

Helmke, Andreas; Helmke, Tuyet; Schrader, Friedrich-Wilhelm; Wagner, Wolfgang; Klieme, Eckhard; Nold, Günter; Schröder, Konrad

Wirksamkeit des Englischunterrichts

Klieme, Eckhard [Hrsg.]: Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim u.a. : Beltz 2008, S. 382-397

urn:nbn:de:0111-opus-31620

in Kooperation mit:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen

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

Mit dem Gebrauch von pedocs und der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Kontakt:

peDOCS

Deutsches Institut für Internationale Pädagogische Forschung (DIPF)

Informationszentrum (IZ) Bildung

Schloßstr. 29, D-60486 Frankfurt am Main

eMail: pedocs@dipf.de

Internet: www.pedocs.de

Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Eckhard Klieme / Günter Nold / Konrad Schröder

32 Wirksamkeit des Englischunterrichts

Viele Themen der Unterrichtsqualität werden in der deutschen Englischdidaktik (ebenso wie in der internationalen Diskussion zu „English as a Foreign Language“) seit langem ausführlich und vielfach kontrovers diskutiert (Bausch/Christ/Krumm 2002; Butzkamm 1998, 2004; Klippel 2004; Timm 1998). Sondiert man aussagekräftige empirischen Studien zur Qualität des *Englischunterrichts*, dann wird man allerdings – zumindest im deutschen Sprachraum – kaum fündig: In internationalen Leistungsstudien dominieren mathematische und naturwissenschaftliche Inhalte. Zu den wenigen einschlägigen large-scale-Studien zählen die betagte IEA Study „The teaching of English as a Foreign Language“ (Lewis/Massad 1975), die Konstanzer Schulleistungsstudien der Forschungsgruppe Fend (Fend 1984, 1998; Haenisch 1979; Lukesch 1979) sowie die Studien von Achtenhagen (Wienold u.a. 1985), BIJU (Gruehn 2000), TOSCA (Köller/Trautwein 2004) und dem Schulinspektorat Graubünden (2004). Allerdings sind diese Studien mit Ausnahme von BIJU querschnittlich angelegt, so dass die Erkenntnismöglichkeiten – sofern überhaupt der Unterricht erfasst wurde – beschränkt sind.

Im Prinzip käme auch die empirische Unterrichtsforschung zum *Zweitsprachenerwerb* in anderen Ländern als Quelle für die Entwicklung von Instrumenten und Hypothesen in Frage. Denn für die Frage der Unterrichtsqualität ist es eigentlich nicht ausschlaggebend, ob sich der Fremdsprachenunterricht auf „English as a Foreign Language“ in Deutschland oder in Vietnam, auf die Fremdsprache Spanisch in den USA oder Deutsch in Polen bezieht. Aber auch hier ist die Forschungslage sehr eingeschränkt (Judd/Tan/Walberg 2001).

Am ergiebigsten für dieses Kapitel sind die Studien, die – mit fachdidaktischem Hintergrund – den Unterricht mit Hilfe standardisierter Verfahren, insbesondere durch *Unterrichtsbeobachtungen*, erfasst und mit abhängigen Variablen in Beziehung gebracht haben. Hierzu gibt es speziell im angloamerikanischen Raum eine große Zahl von Instrumenten und Studien (Waxman/Tharp/Hilberg 2004). Für dieses Kapitel relevant sind vor allem die Studien, die das Verfahren COLT (Communicative Orientation of Language Teaching) eingesetzt haben (Spada/Fröhlich 1995). Die dort entwickelten Kategorien bilden einen wesentlichen Teil des Fundamentes der DESI-Videostudie und regten auch die Instrumentenentwicklung für die fragebogenbasierte Hauptstudie an. Dies betrifft vor allem diejenigen Elemente, die sich auf die Förderung kommunikativer Kompetenzen von Schülerinnen und Schülern beziehen. Guter Englischunterricht wäre aus dieser Perspektive demnach wie folgt charakterisierbar:

- Längere sprachliche Schüleräußerungen anstatt von Ein-Wort-Sätzen

- Beteiligung von Schülerinnen und Schülern an der Gestaltung und den Inhalten des Unterrichts
- Verwendung von Englisch als Verkehrssprache im Unterricht
- Ermöglichung der Selbstkorrektur von Schülerfehlern
- Balance von „accuracy“ und „fluency“

Auf der Grundlage der in der DESI-Studie verwendeten Kategorien und des dort verwendeten Designs (Zwei-Punkt-Erhebung) lässt sich das in Kapitel 26 skizzierte allgemeine Angebots-Nutzungs-Modell (Rahmenmodell der Unterrichtswirksamkeit) wie in Abbildung 32.1 dargestellt spezifizieren.

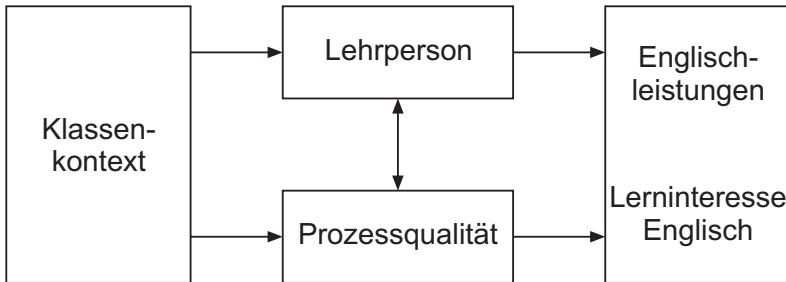


Abbildung 32.1: Spezifikation des Angebots-Nutzungs-Modells für den Englischunterricht.

