabellen mit Angaben über Familien aus der Volkszählung 1961 und 1970, denn diese stammen aus der Totalerhebung.

2) Familien mit mindestens einem Kind der jeweiligen Altersgruppe und ggf. weiteren Kindern anderer Altersgruppen.

1) Die Zahlen wurden der 10 %-Stichprobe der Volkszählung 1961 und 1970 ent-

/ = Kein Nachweis, da das Ergebnis nicht ausreichend genau ist.

Abweichungen in den Summen ergeben sich durch Runden der Zahlen.

Tabelle 10 Tab. L 204 Quelle: Statistisches Bundesamt

Lange Reihen

Tab. L 204: Familien mit Kindern in der Familie nach Familienstand der Bezugsperson und Altersgruppen der Kinder*)
1 000

Zeitpunkt	Ins- gesamt	Ehepaare mit in der Familie lebenden ledigen Kindern	Allein Erziehende										
			zu- sammen	Männer				Frauen					
				zu- sammen	ledig	ver- heiratet getrennt lebend	ver- witwet	ge- schie- den	zu- sammen	ledig	ver- heiratet getrennt lebend	ver- witwet	ge- schie- den
mit mindestens einem Kind unter 15 Jahren²⁾													
Früheres Bundesgebiet													
6.6. 1961 ¹⁾	6 772	6 208	564	48	/	6	27	14	516	123	61	196	137
27.5. 1970 ¹⁾	7 656	7 063	593	96	6	37	28	25	496	86	64	171	176
Mai 1975.....	7 740	7 155	585	72	/	/	29	29	513	80	41	147	245
Mai 1976.....	7 568	6 954	613	79	/	12	27	35	535	82	45	145	262
April 1977.....	7 491	6 868	623	80	/	13	30	34	544	87	43	143	271
April 1978.....	7 333	6 682	651	83	6	17	29	32	568	87	67	143	270
April 1979.....	7 239	6 568	670	94	7	20	31	35	576	90	80	136	271
April 1980.....	7 067	6 399	668	98	7	22	32	37	570	92	93	127	258
April 1982.....	6 661	5 979	683	94	9	25	26	34	588	109	103	107	269
April 1986.....	5 773	5 047	726	88	16	20	16	36	639	165	106	83	285
März 1987.....	5 662	4 944	718	88	15	19	18	36	631	174	103	73	280
April 1988.....	5 709	4 972	737	93	19	20	19	36	644	186	105	69	284
April 1989.....	5 809	5 068	742	93	22	19	16	36	649	192	100	67	290
April 1990.....	6 001	5 232	770	97	20	19	17	42	673	210	109	62	292
April 1991.....	6 133	5 321	812	109	25	21	16	47	704	229	110	60	305
Mai 1992.....	6 234	5 391	843	104	26	20	14	44	739	245	122	57	315
April 1993.....	6 342	5 446	897	115	35	22	15	44	782	270	133	59	319
April 1994.....	6 410	5 469	942	122	39	24	15	45	819	275	144	63	338
April 1995.....	6 473	5 488	984	144	48	25	19	52	841	279	147	63	352
April 1996.....	6 478	5 440	1 039	170	67	29	18	56	869	291	154	65	358
April 1997.....	6 489	5 408	1 081	182	72	32	17	61	899	304	160	66	370
April 1998.....	6 464	5 368	1 096	184	73	31	16	64	912	310	166	60	377
April 1999.....	6 476	5 321	1 154	187	81	29	16	61	967	327	172	64	405
Mai 2000.....	6 521	5 325	1 197	201	86	30	17	68	996	340	183	60	413
April 2001.....	6 502	5 249	1 253	224	97	31	19	77	1 029	351	186	60	432
Deutschland													
April 1991.....	8 160	6 904	1 256	162	49	24	21	69	1 094	413	124	77	480
Mai 1992.....	8 197	6 909	1 288	158	52	23	20	63	1 130	436	140	76	478
April 1993.....	8 237	6 889	1 348	172	61	27	21	62	1 177	469	157	81	470
April 1994.....	8 222	6 827	1 395	169	57	29	20	63	1 226	478	173	82	493
April 1995.....	8 224	6 775	1 449	193	67	31	25	71	1 256	483	187	82	505
April 1996.....	8 149	6 638	1 512	230	95	38	23	74	1 282	498	196	84	504
April 1997.....	8 095	6 540	1 555	242	103	38	22	78	1 313	512	202	85	513
April 1998.....	8 011	6 433	1 578	248	104	38	21	84	1 330	526	212	80	513
April 1999.....	7 959	6 313	1 646	252	114	37	21	80	1 394	553	218	81	542
Mai 2000.....	7 933	6 245	1 688	272	124	38	21	89	1 416	574	223	78	541
April 2001.....	7 849	6 098	1 750	304	144	39	22	99	1 446	587	226	75	558

*) 1961 und 1970 Ergebnis der Volkszählung; sonst Ergebnis des Mikrozensus (1975 aus der EG-Arbeitskräftestichprobe).- 1961 wohnberechtigte Bevölkerung, ab 1970 Bevölkerung am Familienwohnsitz.- Als allein Erziehende zählen auch Väter und Mütter mit volljährigen Kindern.- Einschl. der allein Erziehenden, die Partner/in einer nichtehelichen Lebensgemeinschaft sind.

1) Die Zahlen wurden der 10 %-Stichprobe der Volkszählung 1961 und 1970 ent-

Abweichungen in den Summen ergeben sich durch Runden der Zahlen.

nommen, daher Abweichungen zu den übrigen Tabellen mit Angaben über Familien aus der Volkszählung 1961 und 1970, denn diese stammen aus der Totalerhebung.

2) Familien mit mindestens einem Kind der jeweiligen Altersgruppe und ggf. weiteren Kindern anderer Altersgruppen.

/ = Kein Nachweis, da das Ergebnis nicht ausreichend genau ist.

Tabelle 11 Tab. L 204 Quelle: Statistisches Bundesamt

Bruttoinlandsprodukt, Bruttonationaleinkommen, Volkseinkommen *)