Dieses Modell ist gegenüber dem umfassenderen Modell in Kapitel 26 stark vereinfacht und umfasst lediglich diejenigen Variablen, die in diesem Bericht tatsächlich analysiert werden.

Klassenkontext: Dies sind Merkmale der Klassenzusammensetzung.

Lehrperson: Ausgewählt wurden Merkmale, die einen engen Bezug zur beruflichen Tätigkeit und zur Schülerleistung erwarten lassen.

Prozessqualität: Dieser zentrale Block umfasst verschiedene, mit Hilfe dimensionsanalytischer Methoden voneinander unterscheidbarer Aspekte, die im folgenden Abschnitt detailliert beschrieben werden.

Zielkriterien: Als Zielkriterien fungieren bei den folgenden Auswertungen primär diejenigen beiden Aspekte der englischen Sprachkompetenz, zu denen von beiden Messzeitpunkten zu Beginn und am Ende der neunten Jahrgangsstufe Daten vorliegen, nämlich das Hörverstehen und die Fähigkeit zur Textrekonstruktion (C-Test), (vgl. Kapitel 11 und 14). Außerdem berücksichtigen wir das Lerninteresse im Fach Englisch (Beispielitems: „Ich finde das Fach Englisch cool“; „In meiner Freizeit beschäftige ich mich manchmal mit Englisch über die Hausaufgaben hinaus“; „Ich lege großen Wert darauf, in Englisch gut zu sein“).

32.1 Unterrichtspraxis, Lehrpersonmerkmale und Unterrichtswirksamkeit

Unterrichtspraxis aus Lehrersicht. Wir verzichten hier auf eine komplette Darstellung der Korrelationen zwischen Unterrichtsmerkmalen und Zielkriterien, weil

sich die bei den Lehrern erfragten Aspekte der Alltagspraxis für die Entwicklung der Englischkompetenzen als weitgehend irrelevant herausgestellt haben. Es zeigen sich zwar teilweise sehr hohe einfache Korrelationen des Leistungsniveaus (z.B. mit der Verwendung der deutschen Sprache: $-.67, p < .001$ mit Leistungsdifferenzierung: $-.36, p < .001$ und mit der Verwendung literarischer Texte: $.36, p < .001$). Partialisiert man aber Bildungsgang, die Ausgangsleistungen zu Beginn der neunten Jahrgangsstufe, kognitive Grundfähigkeit, Geschlechts- und Sprachzusammensetzung aus, dann verbleiben lediglich die folgenden statistisch signifikanten, aber zahlenmäßig nur geringen partiellen Korrelationen (für das Lerninteresse im Fach Englisch zeigen sich keine statistisch signifikanten Korrelationen):

- für den C-Test: Übergehen kommunikativer Fehler, $r = .03, p < .05$
- für den Hörverstehenstest: Übergehen grammatischer Fehler, $r = -.08, p < .01$

Merkmale der Expertise und Erfahrung von Lehrkräften. Ergiebiger ist die Befundlage bei den in Kapitel 30 beschriebenen Personmerkmalen. Hierzu berichten wir im Folgenden in Gestalt je eines Strukturgleichungsmodells Ergebnisse zu den beiden Zielkriterien „Hörverstehensleistung“ und „Lerninteresse im Fach Englisch“. Alle Schülermerkmale in den Abbildungen 32.2 und 33.3 beziehen sich auf die Schulklassenebene. Zugrunde liegt ein zweiebenenanalytisches Modell (wir berichten nur die Ergebnisse auf Klassenebene); auspartialisiert wurden Bildungsgang, Sprach- und Geschlechtszusammensetzung.

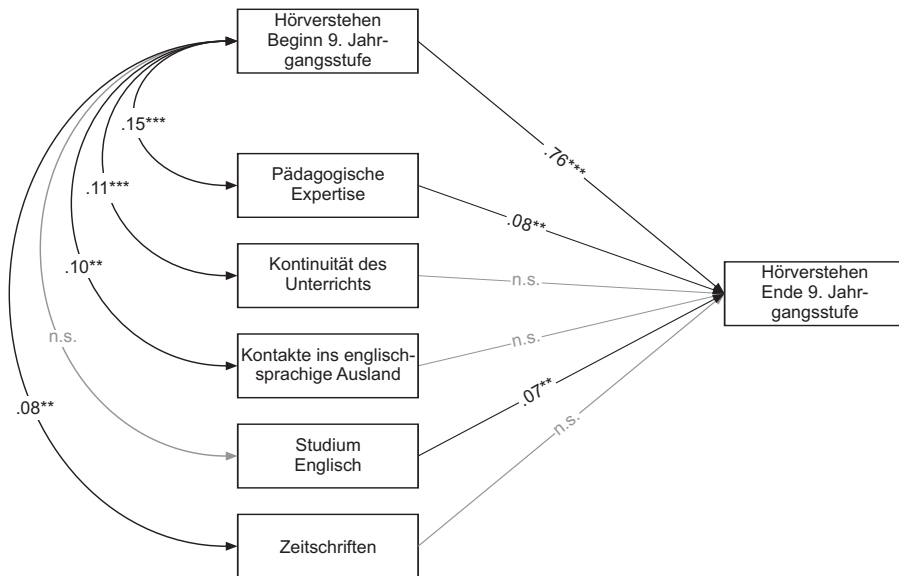


Abbildung 32.2: Lehrpersonenmerkmale und Kompetenzzuwachs im Hörverstehen Englisch in der neunten Jahrgangsstufe.

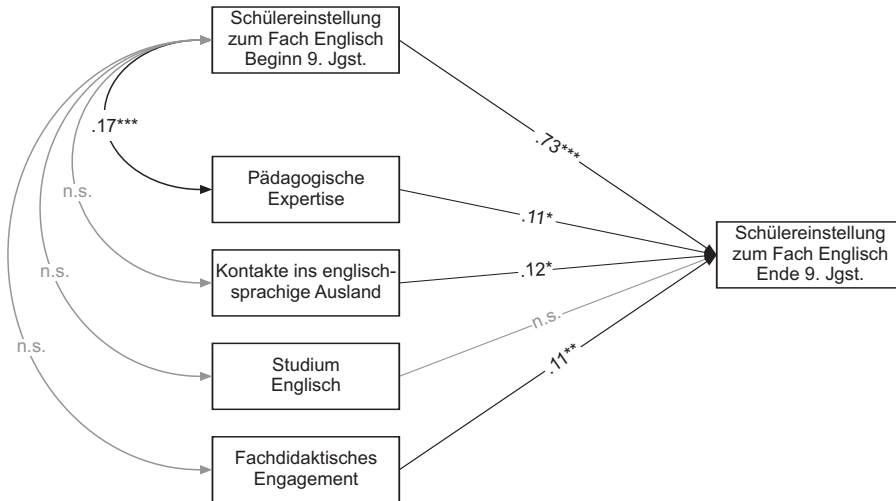


Abbildung 32.3: Lehrpersonenmerkmale und Zuwachs des Lerninteresses im Fach Englisch in der neunten Jahrgangsstufe.

Die Ergebnisse unterstreichen zunächst die bereits zuvor berichtete hohe Stabilität sowohl der Hörverstehensleistung als auch des Lerninteresses im Fach Englisch. Man kann es allerdings auch so sehen: Ein Pfadkoeffizient von $.76$ (Pretest auf Posttest beim Hörverstehen) bedeutet, dass knapp die Hälfte der Leistungsunterschiede ($.76 \times .76 = .58$) zwischen Klassen nicht festgelegt sind und somit Spielräume für unterrichtliche Einflüsse eröffnen. Lehrkräfte mit Ausbilder- oder Mentorstatus („pädagogische Expertise“) erreichen bei beiden Zielkriterien günstigere Ergebnisse. Eine positive motivationale Entwicklung ist insbesondere dann zu verzeichnen, wenn Lehrkräfte fachdidaktisches Interesse an den Tag legen (hier: Bezug einschlägiger Zeitschriften) und Kontakte ins englischsprachige Ausland pflegen. Dieses Ergebnismuster bestätigt empirisch bisher nicht untermauerte Vermutungen der allgemeinen Unterrichtsforschung (Rolle von Engagement, fachdidaktischer Expertise und von Auslandsaufenthalten; Ehrenreich 2004¹).

32.2 Beschreibung und Dimensionalität der Unterrichtsqualität aus Schülersicht

Basis der folgenden Analysen sind 15 Skalen im Schülerfragebogen (vgl. Tabelle 32.1), die unterschiedliche Facetten der Unterrichtsqualität aus Schülersicht abdecken. Sie sollen anhand der in Tabelle 32.1 dargestellten Beispielitems veranschaulicht werden.

¹ Allerdings ging es in dieser Arbeit nur um den Ertrag des Fremdsprachenassistentenjahres.

Tabelle 32.1: Skalen zur Unterrichtsqualität: Konstrukte, Beispielitems, Mittelwerte und Standardabweichungen, Reliabilitäten (Cronbachs α) und Intra-Klassen-Korrelationen (vgl. DESI-Skalenhandbuch).