Jahr	Bruttoinlandsprodukt		Bruttonationaleinkommen (Bruttosozialprodukt)		Volkseinkommen	Bruttoinlandsprodukt		Bruttonationaleinkommen (Bruttosozialprodukt)		Volkseinkommen
	in jeweiligen Preisen	in konstanten Preisen ¹⁾	in jeweiligen Preisen	in konstanten Preisen ¹⁾		in jeweiligen Preisen	in konstanten Preisen ¹⁾	in jeweiligen Preisen	in konstanten Preisen ¹⁾	
	Mrd. EUR ²⁾					Veränderung gegenüber dem Vorjahr in %				
Früheres Bundesgebiet ohne Saarland und Berlin (West)										
1950	49,7	218,2	50,4	223,4	40,1	—	—	—	—	—
1951	61,0	239,3	61,7	244,0	48,7	22,8	9,7	22,4	9,2	21,4
1952	69,8	261,7	70,5	266,2	55,4	14,3	9,3	14,3	9,1	13,6
1953	74,9	285,0	75,6	288,9	59,2	7,4	8,9	7,2	8,5	7,0
1954	80,4	307,1	80,8	309,4	63,4	7,3	7,8	6,9	7,1	6,9
1955	91,9	344,3	92,3	346,3	72,7	14,3	12,1	14,2	11,9	14,7
1956	101,6	370,9	102,0	372,8	80,7	10,5	7,7	10,5	7,7	11,0
1957	110,7	393,4	111,2	395,3	88,5	9,0	6,1	9,0	6,0	9,6
1958	119,0	410,9	119,2	411,9	94,7	7,4	4,5	7,3	4,2	7,1
1959	130,3	443,2	130,3	442,9	103,1	9,6	7,9	9,3	7,5	8,8
1960	146,0	481,4	146,2	481,5	116,1	12,1	8,6	12,2	8,7	12,7
Früheres Bundesgebiet										
1960	154,8	511,3	154,9	511,3	122,8	—	—	—	—	—
1961	169,6	535,0	169,4	533,5	133,3	9,6	4,6	9,4	4,3	8,6
1962	184,5	559,9	184,3	558,4	144,2	8,8	4,7	8,8	4,7	8,2
1963	195,5	575,7	195,4	574,1	152,3	6,0	2,8	6,0	2,8	5,6
1964	214,8	614,0	214,5	611,7	167,3	9,9	6,7	9,8	6,6	9,9
1965	234,8	646,9	234,3	643,7	183,3	9,3	5,4	9,2	5,2	9,5
1966	249,6	664,9	249,2	661,9	194,2	6,3	2,8	6,4	2,8	6,0
1967	252,8	662,9	252,4	660,0	194,7	1,3	-0,3	1,3	-0,3	0,3
1968	272,7	699,0	272,9	697,7	213,8	7,9	5,5	8,1	5,7	9,8
1969	305,2	751,2	305,7	749,4	236,7	11,9	7,5	12,0	7,4	10,7
1970	352,0	897,0	352,9	904,0	274,7	—	—	—	—	—
1971	390,1	926,7	390,5	932,7	302,4	10,8	3,3	10,7	3,2	10,1
1972	427,5	964,8	427,8	968,9	332,2	9,6	4,1	9,5	3,9	9,9
1973	476,7	1 008,8	476,9	1 012,9	372,5	11,5	4,6	11,5	4,5	12,1
1974	513,6	1 013,4	513,7	1 018,0	400,7	7,7	0,5	7,7	0,5	7,6
1975	536,0	1 002,8	536,4	1 007,7	416,9	4,4	-1,0	4,4	-1,0	4,0
1976	583,9	1 052,9	585,4	1 061,0	456,2	8,9	5,0	9,1	5,3	9,4
1977	623,7	1 084,9	624,0	1 090,6	486,2	6,8	3,0	6,6	2,8	6,6
1978	669,3	1 117,4	672,6	1 127,3	524,4	7,3	3,0	7,8	3,4	7,9
1979	722,5	1 163,9	724,1	1 170,8	561,3	7,9	4,2	7,7	3,9	7,0
1980	766,6	1 178,7	769,0	1 186,9	591,0	6,1	1,3	6,2	1,4	5,3
1981	800,2	1 180,0	801,0	1 185,5	613,4	4,4	0,1	4,2	-0,1	3,8
1982	831,8	1 170,7	829,7	1 171,9	632,8	3,9	-0,8	3,6	-1,1	3,2
1983	872,2	1 188,9	873,9	1 195,5	666,2	4,9	1,6	5,3	2,0	5,3
1984	915,0	1 222,5	920,4	1 233,9	702,2	4,9	2,8	5,3	3,2	5,4
1985	955,3	1 249,3	960,4	1 260,8	735,5	4,4	2,2	4,3	2,2	4,7
1986	1 010,2	1 279,5	1 014,1	1 288,3	781,9	5,7	2,4	5,6	2,2	6,3
1987	1 043,3	1 298,3	1 045,7	1 304,5	806,2	3,3	1,5	3,1	1,3	3,1
1988	1 098,5	1 346,8	1 104,1	1 357,2	853,3	5,3	3,7	5,6	4,0	5,8
1989	1 168,3	1 399,5	1 177,1	1 414,4	906,9	6,4	3,9	6,6	4,2	6,3
1990	1 274,9	1 479,6	1 284,5	1 494,9	987,2	9,1	5,7	9,1	5,7	8,9
1991	1 387,1	1 555,0	1 390,4	1 562,0	1 065,4	8,8	5,1	8,2	4,5	7,9
Deutschland										
1991	1 502,2	1 710,8	1 511,1	1 722,5	1 167,1	—	—	—	—	—
1992	1 613,2	1 749,1	1 621,2	1 759,3	1 242,6	7,4	2,2	7,3	2,1	6,5
1993	1 654,2	1 730,1	1 661,0	1 738,2	1 255,7	2,5	-1,1	2,5	-1,2	1,1
1994	1 735,5	1 770,7	1 728,4	1 763,8	1 302,6	4,9	2,3	4,1	1,5	3,7
1995	1 801,3	1 801,3	1 791,8	1 791,8	1 358,6	3,8	1,7	3,7	1,6	4,3
1996	1 833,7	1 815,1	1 825,7	1 808,2	1 381,7	1,8	0,8	1,9	0,9	1,7
1997	1 871,6	1 840,4	1 862,5	1 832,3	1 404,6	2,1	1,4	2,0	1,3	1,7
1998	1 929,4	1 876,4	1 915,4	1 863,6	1 442,2	3,1	2,0	2,8	1,7	2,7
1999	1 978,6	1 914,8	1 965,1	1 902,9	1 468,2	2,6	2,0	2,6	2,1	1,8
2000	2 030,0	1 969,5	2 020,3	1 961,7	1 508,4	2,6	2,9	2,8	3,1	2,7
2001	2 073,7	1 986,2	2 065,6	1 979,6	1 542,2	2,2	0,8	2,2	0,9	2,2
2002	2 110,4	1 989,7	2 108,8	1 989,5	1 571,5	1,8	0,2	2,1	0,5	1,9

*) Für den Zeitraum 1925 bis 1938 auf das frühere Bundesgebiet umgerechnete Ergebnisse des Statistischen Reichsamtes; die Definitionen und Abgrenzungen entsprechen zum Teil nicht den heutigen

Tabelle 12 BIP-LR1 Quelle: Statistisches Bundesamt

Land	Brutto-sozial-produkt pro Kopf (in 1.000 US-Dollar) ¹	Anteil der öffentlichen Ausgaben für das Schulsystem in Prozent des Brutto-sozialprodukts ²	Retentivität im Abschlussjahr der Sekundarstufe II (in % der Alterskohorte) ³	TIMSS Coverage Index (TCI) ³	Relativer Besuch vor-universitärer Bildungswege (in % der Alterskohorte) ³	Mittleres Alter im Abschlussjahr ³	Mathematisch-naturwissenschaftliches Grundbildungsniveau (Mittelwert) ³
Australien	18,0	3,7	72	68	37	17,7	525
Dänemark	28,1	4,8	59	58	33	19,1	528
Deutschland	25,6	2,4	84	78	25	19,5	496
Frankreich	23,5	3,6	85	84	75	18,8	505
Island	24,6	4,8	55	55	45	21,2	541
Italien	19,3	2,9	74*	74*	55	18,7	475
Kanada	19,6	4,6	78	70	54	18,6	526
Neuseeland	13,2	3,2	71	70	70	17,6	525
Niederlande	22,0	3,3	92	78	34	18,5	559
Norwegen	26,5	5,3	85	84	48	19,5	536
Österreich	25,0	4,2	93	76	33	19,1	519
Schweden	23,6	4,9	71	71	47	18,9	555
Schweiz	37,2	3,7	84	82	19	19,8	531
Slowenien	7,1	4,2	93	88	80	18,8	514
Südafrika	3,0	5,1	49	49	49	20,1	352
Tschechien	3,2	3,8	83	78	36	17,8	476
Ungarn	3,8	4,3	65	65	18	17,5	477
USA	25,9	4,0	65	63	35	18,1	471
Zypern	10,4	3,6	61	48	44	17,7	447

¹ Weltbank 1996 (in Mullis u.a., 1998, Tab. 5, S. 23).
² Unesco Statistisches Jahrbuch 1995 (in Mullis u.a., 1998, Tab. 5, S. 23).
³ TIMSS/III.
* Vier Regionen sind ausgeschlossen.