Konstrukt	Beispielitem(s)	M	SD	α	IKK
Strukturiertheit	Am Ende der Stunde fasst mein Englischlehrer das Wichtigste zusammen.	2.25	0.72	.77	.15
Positive Fehlerkultur	Wenn ich im Englischunterricht Fehler mache, bespricht mein Englischlehrer diese mit mir auf eine Art und Weise, dass es mir wirklich etwas bringt.	2.80	0.55	.89	.20
Unterrichtsklima	Ich komme mit meinem Englischlehrer gut aus.	2.66	0.74	.83	.16
Motivierung thematisch	Mein Englischlehrer kann auch trockene Themen wirklich interessant machen.	2.40	0.82	.79	.20
Motivierung instrumentell	Mein Englischlehrer weist darauf hin, wie nützlich Englisch im Alltag ist.	2.79	0.75	.83	.12
Schülerorientierung	Mein Englischlehrer ermuntert mich, meine persönlichen Erfahrungen in den Unterricht einzubringen.	2.65	0.66	.87	.22
Unterstützung	Mein Englischlehrer achtet darauf, dass ich im Unterricht mitkomme.	2.80	0.70	.84	.17
Verständlichkeit	Mein Englischlehrer drückt sich klar und deutlich aus.	2.82	0.70	.76	.19
Wichtigkeit von Kommunikation	Wie wichtig ist deinem Englischlehrer, dass ihr euch auf Englisch unterhalten könnt?	3.04	0.64	.82	.19
Wichtigkeit von Korrektheit	Wie wichtig ist deinem Englischlehrer, dass eure Aussprache korrekt ist / eure Sätze beim Sprechen grammatikalisch korrekt sind?	3.16	0.62	.87	.17
Integration der Hausaufgaben	Wird im Unterricht auf den Inhalt der Hausaufgaben eingegangen?	3.61	0.66	.83	.26
Management	Mein Englischlehrer sorgt dafür, dass die Schüler die ganze Stunde über aufpassen.	2.64	0.81	.77	.22
Störungsfreiheit	Mein Englischlehrer muss lange warten, bis Ruhe eintritt.	2.75	0.69	.87	.23
Tempodruck	Wenn mein Englischlehrer mir eine Frage gestellt hat, lässt er / sie mir genug Zeit zum Überlegen.	2.03	0.76	.79	.10
Verwendung der deutschen Sprache	Denk einmal an deinen Englischunterricht in der vergangenen Zeit. Wie oft wird im Englischunterricht die deutsche Sprache verwendet? bei Störungen der Disziplin in der Klasse	2.79	0.67	.89	.29

Anmerkung: In der Tabelle wird aus Platzgründen die Formulierung „mein Englischlehrer“ verwendet; im Fragebogen ist dagegen selbstverständlich von „mein Englischlehrer / meine Englischlehrerin“ die Rede.

Zusätzlich setzten wir ausgewählte Items zum Thema „Umgang mit Schülerfehlern“ aus Schülersicht ein, die sich nicht zu Skalen zusammenfassen lassen (vgl. Kapitel 31), und zwar den Umgang mit grammatischen und mit Aussprachefehlern (*Ignorieren / Gelegenheit zur Selbstkorrektur geben / Korrigieren*).

Zur Klärung der Dimensionalität führten wir explorative Faktorenanalysen durch. Es wurden alle in Skalenform vorliegenden Schülerperzeptionen des Englischunterrichts einbezogen, mit Ausnahme der auf spezifische Kompetenzen und Testdomänen bezogenen Lerngelegenheiten. Letztere werden je nach Zielkriterium als weitere Erklärungsvariable in die Modelle einbezogen. Tabelle 32.2 berichtet die Interkorrelationen der sechs Merkmale.

Tabelle 32.2: Signifikante Interkorrelationen der auf Schülerangaben basierenden Unterrichtsmerkmale². Nicht signifikante Korrelationen ($p \geq .05$) sind nicht dargestellt.

Unterrichtsmerkmale	Merkmale				
	(2)	(3)	(4)	(5)	(6)
(1) Prozessqualität					
Positive Fehlerkultur					
Unterrichtsklima					
Motivierung thematisch	.76***	.73***	-.54***	-.27**	.68***
Schülerorientierung					
Unterstützung					
Verständlichkeit					
(2) Anspruchsniveau					
Wichtigkeit von Kommunikation					
Wichtigkeit von Korrektheit		.72***	-.61***	-.47***	.26**
Motivierung instrumentell					
Integration der Hausaufgaben					
(3) Klassenführung					
Management			-.40**	-.40**	.46***
Störungsfreiheit					
(4) Tempodruck				.27**	
(5) Verwendung der deutschen Sprache					
(6) Strukturiertheit					

Anmerkungen: Signifikanzniveaus: * $p < .05$, ** $p < .01$, *** $p < .001$ (zweiseitige Tests).

Es ergeben sich sechs Faktoren (die zugehörigen Subskalen tauchen unter dem fett gedruckten Namen des Faktors auf), wobei nur die drei ersten Faktoren mehrere Skalen umfassen: Prozessqualität, Anspruchsniveau und Effizienz der Klassenführung. Die drei restlichen Faktoren sind identisch mit einzelnen Skalen (Tempodruck, Verwendung der deutschen Sprache und Strukturiertheit).

32.3 Zur Rolle von Schulklassenunterschieden

Analysen zur Wirksamkeit des Unterrichts sind nur dann aussichtsreich, wenn sich Schulklassen im Hinblick auf die erfassten Unterrichtsmerkmale wie auch die

2 Die konfirmatorische Faktorenanalyse erfolgte in einem Zwei-Ebenen-Modell, wobei Faktoren lediglich auf der Klassenebene modelliert wurden. Auf der Individualebene wurden die Korrelationen sämtlicher Skalen untereinander modelliert. Modellfit: $CFI = .985$; $TLI = .958$; $RMSEA = .023$; $SRMR_{between} = .096$.

Zielkriterien überhaupt hinlänglich voneinander unterscheiden. Das Ausmaß der Schulklassenunterschiede lässt sich durch die *Intraklassen-Korrelation* charakterisieren. Dieses Maß setzt die Varianz zwischen Klassen zu der Gesamtvarianz in Beziehung. Ein hoher IKK-Wert bedeutet, dass sich ein großer Prozentsatz der Gesamtvarianz auf Unterschiede zwischen Klassen zurückführen lässt, mit anderen Worten: dass es bei dem entsprechenden Merkmal große Unterschiede zwischen Klassen gibt, deren Ursache in Unterschieden in der Klassenzusammensetzung, der Unterrichtsqualität oder in beiden liegen können. Wegen der Konfundierung von Leistungsniveau und Bildungsgang erfolgt die Analyse global (bildungsgang-übergreifend) sowie separat für die vier Bildungsgänge.

Tabelle 32.3: *Englischleistung und Leistungszuwächse: Unterschiede zwischen Klassen („Intraklassen-Korrelationen“) nach Bildungsgang.*

Leistung: Niveau	9. Jgst.	HS	RS	IGS	GY	Gesamt
C-Test	Beginn	.23	.37	.55	.28	.65
	Ende	.24	.33	.50	.29	.64
Hörverstehen	Beginn	.36	.55	.64	.46	.71
	Ende	.40	.57	.62	.48	.71
Leseverstehen	Ende	.19	.37	.48	.32	.58
Soziopragmatik	Beginn	.22	.39	.65	.28	.67
Grammatik	Ende	.23	.44	.52	.40	.72
Textproduktion	Ende	.28	.39	.60	.34	.68
Englischleistung (global)	Ende	.29	.48	.67	.40	.77
Leistung: Zuwachs						
C-Test		.12	.08	.30	.09	.12
Hörverstehen		.59	.50	.36	.65	.58
Lerninteresse						
Lerninteresse	Beginn	.04	.05	.05	.07	.08
	Ende	.03	.06	.05	.09	.09
Zuwachs		.03	.03	.08	.03	.03

Die in Tabelle 32.3 berichteten Ergebnisse zeigen: Während sich Schulklassen – abgesehen vom Bildungsgang IGS – hinsichtlich des Zuwachses beim *C-Test* nur wenig voneinander unterscheiden, gibt es erhebliche Klassenunterschiede beim Zuwachs für das *Hörverstehen*. Das Erklärungspotenzial des Unterrichts und von Merkmalen der Klassenzusammensetzung hängt also deutlich von dem jeweils betrachteten Zielkriterium ab. Bei einem integrativen und globalen Instrument wie dem *C-Test*, dessen erfolgreiche Bearbeitung ein breites Spektrum unterschiedlicher Kompetenzen erfordert (insbesondere Wortschatz, Grammatik und Lesestrategien), spielen Unterschiede in der Unterrichtsqualität offensichtlich eine geringere Rolle als beim Hörverstehen. Beim Lerninteresse sind die Klassenunterschiede deutlich geringer (ohne Tabelle). Die folgenden Analysen werden sich deshalb schwerpunktmäßig

mit der Rolle der Unterrichtsqualität für den Kompetenzzuwachs beim Hörverstehen beschäftigen.

32.4 Unterrichtsqualität und Englischleistung

Welche Rolle spielen Merkmale der Unterrichtsqualität aus Schülersicht für die Entwicklung sprachlicher Kompetenzen in der neunten Jahrgangsstufe? Die Merkmale des Unterrichts in Tabelle 32.4 entsprechen der Anordnung in Tabelle 32.2.

Tabelle 32.4: Korrelationen der Unterrichtsvariablen mit dem C-Test und dem Hörverstehenstest am Ende der neunten Jahrgangsstufe (Klassenebene). Nicht signifikante Korrelationen ($p \geq .05$) sind nicht dargestellt.

	C-Test			HV-Test		
	einfache Korrelation	Partialkorrelation mit dem Posttest	Partialkorrelation mit dem Zuwachs	einfache Korrelation	Partialkorrelation mit dem Posttest	Partialkorrelation mit dem Zuwachs
1 Prozessqualität						
Positive Fehlerkultur	.24***			.25***	.07**	.17**
Unterrichtsklima					.06*	.13*
Motivierung thematisch					.05*	.12*
Schülerorientierung	.23***	.02*	.12*	.19***		
Unterstützung						
Verständlichkeit	.26***	.02*	.14*	.22***	.06**	.15**
2 Anspruchsniveau						
Wichtigkeit: Kommunikation	.63***			.64***	.08**	.19**
Wichtigkeit: Korrektheit	.58***			.58***	.08**	.20**
Motivierung instrumentell	.23***	.04**	.21**	.25***	.06*	.15*
Integration der Hausaufgaben	.48***			.50***	.07**	.16**
3 Klassenführung						
Management	.15*			.18**	.06*	.14*
Störungsfreiheit	.35***			.34***	.08**	.18**
4 Tempodruck	-.60***			-.52***		
5 Verwendung deutscher Sprache	-.44***			-.53***	-.09***	-.22***
6 Strukturiertheit	-.39***			-.40***		

Anmerkung: Mehrebenenkorrelationen mit Mplus; Kontrollvariablen: Leistung zu Beginn der neunten Jahrgangsstufe, Geschlecht, Erstsprache, kognitive Grundfähigkeit, Bildungsgang. Signifikanzniveaus (jeweils einseitiger Test): * $p < .05$, ** $p < .01$, *** $p < .001$.