IEA. Third International Mathematics and Science Study. © TIMSS/III-Germany

Tabelle 13 Ausgewählte Kennziffern der Bildungssysteme Quelle: TIMSS/III

COUNTRY	GROUP: TOTAL WITHOUT G_MIG **)			G_MIG: MOTHER AND FATHER BORN: <OTHER> ***)			PERCENT OF STUDENTS IN G_MIG	HISEI DIFF [G_MIG - NON MIG]	GRADE DIFF [G_MIG - NON MIG]
	HISEI HIGHEST IN. SOCIO- ECON. INDEX	GRADE	NUMBER OF STUDENTS	HISEI HIGHEST IN. SOCIO- ECON. INDEX G_MIG	GRADE G_MIG	NUMBER OF STUDENTS G_MIG			
PISA 2000*)	mean	mean		mean	mean		H_Diff	G_Diff	
ALBANIA	45,80	9,80	4815	42,86	9,71	34	0,70%	-2,95	-0,09
ARGENTINA	45,51	9,66	3678	36,14	9,38	79	2,10%	-9,37	-0,28
AUSTRALIA	52,61	10,10	3999	52,28	10,13	1038	20,61%	-0,33	0,03
AUSTRIA	49,95	9,52	4212	39,92	9,03	457	9,79%	-10,03	-0,50
BELGIUM	50,56	9,72	5788	40,66	9,21	743	11,38%	-9,90	-0,51
BRAZIL	43,01	8,26	4653	36,53	7,73	22	0,47%	-6,48	-0,53
BULGARIA	49,62	9,12	4439	56,00	8,88	17	0,38%	6,38	-0,24
CANADA	51,07	9,80	25232	52,64	9,86	3596	12,47%	1,57	0,06
CHILE	40,57	9,55	4719	47,46	9,31	13	0,27%	6,89	-0,24
CZECH REPUBLIC	49,22	9,58	5095	45,22	9,22	51	0,99%	-4,00	-0,36
DENMARK	50,12	8,97	3911	43,88	8,89	244	5,87%	-6,24	-0,07
FINLAND	50,16	8,89	4736	48,43	8,61	57	1,19%	-1,73	-0,28
FRANCE	49,33	9,52	4010	41,16	9,30	538	11,83%	-8,16	-0,23
GERMANY	51,25	9,13	4223	41,01	8,77	708	14,36%	-10,24	-0,36
GREECE	47,46	10,16	4463	45,53	9,53	144	3,13%	-1,93	-0,62
HONG KONG	44,67	10,06	2439	39,30	9,57	1879	43,52%	-5,37	-0,48
HUNGARY	49,30	9,34	4733	53,27	9,23	80	1,66%	3,97	-0,11
ICELAND	52,78	10,00	3296	45,25	10,00	25	0,75%	-7,53	0,00
INDONESIA	34,90	9,27	7009	28,64	8,80	30	0,43%	-6,26	-0,47
IRELAND	48,56	9,51	3673	54,05	9,59	85	2,26%	5,49	0,08
ISRAEL	56,13	9,97	2823	54,75	9,89	1042	26,96%	-1,38	-0,09
ITALY	47,12	9,88	4872	40,53	9,38	39	0,79%	-6,60	-0,49
JAPAN	50,62	10,00	4890	36,00	10,00	7	0,14%	-14,62	0,00
LATVIA	51,42	9,47	2802	47,27	9,44	888	24,07%	-4,15	-0,04
LIECHTENSTEIN	49,19	8,89	244	40,64	8,71	61	20,00%	-8,55	-0,18
LUXEMBOURG	49,01	9,17	2289	36,24	8,83	1113	32,72%	-12,76	-0,35
MEXICO	43,54	9,48	4167	35,47	8,74	137	3,18%	-8,07	-0,74
NETHERLANDS	52,87	9,48	2217	40,98	9,12	262	10,57%	-11,89	-0,36
NEW ZEALAND	51,51	10,98	2871	54,57	10,96	661	18,71%	3,06	-0,02
NORWAY	54,45	10,00	3836	48,27	9,94	217	5,35%	-6,19	-0,06
PERU	41,12	9,39	4158	49,75	8,91	11	0,26%	8,63	-0,48
POLAND	44,73	9,00	3577	51,50	9,00	8	0,22%	6,77	0,00
PORTUGAL	44,29	9,34	4362	50,88	9,03	141	3,13%	6,59	-0,31
ROMANIA	47,69	9,06	4754	61,63	8,78	9	0,19%	13,94	-0,28
RUSSIAN FEDERATION	49,64	9,72	5944	49,64	9,68	280	4,50%	0,00	-0,04
SPAIN	45,16	9,72	5946	41,29	9,49	120	1,98%	-3,87	-0,23
SWEDEN	51,09	8,99	3922	46,60	8,93	443	10,15%	-4,48	-0,05
SWITZERLAND	50,90	8,96	4809	40,28	8,81	1141	19,18%	-10,62	-0,15
THAILAND	34,39	9,37	5231	32,09	8,35	23	0,44%	-2,30	-1,02
MACEDONIA	47,32	9,28	4208	43,12	9,18	119	2,75%	-4,21	-0,09
UNITED KINGDOM	50,67	10,71	8344	50,35	10,64	513	5,79%	-0,32	-0,07
UNITED STATES	52,26	9,58	3058	45,25	9,47	558	15,43%	-7,01	-0,10

*) DATABASE: INSTUD_READ (WITHOUT CONSID. OF. W_FSTUWT - STUD. POP. WEIGHT -)

download http://pisaweb.acer.edu.au/oeecd/oeecd_pisa_data_s1.html

VALID CASES 216080

INVALID CASES 12704

TOTAL 228784

KOREA: INVALID CASES N = 4982 (100%)

***) G_MIG: ST16Q03 COUNTRY OF BIRTH: MOTHER AND FATHER BORN <OTHER>

***) GROUP: TOTAL WITHOUT G_MIG

Tabelle 14 Sozioökonomische Differenz und Schulalterdifferenz

7. Dokumentation (Programmbeispiele)

```
*-----*
*
*           PISA 2000-E
*         ISEI GYMNASIUM-REALSCHULE
*
*-----*

SET PRINTBACK=ON MPRINT=ON.

DEFINE !stat1 (size=!CHAREND('/') /vnames=!CMDEND)
!DO !s !IN (!size)

*GET FILE='C:\ ... <path>...PISA_OVE_9KL_short.sav'.
GET FILE='C:\ ... <path>...PISA-OVE-9KL.sav'.

FILTER OFF.
USE ALL.
SELECT IF(st_bl = !s).
EXE .

DESCRIPTIVES VARIABLES=!vnames /SAVE.

SAVE OUTFILE=!QUOTE(!CONCAT('C:\ ... <path>...bundland',!s,'.sav')).
!DOEND

!ENDDEFINE.

!stat1 size = 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 /vnames=isei.

DEFINE !stat18 (size=!CHAREND('/') /vnames=!CMDEND).
GET FILE='C:\ ... <path>...bundland31.sav'.
!DO !s !IN (!size)
ADD FILES /FILE=* /FILE= !QUOTE(!CONCAT('C:\ ... <path>...bundland',!s,'.sav')).
!DOEND
comp sum_ =1.

AGGREGATE
/OUTFILE='C:\ ... <path>...Gynmasium_Reals.sav'
/BREAK=st_bl schulf
/zisei_1 = MEAN(zisei)
/sum = SUM(sum_).

GET FILE='C:\ ... <path>...Gynmasium_Reals.sav'.

FILTER OFF.
USE ALL.

SELECT IF(ANY(schulf,3,5)).
EXE .

SORT CASES BY st_bl .

CASESTOVARS
/ID = st_bl
/GROUPBY = VARIABLE .

CORRELATIONS /VARIABLES=v1 with v2 /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE .
NONPAR CORR /VARIABLES=v1 with v2 /PRINT=SPEARMAN TWOTAIL NOSIG /MISSING=PAIRWISE .
GRAPH /SCATTERPLOT(BIVAR)=v2 WITH v1 BY st_bl (IDENTIFY) /MISSING=LISTWISE .

!ENDDEFINE.

!stat18 size = 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 /vnames=isei.

SAVE OUTFILE='C:\ ... <path>...REAL_GYM.sav' /COMPRESSED.
```

```

*-----*
*
*           PISA 2000
*       Decken/Bodeneffekt/Disparität
*
*-----*

```

```

SET PRINTBACK = on.
DEFINE !f$_ (f01$_ = !tokens(1)
            /f_02$_ = !tokens(1)
            /f_03$_ = !tokens(1)
            /f_04$_ = !tokens(1)
            /f_05$_ = !tokens(1)
            /f_06$_ = !tokens(1)
            /f_07$_ = !tokens(1)
            /f_08$_ = !tokens(1)
            /f_09$_ = !tokens(1)
            /f_10$_ = !tokens(1)
            /f_11$_ = !tokens(1)).

title !f_10$_.
GET FILE=!f_03$_.
WEIGHT OFF.
comp sum_ = 1.
USE ALL.

*COMPUTE filter_$=(!f_07$_ = !f_08$_).
*FILTER BY filter_$.
WEIGHT BY !f_11$_.
RANK VARIABLES=hisei (A) BY country /NTILES(4) /PRINT=YES /TIES=MEAN .
title 'Check 25%'.
subtitle 'HISEI 25%'.
CROSSTABS /TABLES=country BY nhisei /FORMAT= AVALUE TABLES /CELLS= COUNT ROW .

TEMPORARY.
COMP p_r = MEAN(pv1read to pv5read) .

AGGREGATE /OUTFILE=!f_04$_ /BREAK=country nhisei /pv1rea_1 = MEAN(p_r)/sum =SUM(sum_).

GET FILE=!f_04$_.
ren var (pv1rea_1 = pv).
FILTER OFF.
USE ALL.
SELECT IF(nhisei > 0).
EXE .

CASESTOVARS /ID = country /GROUPBY = VARIABLE .

COMPUTE diff = pv.4 - pv.1 .

SAVE OUTFILE=!f_04$_.

GET FILE=!f_03$_.
WEIGHT OFF.
comp sum_ =1.
USE ALL.
*COMPUTE filter_$=(!f_07$_ = !f_08$_).
*FILTER BY filter_$.
WEIGHT BY !f_11$_.

COMP p_r = MEAN(pv1read to pv5read) .
RANK VARIABLES=p_r (A) BY country /NTILES(4) /PRINT=NO /TIES=MEAN .

ren var np_r = npv.

title 'Check 25%'.

CROSSTABS /TABLES=country BY npv /FORMAT= AVALUE TABLES /CELLS= COUNT ROW .

TEMPORARY.
COMP p_r = MEAN(pv1read to pv5read) .
AGGREGATE
/OUTFILE=!f_05$_
/BREAK=country npv

```

```

/pv1 = MEAN(p_r)
/sum=SUM(sum_).
GET FILE=!f_05$.

CASESTOVARS /ID = country /GROUPBY = VARIABLE .

ren var (pv1.1 to pv1.4 = quant1 to quant4).

SAVE OUTFILE=!f_05$.

MATCH FILES /FILE=* /FILE=!f_04$_/BY country.
EXE.

SAVE OUTFILE=!f_06$.
GET FILE=!f_06$.
ERASE FILE =!f_04$.
ERASE FILE =!f_05$.

IF (ANY(country,76 ,756,752,724,643,620,616,578,56 ,554,528,484,442,
438,428,410,40 ,392,380,372,36 ,352,348,300,276,250,246,208,203,124)) sel_ = 1 .

RECODE sel_ (MISSING=2) .
COMP N_total = SUM(sum.1,sum.2,sum.3,sum.4) .
exe.

IF (n_total > !f_09$_ & sel_ = 1) filter = 1 .
EXE .

USE ALL.
FILTER BY filter.
EXE .

CORRELATIONS /VARIABLES=quant1 with diff /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE .
NONPAR CORR /VARIABLES=quant1 with diff /PRINT=SPEARMAN TWOTAIL NOSIG /MISSING=PAIRWISE .
GRAPH /SCATTERPLOT(BIVAR)=quant1 WITH diff BY country (IDENTIFY) /MISSING=LISTWISE .

!ENDDEFINE.

!f$
f_10$ = 'PISA 2000'
f_03$ = 'C:<path>\<infile>.sav'
f_04$ = 'C:<path>\f_12$.sav'
f_05$ = 'C:<path>\f_13$.sav'
f_06$ = 'C:<path>\f_14$.sav'
f_07$ = st02q01
f_08$ = 10
f_09$ = 2
f_11$ = W_FSTUWT.

```

```

*-----*
*                PISA 2000/2003
*                MIGRATION
*-----*

SET PRINTBACK = on.
DEFINE !mig (v1 = !tokens (1)
            /v2 = !tokens (1)
            /v3 = !tokens(1)
            /v4 = !tokens(1)
            /v5 = !tokens(1)
            /v6 = !tokens(1)
            /infile = !tokens(1)
            /sub = !tokens(1)).

title 'PISA: MIG - GRADE'.
subt !sub.
GET FILE=!infile.
WEIGHT OFF.
USE ALL.

IF (!v1 = 1 & !v2 = 1 & !v3 = 1) m_08 = 1 .
IF (!v1 = 1 & !v2 = 1 & !v3 = 2) m_08 = 1 .
IF (!v1 = 1 & !v2 = 2 & !v3 = 1) m_08 = 1 .
IF (!v1 = 1 & !v2 = 2 & !v3 = 2) m_08 = 2 .

IF (!v1 = 2 & !v2 = 1 & !v3 = 1) m_08 = 1 .
IF (!v1 = 2 & !v2 = 1 & !v3 = 2) m_08 = 1 .
IF (!v1 = 2 & !v2 = 2 & !v3 = 1) m_08 = 1 .
IF (!v1 = 2 & !v2 = 2 & !v3 = 2) m_08 = 2 .
VAR LAB m_08 'Migrationshintergrund' .

comp sum = 1.

val lab m_08 1 'Mind. ein Elternteil ohne Migration*' 2 'Beide Eltern mit Migration*'.

USE ALL.

COMPUTE filter_$=(m_08 = !v4).
FILTER BY filter_$ .

AGGREGATE OUTFILE='C:\...<path>...\break_oM.sav' /BREAK=country
           /grade_oM = MEAN(!v6) /hisei_oM = MEAN(hisei) / sum_oM = SUM(sum).
USE ALL.
COMPUTE filter_$=(m_08 = !v5).

FILTER BY filter_$ .

AGGREGATE OUTFILE='C:\...<path>...\break_mM.sav' /BREAK=country
           /grade_mM = MEAN(!v6) /hisei_mM = MEAN(hisei) / sum_mM = SUM(sum).

GET FILE='C:\...<path>...\break_oM.sav'.
SORT CASES BY country (A) .

MATCH FILES /FILE=* /FILE='C:\...<path>...\break_mM.sav' /BY country.
COMP H_diff = hisei_mm - hisei_om .
COMP G_diff = grade_mm - grade_om .

formats h_diff (f5.3).
formats g_diff (f5.3).

compute percent = sum_mm/(sum(sum_om,sum_mm)).

FILTER OFF.
USE ALL.
SELECT IF(percent >= 0.05).
EXE .

CORRELATIONS VARIABLES=h_diff with g_diff / PRINT=TWOTAIL NOSIG / MISSING=PAIRWISE .
NONPAR CORR VARIABLES=h_diff with g_diff /PRINT=SPEARMAN TWOTAIL NOSIG /MISSING=PAIRWISE .
GRAPH /SCATTERPLOT(BIVAR)=h_diff WITH g_diff BY country (NAME) /MISSING=LISTWISE .

SAVE OUTFILE='C:\...<path>...\hisei_grade_mig_noMig.sav' .