Die Ergebnisse machen deutlich:

a) Erwartungsgemäß gibt es markante Unterschiede zwischen den unbereinigten und den bereinigten Korrelationen. Beispiel „Tempodruck“: In leistungsstärkeren Klassen (dies gilt sowohl für den C-Test als auch für das Hörverstehen) kommt Überforderung durch unangemessen hohe Geschwindigkeit seltener vor (Korrelationen in Höhe von $-.60$ und $-.52$). Dagegen leistet der Tempodruck keinen Erklärungsbeitrag für Klassenunterschiede im Kompetenzzuwachs im Laufe der neunten Klassenstufe.

b) Beim *C-Test* finden sich lediglich drei Unterrichtsmerkmale, denen nach Bereinigung um Bildungsgang- und Kontextunterschiede überhaupt ein signifikanter Erklärungswert zukommt: Motivierung, Verständlichkeit und Schülerorientierung. Dies war in Anbetracht der vergleichsweise geringen Klassenunterschiede im Zuwachs (vgl. Tabelle 32.3) zu erwarten gewesen. Bei der Suche nach unterrichtlichen Bedingungen des Zuwachses beim *Hörverstehen* wird man dagegen in wesentlich höherem Maße fündig. Mit Ausnahme des Tempodrucks, der Strukturiertheit sowie der Schülerorientierung und der Unterstützung sind bei allen Unterrichtsmerkmalen signifikante Ergebnisse zu verzeichnen.

32.5 Bildungsgangprofile der Unterrichtsqualität

Welche Rolle spielen Bildungsgangunterschiede und Merkmale der Schülerschaft für die Qualität des Unterrichts? Dieser Frage soll in zwei Schritten nachgegangen werden: Zunächst werden die vorgefundenen Bildungsgangunterschiede hinsichtlich der sechs zentralen Merkmale der Unterrichtsqualität beschrieben.

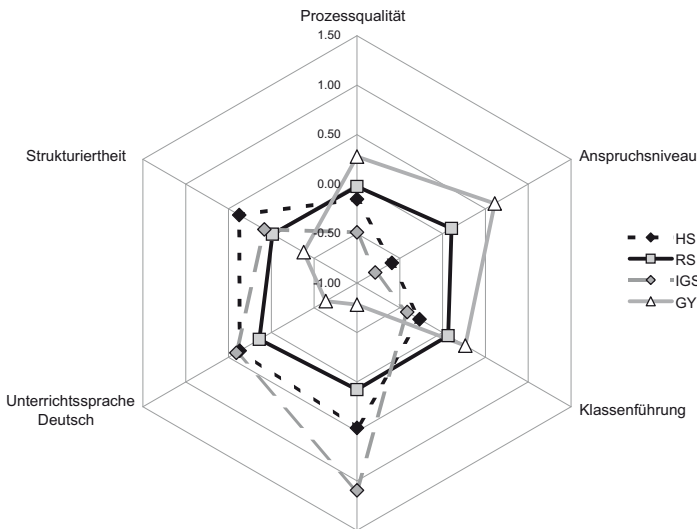


Abbildung 32.4: Bildungsgangunterschiede in den sechs zentralen Merkmalen der Qualität des Englischunterrichts aus Schülersicht.

Hier zeigt Abbildung 32.4, dass Schülerinnen und Schüler, die den Hauptschulabschluss erwerben wollen, und mehr noch Schüler/innen aus Integrierten Gesamtschulen übereinstimmend einen hohen Tempodruck sowie einen stark strukturierten Unterricht berichten. Erwartungsgemäß wird das Anspruchsniveau in Gymnasien am höchsten eingeschätzt, ebenso die Prozessqualität des Unterrichts sowie die Klassenführung.

In einem zweiten Schritt wird dem Sachverhalt Rechnung getragen, dass sich der Unterricht je nach Bildungsgang mit unterschiedlichen Eingangsvoraussetzungen konfrontiert sieht, und zwar in Bezug auf das kognitive Fähigkeits- und Vorkenntnisniveau, den Schüleranteil mit Deutsch als Erstsprache und den Mädchenanteil innerhalb der Klasse. Kontrolliert man die Bildungsgangunterschiede um diese Merkmale, dann erhält man den „Netto-Effekt“ der Bildungsgangunterschiede, d.h. Aussagen darüber, welchen Erklärungswert der Bildungsgang über die jeweils unterschiedlichen Hintergrundmerkmale hinaus für Unterschiede in der Unterrichtsqualität besitzt.

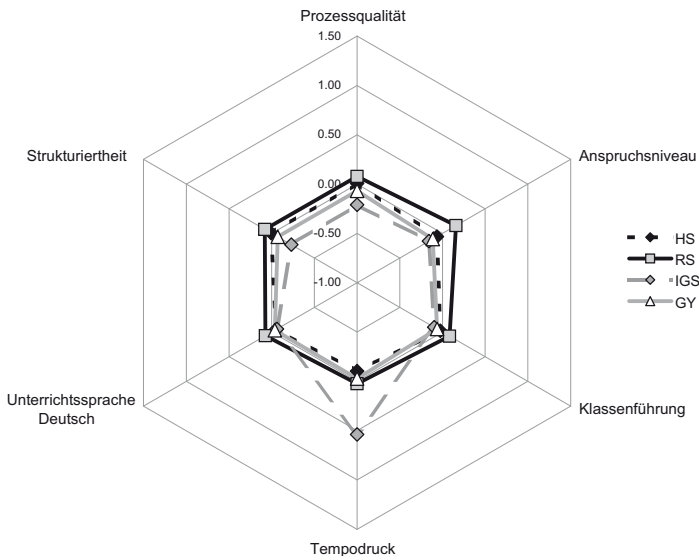


Abbildung 32.5: Um Unterschiede im Kontext (Klassenzusammensetzung) bereinigte Bildungsgangunterschiede in den sechs zentralen Merkmalen der Qualität des Englischunterrichts aus Schülersicht.