```

```
GET FILE='C:\...<path>...\hisei_grade_mig_noMig.sav'.
```

```
erase file = 'C:\...<path>...\break_mM.sav'.
```

```
erase file = 'C:\...<path>...\break_oM.sav'.
```

```
!ENDDDEFINE.
```

```
/* PISA 2000 READ */.
```

```
!mig v1 = ST16Q01  
v2 = ST16Q02  
v3 = ST16Q03  
v4 = 1  
v5 = 2  
v6 = st02q01  
sub = 'READING 2000'  
infile = 'C:\...<path>...\<filename>.sav' .
```

```
/* PISA 2003 */.
```

```
!mig v1 = ST15Q01  
v2 = ST15Q02  
v3 = ST15Q03  
v4 = 1  
v5 = 2  
v6 = st01q01  
sub = 'READING 2003'  
infile = 'C:\...<path>...\<filename>.sav' .
```

```
/* PISA 2000 MATH*/.
```

```
!mig v1 = ST16Q01  
v2 = ST16Q02  
v3 = ST16Q03  
v4 = 1  
v5 = 2  
v6 = st02q01  
sub = 'MATH 2000'  
infile = 'C:\...<path>...\<filename>.sav' .
```

8. Literatur

ACKEREN, I.V., KLEMM, K.: TIMSS, PISA, LAU, MARKUS Ein aktueller Überblick über Typen und Varianten von Schulleistungsstudien. In: Pädagogik 52, 2000, Heft 12,

ADAMS, R., WU, M. (Hrsg.): PISA 2000 Technical Report Download, www.acer.com

ALKIN, M. C., & SOLMON, L. C. (Hrsg.). The costs of evaluation. Beverly Hills, CA: Sage. 1983

ALKIN, M. C., & SOLMON, L. C. (Hrsg.) The costs of evaluation. Beverly Hills, CA: Sage. 1983

ARNOLD, K.H.: Fairness bei Schulsystemvergleichen. Münster 1999

ARTELT, C., SCHNEIDER, W. UND SCHIEFELE, U.: Ländervergleich zur Lesekompetenz. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich Opladen 2001

ARTELT, C., STANAT, P., SCHNEIDER, W. UND SCHIEFELE, U.: Lesekompetenz: Testkonstruktion und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

AURIN, K. (Hrsg.): Schulvergleich in der Diskussion. Stuttgart 1987

BACKHAUS, K., ERICHSON, B., PLINKE, W., WEIBER, R.: Multivariate Analysemethoden. Berlin 1996

BARNETT, S., FREDE, E. C., COX, J. O., & BLACK, T.: Using cost analysis to improve early childhood programs. In W. S. Barnett (Ed.), Cost analysis for education decisions: Methods and examples. (Vol. 4, pp. 145-181). Greenwich, CT: JAI. 1994

BAUER, K.O.: Erziehungsbedingungen von Sekundarschulen. Weinheim 1980

BAUMERT, J., ARTELT, C., CARSTENSEN, C.H., SIBBERNS, H., STANAT, P.: Untersuchungsgegenstand, Fragestellungen und technische Grundlagen der Studie. In:

Deutsches PISA-Konsortium (Hrsg.) PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich Opladen 2002

BAUMERT, J., ARTELT, C., KLIEME, E., NEUBRAND, M., PRENZEL, M., SCHIEFELE, U., SCHNEIDER, W., SCHÜMER, G., STANAT, P., TILLMANN, K.-J.: 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
Download: www.mpib-berlin.mpg.de

BAUMERT, J., BOS, W., WATERMANN, R.: Mathematische und naturwissenschaftliche Grundbildung im internationalen Vergleich. In: BAUMERT, J., BOS, W., Lehmann (Hrsg.) TIMSS, III Dritte Internationale Mathematik- und Naturwissenschaftsstudie, Opladen 2000, Band 1

BAUMERT, J., BOS, W., LEHMANN, R. (Hrsg.): TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie. Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen 2000

BAUMERT, J., KÖLLER, O., SCHNABEL, K.U.: Schulformen als differentielle Entwicklungsmilieus – eine ungehörige Fragestellung?. In: Gewerkschaft Erziehung und Wissenschaft (Hrsg.): Messung sozialer Motivationen – eine Kontroverse, Frankfurt, M. 2000,

BAUMERT, J., LEHMANN, R.: TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen 1997

BAUMERT, J., SCHÜMER, G.: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 - Die Länder der Bundesrepublik im Vergleich. Opladen 2002

BAUMERT, J., SCHÜMER, G.: Schulformen als selektionsbedingte Lernmilieus.. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000- Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen 2001b

BAUMERT, J., KÖLLER, O., LEHRKE, BROCKMANN: Anlage und Durchführung der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur Sekundarstufe II (TIMSS, III) – Technische Grundlagen. In: BAUMERT, J., BOS, W., Lehmann (Hrsg.) TIMSS, III Dritte Internationale Mathematik- und Naturwissenschaftsstudie, Opladen 2000, Band 1

BAUMERT, J., SCHÜMER, G.: Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

BAUMERT, J., STANAT, P., DEMMRICH, A.: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

BAUMERT, J., WEIß, M.: Föderalismus und Gleichwertigkeit der Lebensverhältnisse. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich Opladen 2002

BAUMERT, J.: TIMSS-Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. In: List: TIMSS Mathematische-naturwissenschaftliche Kenntnisse deutscher Schüler auf dem Prüfstand Köln 1998

BECKER, E.: Von der Zukunftsinvestition zur Effektivitätskontrolle des Bildungssystems, In: Radtke, Weiß (Hrsg.): Schulautonomie, Wohlfahrtsstaat und Chancengleichheit Opladen 2000

BETTS, J.R. : Is there a link between school inputs and earnings? In: Burtless (Hrsg.): Does money matter? The effect of school resources on student achievement and adult success. Washington 1996

BILS, M., Klenow, P.J.: Does Schooling Cause Growth? American Economic Review 90/5 1160-1183

BISHOP, J.H., WÖBMAN, L.: Institutional Effects in a Simple Model of Educational Production, Kiel Working Paper No.1085, 2001

BORTZ, J., LIENERT, G.A., BOEHNKE, K. : Verteilungsfreie Methoden in der Biostatistik. Berlin 1999

BOS, W., LANKES, E.M., SCHWIPPERT, K., VALTIN, R., VOSS, A., BADEL, I., PLABMEIER, N. : Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich: In Erste Ergebnisse aus IGLU Münster 2003

BOS, W., VALTIN, R., LANKES, E.M., SCHWIPPERT, K., VOSS, A., BADEL, I., PLABMEIER, N. : Lesekompetenzen am Ende der vierten Jahrgangsstufe in einigen Ländern der Bundesrepublik Deutschland im nationalen Vergleich. In: BOS, W., LANKES, E.M., Prenzel, M., SCHWIPPERT, K., VALTIN, R., WALTHER, G. (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II) Münster 2004

BOSCH K.: Großes Lehrbuch Statistik. München 1996

BREWER, D.J., KROP, C., GILL, B.P., und REICHHARDT, R.: Estimating the cost of national class size reductions under different policy alternatives. In: Educational Evaluation and Policy Analysis. 21/2

BRÜGELMANN, H.: Besserwisser und Alleskönner, GEWERKSCHAFT ERZIEHUNG UND WISSENSCHAFT HAUPTVORSTAND 11/12/2001

BÜNING H., TRENKLER: Nichtparametrische statistische Methoden. Berlin 1994

BÜNING, H.: Robuste und adaptive Tests. Berlin 1991

CATERALL, J. S.: On the social costs of dropping out of school. The High School Journal. 81(1), 1987

CHAMBERS, J.: The development of a cost of education index. Journal of Education Finance. 5/3, 1980

CHAMBERS, J., FOWLER, W.J.: Public school teacher cost differences across the United States. Washington, DC, NCFES 1995

CHAMBERS, J., & PARRISH, T.. Developing a resource cost database. In W. S. Barnett (Ed.), Cost analysis for education decisions: Methods and examples (Vol. 4, pp. 23-44). Greenwich, CT: JAI. 1994

COLLANI, E.v.: OECD PISA – An Example of Stochastic Illiteracy? Economic Quality Control Vol 16, 2001, No. 2, 227-253

CORTINA K., Leschinsky, A.: zur sozialen Einbettung bildungspolitischer Trends in der Bundesrepublik. In: Cortina, Baumert, Leschinsky, Mayer, Trommer (Hrsg.): Das Bildungswesen in der Bundesrepublik Deutschland Hamburg 2003

DALGAARD, B. R., LEWIS, D. R. und BOYER, C. M.: Cost and effectiveness considerations in the use of computer-assisted instruction in economics. Journal of Economic Education. 15/4

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

DENISON, E.F.: Why Growth Rates Differ: Postwar Experience in Nine Western Countries. Washington 1967

DEUTSCHES PISA-KONSORTIUM (Hrsg.): PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen 2001

DOWNES, T.A., POGUE, T.F.: Adjusting school aid formulas for the higher cost of education disadvantaged students. National Tax Journal, 47/1, 1994

DUNCOMBE, W., RUGGIERO, J. und YINGER, J.: Alternative approaches to measuring the cost of education. In: Ladd (Hrsg.): Holding schools accountable: Performance-based reform in education. 1996

ENTORF, H., MINOIUJ, N.: PISA Results: What a Difference Immigration Law Makes IZA DP No. 1021 Institute for the Study of Labor February 2004

FEND, H., KNÖRZER, W., NAGL, W., SPECHT, W., VÄTH-SZUSDZIARA, R.: Gesamtschule und dreigliedriges Schulsystem – eine Vergleichsstudie über Chancengleichheit und Durchlässigkeit. Stuttgart 1976a FORSCHUNGSGRUPPE SCHULEVALUATION: Gewalt als soziales Problem in Schulen. Untersuchungsergebnisse und Präventionsstrategien. Opladen 1998

FEND, H.: Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs. Weinheim, Basel 1982

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 http://ssrn.com/abstract_id=323599

Fertig, M., Wright, R.E.: School Quality, Educational Attainment and Aggregation Bias, IZA DP No. 994, Januar 2004

FLETCHER; J.D. : Effectiveness and cost of interactive videodisc instruction in defense training and education. IDA Paper P-2372. Alexandria 1990

FODERS, F., SIEBERT, H. (Hrsg.): Bildungspolitik für den Standort D. Berlin, Heidelberg 2001

FREYMAN VON: PISA in Finnland. In: MUT 9, 2003

FUENTE, A., CICCONE, A. (Europäische Kommission): Das Humankapital in der wissensbasierten globalen Wirtschaft. Abschlussbericht. Luxemburg 2002

GEDIGA G.: Skalierung. Münster 1998

GRAF, ORTSEIFEN: Statistische und graphische Datenanalyse mit SAS. Heidelberg 1995

GREENBERG, D. H., & APPENZELLER, U. Cost analysjs step by step. New York Research Corporation. 1998

GROEBEN, A.: Schulen für ein neues Verständnis von Leistung und Qualität. In: Pädagogik, 52, 2000, Heft 12,

GUNDLACH, E., WÖBMANN: The Fading Productivity of Schooling in East Asia, Kiel Working Paper No.945 1999

GUNDLACH, E., WÖBMANN, L.: Die sinkende Produktivität der Schulen in OECD-Ländern, In: WEIß, M., WEISHAUPT, H.: Bildungsökonomie und Neue Steuerung. Frankfurt a.M. 2000

GUNDLACH, E., Rudman, D., Wößmann, L.: Second Thoughts on Developing Accounting. Applied Economics 34/11, 2002

HAGEMEISTER, V.: Was wurde bei TIMSS erhoben? In: Die Deutsche Schule, 91/1999

HANUSHEK, E.A., KIMKO, D.D.: Schooling, Labor-Force Quality and the Growth of Nations. American Economic Review 90/5. 2000

HANUSHEK, E.A., LUQUE, J.A.: Efficiency and Equity in Schools Around the World, NATIONAL BUREAU OF ECONOMIC RESEARCH, INC. Stanford University, Cambridge, WORKING PAPER 8949, 2002

HANUSHEK, E.A.: The Long Run Importance of School Quality, NATIONAL BUREAU OF ECONOMIC RESEARCH, INC. Stanford University, Cambridge, WORKING PAPER 9071, 2002

HARTUNG, J.: Statistik. München 1999

HÄSSIG, M.: Grundlagen der Informatik, Biometrie und Epidemiologie. Departement für Fortpflanzungskunde MPH FVH 1999

HELSPER, W., BÖHME, J., KRAMER, R.-T., LINGKOST, A.: Schulkultur und Schulmythos. Rekonstruktionen zur Schulkultur I. Opladen 2001

HANSEN, H.R., NEUMANN, G.: Wirtschaftsinformatik. Weinheim Basel 2001

HORSTKEMPER, M.: Schule, Geschlecht und Selbstvertrauen. Eine Längsschnittstudie über Mädchensozialisation in der Schule. Weinheim, München 1987

JÜRGES, H., SCHNEIDER, K., BÜCHEL, K.: The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany, München 2003

KAHL, R. (2001): Der Studienrat. Am 4. Dezember verteilt die internationale PISA-Studie Zensuren an die Schule. Jürgen BAUMERT, J. verantwortet den deutschen Teil. In: Die Zeit, Nr. 49 v. 29.11.01, 46

KLEMM, K. (2001): Unfähig, die Schwächen der Schüler zu erkennen. Die PISA-Studie kreidet Deutschland den Sonderweg im Bildungssystem sowie grundsätzliche Fehler im Unterricht an. In: Frankfurter Rundschau v. 6.12.2001

KLEMM, K.: Warten auf den nächsten Happen. Warum die Internationale Mathe-Physik-Studie für den Schulstreit in Deutschland ungeeignet ist. In: Frankfurter Rundschau vom 28. 5. 1998

KLIEME, E., BAUMERT, J., BOS, W.: Mathematisch-naturwissenschaftliche Grundbildung im internationalen Vergleich. In: BAUMERT, J., BOS, W., Lehmann (Hrsg.) TIMSS, III Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Opladen 2000, Band 1

KLIEME, E., BAUMERT, J., KÖLLER, O., BOS, W.: Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung von Kompetenzen. In: BAUMERT, J., BOS, W., LEHMANN R. (Hrsg.) TIMSS, III Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Opladen 2000, Band 1

KLIEME, E., NEUBRAND, M., LÜDTKE, O. : Mathematische Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

KLIEME, AVENARIUS, BLUM, DÖBRICH, GRUBER, PRENZEL, REISS, RIQARTS, ROST, TENORTH, VOLLMER, LAUKART, PAETZ: Zur Entwicklung nationaler Bildungsstandards – Eine Expertise

KMK: Kultusministerkonferenz erzielt Einigung mit Lehrerverbänden über Konsequenzen aus der PISA-Studie. Sieben Handlungsfelder stehen zunächst im Zentrum. Pressemitteilung v. 5.12.01. Kultusministerkonferenz: Bonn 2001