Wie Abbildung 32.5 dokumentiert, fallen mit einer Ausnahme (dem Tempodruck in der IGS) die Muster der vorher stark abweichenden Bildungsgangprofile nahezu zusammen. Dies lässt sich so interpretieren, dass der auf den ersten Blick starke Bildungsgangunterschied zwischen den Unterrichtsprofilen weitgehend durch Merkmale der je nach Bildungsgang unterschiedlich zusammengesetzten Schülerschaft vermittelt wird.

32.6 Unterrichtsqualität und Englischleistung: Komplexe Modellierung

Wir verlassen jetzt die Ebene einfacher Zusammenhänge und analysieren das Zusammenwirken von Kontext, Unterrichtsqualität und den Englischleistungen im Nachtest. Da es sehr unterschiedliche Strategien sowohl der kausalen Modellierung als auch der Berücksichtigung des Mehrebenencharakters (Individual- und Klassenebene) gibt und die folgenden Analysen den Kern der Aussagen zur Unterrichtswirksamkeit darstellen (vgl. Kapitel 26), erscheint es uns nötig, die Logik und Technik des hier verwendeten Verfahrens kurz zu skizzieren. Sofern nicht ausdrücklich etwas anderes gesagt wird, verwenden wir das Statistikprogramm *Mplus* (vgl. Muthén/Muthén 2003).

Abbildung 32.6 stellt die Hörverstehensleistung am Ende der neunten Jahrgangsstufe in Abhängigkeit von der Unterrichtssprache und den Eingangsleistungen zu Beginn der neunten Jahrgangsstufe auf Schulklassenebene dar. Aus diesen Zusammenhängen wurden die Effekte der Unterschiede von Bildungsgang, Erstsprache, kognitiver Grundfähigkeit und Mädchenanteil herausgerechnet.

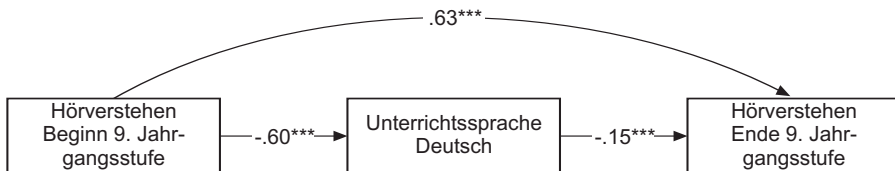


Abbildung 32.6: Direkte und indirekte – über das Merkmal „Unterrichtssprache Deutsch“ vermittelte – Wirksamkeit der anfänglichen Hörverstehenskompetenz (unter Kontrolle der Kontextmerkmale Bildungsgang, kognitive Grundfähigkeit, Mädchenanteil und Anteil von Schülern mit Deutsch als Erstsprache).

Inhaltlich bedeutet dieses Ergebnis, dass die Hörverstehensleistung von Schulklassen *erstens* äußerst stabil ist („Positionsstabilität“): Klassen, die zu Beginn der neunten Jahrgangsstufe zu den Besten gehören, sind mit hoher Wahrscheinlichkeit auch am Ende der neunten Jahrgangsstufe unter den Besten zu finden ($\beta = .63$). *Zweitens* ist die Hörverstehensleistung am Ende der neunten Jahrgangsstufe um so geringer, je häufiger im Unterricht Deutsch gesprochen wird ($\beta = -.15$). *Drittens*: Ob im Englischunterricht Deutsch oder Englisch gesprochen wird, hängt aber in erheblichem Maße vom Eingangsniveau ab: Je höher die Eingangsleistung, umso weniger wird Deutsch gesprochen ($\beta = -.60$), wobei Bildungsgangunterschiede wohlgermerkt herausgerechnet sind!

32.7 Unterrichtsqualität und Lerninteresse

Der Erfolg eines Unterrichts bemisst sich nicht nur nach dem Leistungs- oder Kompetenzzuwachs. Wir haben exemplarisch für die nicht-kognitiven Zielkriterien das Lerninteresse für das Fach Englisch ausgewählt, das zu Beginn und am Ende

der neunten Jahrgangsstufe erhoben wurde. Tabelle 32.5 berichtet zunächst die Zusammenhänge mit den Unterrichtsvariablen.