KÖLLER, O., TRAUTWEIN, U. (2001): Evaluation mit TIMSS-Instrumenten. Untersuchungen in der 8. Jahrgangsstufe an fünf Gesamtschulen. In: Die Deutsche Schule, 93. Jg., H. 2, 167-185

KRAUS, J.: Apokalyptische Inszenierung und dilettantischer Leichtsinn. In: DEUTSCHER LEHRERVERBAND (DL) AKTUELL Zur Diskussion um "TIMSS/III" 1998 Download: www.lehrerverband.de/timss.htm

KRAUS, J.: BILDUNGSPOLITISCHE SchlussFOLGERUNGEN AUS DER TIMSS-STUDIE DEUTSCHLAND – THESEN. In: LIST, J: TIMSS Mathematisch-Naturwissenschaftliche Kenntnisse Deutscher Schüler auf dem Prüfstand. Köln 1998

KRAUSS, B. (2001): Eine nationale Katastrophe. Deutsche Schulen im PISA-Test. In: Süd-deutsche Zeitung v. 6.12.01

KRUEGER; A.B., WHITMORE,D.M.: The effect of attending a small class in the early grades on college test taking an middle school test results: Evidence from Project STAR. National Bureau of Economic Research Working Paper No. 7656. Cambridge 2000

KUNTER, M., SCHÜMER, G., ARTELT, C., BAUMERT, J., KLIEME, E., NEUBRAND, M., PRENZEL, M., SCHIEFELE, U., SCHNEIDER, W., STANAT, P., TILLMANN, K.-J., WEIß, M.: PISA 2000 Dokumentation der Erhebungsinstrumente
Download: [www.mpib-berlin.mpg.de, pisa](http://www.mpib-berlin.mpg.de/pisa)

LANKES, E.M., BOS, W., MOHR, PLABMEIER, N., SCHWIPPERT, K., SIBBERNS, H., VOSS, A.: Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU) und ihrer Erweiterung um Mathematik und Naturwissenschaften (IGLU-E). In: Erste Ergebnisse aus IGLU. Münster 2003

LANKES, E.M., BOS, W., MOHR, PLABMEIER, N., SCHWIPPERT, K. : Lehr- und Lernbedingungen in den Teilnehmerländern. In: Erste Ergebnisse aus IGLU. Münster 2003

LEHMANN, R. H., u.a.: Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Behörde für Schule, Jugend und Berufsbildung: Hamburg. 1997

LEHMANN, R. H., u.a. (1999): Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern an Hamburger Schulen - Klasse 7. Behörde für Schule, Jugend und Berufsbildung: Hamburg. Lind, G. (2002): PISA: Kritische Fragen sind nötig. Minist. Für Bildung u. Wissenschaft (Baden Württemberg), H. 1. 2002

LEHMANN, R., PEEK, R., GÄNSFUß, R., HUSFELD, V.: Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9. Hrsg. von der Behörde für Bildung und Sport. Hamburg 2002

LEHMANN, R., PEEK, R., GÄNSFUß, R.: Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Hrsg. von der Behörde für Schule, Jugend und Berufsbildung, Hamburg 1997

LEVESQUE, R., SPSS® Programming and Data Management. A Guide for SPSS® and SAS® Users. Chicago. Download: <http://www.spss.com>

LEVIN, H. M. Cost-effectiveness in evaluation research. In M. Guttentag & E. Struening (Hrsg.), Handbook of evaluation research (Vol. 2, pp. 89-122). Beverly Hills, CA: Sage. 1975

LEVIN, H. M.: Cost analysis. In N. Smith (Ed.), New techniques for evaluation. Beverly Hills, CA: Sage. 1981

LEVIN, H. M.: Cost-effectiveness and educational policy. Educational Evaluation and Policy Analysis, 10(1), 51-69. 1988

LEVIN, H. M.: Cost-effectiveness at quarter century. In M. W. McLaughlin & D. C. Phillips (Hrsg.), Evaluation and education at quarter century (pp. 188-209). Chicago: University of Chicago Press. 1991

LEVIN, H. M.: Cost-benefit analysis. In M. Carnoy (Ed.), International encyclopedia of economics of education (2nd ed., pp. 360-364). Oxford, UK: Pergamon. 1995

LEVIN, H. M.: The cost to the nation of inadequate education [Report prepared for the Select Senate Committee on Equal Educational Opportunity, 92nd Congress, 2nd Session]. Washington, DC: Government Printing Office. 1972.

LEVIN, H.M., McEWAN, P.J: Cost Effectiveness Analysis, Thousand Oaks, California 2000

LEVIN, H.M., GLASS, G.V. und MEISTER, G.R.: Cost-effectiveness of computer-assisted instruction. Evaluation Review. 11/1, 1987

LEVIN, H., WOO, L: An evaluation of the costs of computer-assisted instruction. Economics of Education Review, 1/1, 1981

LEWIS, D.R., STOCKDILL, S.J. und TURNER, T.C.: Cost-effectiveness of micro-computers in adult basic reading. Adult Literacy and Basic Education, 14/2, 1990

LIENERT, G.A., RAATZ, U.: Testaufbau und Testanalyse. Weinheim 1998

LIND, G.: Zur Messung sozialer Motivationen in der BIJU-Studie. In: Gewerkschaft Erziehung und Wissenschaft (Hrsg.): Messung sozialer Motivationen – eine Kontroverse. Frankfurt, M. 2000

LINNAKYLÄ, P.: Arbeitsgruppe Internationale Vergleichsstudie: Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. Berlin 2003

ROBITZSCH, A.: Modellierung von Stichprobenausfall in PISA, In: BAUMERT, J., ARTELT, C., KLIEME, E., NEUBRAND, M., PRENZEL, M., SCHIEFELE, U., SCHNEIDER, W., SCHÜMER, G., STANAT, P., TILLMANN, K.-J.: 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 Download: www.mpib-berlin.mpg.de

MARKSTAHLER, J., STEFFENS, U. (1999): Das OECD-Projekt "PISA" – ein Beitrag zur Qualitätssicherung von Schule. In: Pro Schule (Hessisches Landesinstitut für Pädagogik: Wiesbaden), 1, 99, 16-20. KLAFKI, W.: Verändert Schulforschung die Schulwirklichkeit. In: Zeitschrift für Pädagogik, 30, 1983

MARTIN, M.O., KELLY, D.L. (Hrsg.): TIMSS Technical Report 1996 Vol. 1 Download: www.bc.edu

MERKENS H.: Internationale Herausforderungen an Bildung im 21. Jahrhundert download: <http://www.phil-fak.uni-duesseldorf.de/oasien/oasien/japan/Bildungssymposium/Reden+Folien/Merkens-dt.htm>

MERKENS, H.: Einführung in die Statistik für Pädagogen, Trier 1975

MESSENER R.: PISA und Allgemeinbildung In: Zeitschrift für Pädagogik 3/2003

MONK, D. H., & KING, J. A. Cost analysis as a tool for education reform. In S. L. Jacobson & R. Berne (Eds.), Reforming education: The emerging systemic approach (pp. 131-150). Thousand Oaks, CA: Corwin Press. 1993.

MONK, D. H. The costs of pupil performance assessment: A summary report. Journal of Education Finance, 20, 363-371. 1995

MULLIS, I.V.S., MARTIN, M.O., BEATON, A.E., GONZALEZ E.J., KELLY, D.L., SMITH, T.A.: Mathematics Achievement in the Primary School Years (TIMSS), Boston College 1997 <http://www.csteep.bc.edu/timss>

NAYLOR, F.: OECD: The Trojan Horse Within. Campaign for Real Education Download <http://www.credyork.freereserve.co.uk/index.htm>

NEUBRAND, M., KLIEME, E.: Mathematische Grundbildung. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. Opladen 2001

NIEMIEC; R.P., SIKORSKI, M.F. und WALBERG, H.J.: Comparing the cost-effectiveness of tutoring and computer-based instruction. Journal of Educational Computing Research. 5/4, 1989, S. 395-407

OECD (Hrsg.) (im Auftrag Bundesministerium für Bildung und Forschung) : Lernen für das Leben – Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000

OEHLERT, P.: Der Leistungsvergleich zwischen traditionellen Schulen und Gesamtschulen in Nordrhein-Westfalen. In: Zeitschrift für Pädagogik, 26, 1980,

PEISERT, H.: Soziale Lage und Bildungschancen in Deutschland, 1967

PICHT, G.: Die deutsche Bildungskatastrophe. Olten, Freiburg 1964

POKROPP, F.: Stichproben: Theorie und Verfahren. München 1996

POHLENZ, F.L.: Educational Attainment as a Proxy for Human Capital in Models of Growth and Development: A Critical Survey. (Diss) Bamberg 2000

PRAIS, S. J. Oxford Review of Education, Vol. 29, June, 2003

PRENZEL, M., GEISER H., LANGEHEINE, R., LOBEMEIER, K.: Naturwissenschaftliche Kompetenz am Ende der Grundschulzeit: Vergleiche zwischen einigen Ländern der Bundesrepublik Deutschland,. In: BOS, W., LANKES, E.M., PRENZEL, M., SCHWIPPERT, K., VALTIN, R., WALTHER, G. (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II) Münster 2004

PRENZEL, M., ROST, J., SENKBEIL, M., HÄUßLER, P., KLOPP, A.: Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

PRENZEL M., BAUMERT J., BLUM W., LEHMANN R., LEUTNER D., NEUBRAND M., PEKRUN R., ROLFF H.G., ROST J. und SCHIEFELE U. (Hrsg.) PISA-Konsortium Deutschland PISA 2003 Ergebnisse des zweiten internationalen Vergleichs Zusammenfassung

PREUß, O.: Illusionen der Schulforschung. In: Jahrbuch für Lehrer, Bd. 7., Reinbek 1982

RASCHERT, J.: Gesamtschule: ein gesellschaftliches Experiment. Stuttgart 1974

RICE, J. K.. Cost analysis in education: Paradox and possibility. Educational Evaluation and Policy Analysis, 19(4),309-317. 1997

ROEDER, P.M.: Die Sekundarstufe I im Schulsystemvergleich. In: Zeitschrift für Pädagogik, 26, 1980,

ROEDER, P.M.: Gute Schule unter schlechten Rahmenbedingungen – Fallstudien an Berliner Hauptschulen. In: Steffens, U., Bargel, T. (Hrsg.): Erkundungen zur Wirksamkeit und Qualität von Schule. Beiträge aus dem Arbeitskreis Qualität von Schule. Heft 1. Wiesbaden-1987

ROST, J.: Testtheorie Testkonstruktion. Bern 1996

ROTHENBER, J.: Cost-benefit analysis: a methodological exposition. In M. Guttentag & E. Struening (Hrsg.), Handbook of evaluation research (Vol. 2, pp. 89-122). Beverly Hills, CA: Sage. 1975

ROYL, W., LIND, G., RÖPCKE, B., VOGEL-KRAHFORST, E.: Lernerfolgsmessung im Rahmen der wissenschaftlichen Begleituntersuchungen zu den schleswig-holsteinischen Gesamtschulversuchen. Hrsg. vom Kultusministerium des Landes Schleswig-Holstein. Kiel 1978

SCHLITTEGEN, R.: Statistik. München 1995

SCHLÖMERKEMPER, J.: Schülerleistungen in integrierten und gegliederten Systemen. In: F. Bohnsack (Hrsg.): Kooperative Schulen. Weinheim 1978

SCHNABEL, SCHWIPPERT, K. : Einflüsse sozialer und ethnischer Herkunft beim Übergang in die Sekundarstufe II und den Beruf. In: Baumert, J., Bos, W., Lehmann (Hrsg.) TIMSS, III Dritte Internationale Mathematik- und Naturwissenschaftsstudie, Opladen 2000, Band1

SCHORB, A. (Hrsg.): Schulversuche mit Gesamtschulen in Bayern. Ergebnisse der wissenschaftlichen Begleitung. Stuttgart 1977

SCHÜMER, G., WEIß, M., STEINERT, B., BAUMERT, J., TILLMANN, K.-J., MEIER U.: Institutionelle Bedingungen schulischen Lernens im internationalen Vergleich. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen 2001

SCHÜMER, G., WEIß, M., STEINERT, B., BAUMERT, J., TILLMANN, K.-J., MEIER: Lebens- und Lernbedingungen von Jugendlichen. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

SIBBERNS, H., BAUMERT, J.: Stichprobenziehung und Stichprobengewichtung. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich Opladen 2001

SMITH, N. L., & SMITH, J. K. Statelevel evaluation uses of cost analysis: A national descriptive survey. In J. S. Catterall (Ed.), Economic evaluation of public programs (pp. 83-97). San Francisco: Jossey-Bass. 1985

SPRENGER, U.: Die unterschiedlichen Fördereffekte der verschiedenen Schulformen. In: Bildung aktuell, Heft 5, 2000 (51. Jg.)

STANAT, P., KUNTER, M.: Geschlechterunterschiede in Basiskompetenzen. In: Deutsches PISA-Konsortium (Hrsg.) PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen 2001

STECHER, B. M., & KLEIN, S. P.. The cost of science performance assessments in large-scale testing programs. Educational Evaluation and Policy Analysis, 19(1),1-14. 1997

TERHART, E.: Nach PISA. Hamburg 2002

TILLMANN, K.-J., MEIER, U.: Schule, Familie und Freunde – Erfahrungen von Schülerinnen und Schülern in Deutschland. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000 Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen 2001

WALTHER, G., GEISER, LANGEHEINE, LOBEMEIER: Mathematische Kompetenzen am Ende der vierten Jahrgangsstufe in einigen Ländern der Bundesrepublik Deutschland. In: BOS, W., LANKES, E.M., PRENZEL, M., SCHWIPPERT, K., VALTIN, R., WALTHER, G. (Hrsg.): IGLU Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich (II). Münster 2004

WATERMANN, R., BAUMERT, J.: Mathematische und naturwissenschaftliche Grundbildung beim Übergang von der Schule in den Beruf. In: BAUMERT, J., BOS, W., Lehmann (Hrsg.) TIMSS, III Dritte Internationale Mathematik- und Naturwissenschaftsstudie, Opladen 2000, Band 1

WEIß, M.: Effizienzforschung im Bildungsbereich. Diss. Berlin, 1981

WEISBROD, B. A.. Preventing high school dropouts. In R. Dorfman (Hrsg.), Measuring benefits of government investments (S. 117-148). Washington, DC: Brookings Institution. 1965

WEISHAUPT, H., ZEDLER, P.: Aspekte der aktuellen Schulentwicklung in den neuen Ländern. In: Rolff, H.-G., Bauer, K.O., KLEMM, K., Pfeiffer, H., Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung. Band 8. Weinheim und München 1994

WEIß, WEISHAUPT (Hrsg.): Bildungsökonomie und Neue Steuerung. München 2000

WÖBMANN, L.: How Central Exams Affect Educational Achievement: International Evidence from TIMSS and TIMSS-Repeat* - Paper prepared for the conference Taking Account of Accountability: Assessing Politics and Policy John F. Kennedy School of Government Harvard University June 10 – 11, 2002

WÖBMANN, L.: Familiärer Hintergrund, Schulsystem und Schülerleistungen im internationalen Vergleich, Aus Politik und Zeitgeschichte (B 21-22/2003)

WÖBMANN, L.: Schooling and the Quality of Human Capital. Kieler Studien. Berlin 2002

WÖBMANN, L., WEST, M.R.: Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS, Kiel Working Paper No.1099, 2002