Tabelle 32.5: Korrelationen der Unterrichtsvariablen mit dem Lerninteresse Englisch am Ende der neunten Jahrgangsstufe. Nicht signifikante Korrelationen ($p > .05$) sind nicht dargestellt.

	einfache Korrelation	Partialkorrela- tion mit dem Lerninteresse Ende 9. Jg.St.	Partialkorrela- tion mit dem Zuwachs
1 Prozessqualität			
Positive Fehlerkultur	.68***	.22**	.43**
Unterrichtsklima	.56***	.25***	.47***
Thematische Motivierung	.62***	.30***	.58***
Schülerorientierung	.62***	.24***	.45***
Unterstützung	.56***	.23**	.45**
Verständlichkeit	.70***	.21**	.41**
2 Anspruchsniveau			
Wichtigkeit: Kommunikation	.73***		
Wichtigkeit: Korrektheit	.53***		
Instrumentelle Motivierung	.47***	.17*	.32*
Integration der Hausaufgaben	.40***		
3 Klassenführung			
Management	.44***	.13*	.25*
Störungsfreiheit	.44***	.14*	.27*
4 Tempodruck	-.59***		
5 Verwendung der deutschen Sprache	-.49***		
6 Strukturiertheit	.18*	.15*	.29*

*Anmerkung: Mehrebenenkorrelationen mit Mplus; Kontrollvariablen: Einstellung zum Fach Englisch zum Beginn der neunten Jahrgangsstufe, Geschlecht, Erstsprache, kognitive Grundfähigkeit, Bildungsgang; Signifikanzniveaus (jeweils einseitiger Test): * $p < .05$, ** $p < .01$, *** $p < .001$.*

Trotz der in Tabelle 32.3 geschilderten relativ geringen Klassenunterschiede (Intraklassen-Korrelationen) im Lerninteresse und dessen Zuwachs sind die Zusammenhänge zwischen Unterrichtsqualität und Entwicklung des Lerninteresses insgesamt noch höher als bei den Testleistungen. Einer der Gründe dafür könnte sein, dass hier Prädiktor und Zielkriterium der gleichen Quelle – dem Schülerfragebogen – entstammen. Vor allem die Komponenten des ersten Faktors, der klassische Merkmale der Unterrichtsqualität umfasst, leisten einen beachtlichen Beitrag zur Vorhersage von Unterschieden in der Entwicklung des Lerninteresses.

Vergleicht man die Ergebnismuster in den Tabellen 32.4 und 32.5 miteinander, dann wird deutlich: Die Frage, was „guten Unterricht“ ausmacht, ist je nach betrachtetem Zielkriterium unterschiedlich zu beantworten. Es gibt einen Satz von Merkmalen der Prozessqualität, die gleichermaßen für die Entwicklung des

Hörverstehens und des Lerninteresses relevant sind, es zeigen sich jedoch auch kriterienspezifische Unterschiede, beispielsweise bei der Strukturiertheit.

32.8 Klassengröße und Heterogenität der Vorkenntnisse

Wenn bisher von relevanten Kontextmerkmalen die Rede war, dann dienten diese im Wesentlichen der „Bereinigung“ der inhaltlich interessierenden Pfade (insbesondere vom Unterricht auf die Leistung) durch zentrale Kontrollvariablen. Im Folgenden möchten wir zwei weitere Merkmale untersuchen, deren Bedeutung für den Unterricht und die Leistung in der Bildungspolitik diskutiert wird: die Klassengröße und die Heterogenität der (klasseninternen) Leistungen.

Die meisten bisherigen Studien haben keinen Zusammenhang zwischen *Klassengröße* und Testleistung nachweisen können. Allerdings handelte es sich a) zu meist um Querschnittsstudien, wurde b) der Unterricht nicht oder nur am Rande thematisiert, und ging es c) überwiegend um mathematisch-naturwissenschaftlichen Unterricht. Im Fremdsprachenunterricht, bei dem die Kommunikation eine überragende Rolle spielt, könnten sich die Dinge anders darstellen. Unsere Hypothese ist deshalb, dass zahlenmäßig kleinere Klassen – alle anderen Kontextfaktoren konstant gehalten – eine günstige Bedingung für guten Unterricht und damit mittelbar auch für den Leistungszuwachs sind.

Zweitens überprüfen wir, ob es zutrifft, dass die *Heterogenität der Vorkenntnisse* zu Beginn der neunten Jahrgangsstufe eine Rolle für den Leistungszuwachs spielt, ob also z.B. eine hohe Leistungsstreuung – alle anderen Kontextfaktoren wiederum konstant gehalten – einen Risikofaktor für die Leistungsentwicklung darstellt.

Entgegen vielfach gefundenen Ergebnissen zum Mathematikunterricht – beispielsweise im Projekt MARKUS (Helmke/Jäger 2002) – zeigt sich hier durchaus ein klares Muster: Die Klassengröße geht eindeutig mit einem ungünstigeren Unterrichtsprofil einher (vgl. Abbildung 32.7). Das um die Hörverstehensleistung (exemplarisch für zwei Unterrichtsmerkmale) erweiterte Pfadmodell (ohne Abbildung) unterstützt die Annahme einer indirekten Wirkungsweise der Klassengröße: Je kleiner die Klasse, desto verständlicher wird der Unterricht wahrgenommen und desto häufiger ist Englisch die Unterrichtssprache. Diese beiden Unterrichtsvariablen wiederum wirken sich auf die Hörverstehensleistung aus.

Wirkt sich die Heterogenität der Vorkenntnisse auf den Unterricht und den Leistungszuwachs aus? Um dieser Frage nachzugehen, haben wir in einem Zweiebenenmodell den Zusammenhang zwischen der Heterogenität der Englischleistungen zu Beginn der neunten Jahrgangsstufe und dem Leistungszuwachs berechnet, unter Einbezug der folgenden Kontext- und Hintergrundmerkmale: Bildungsgang (GY vs. andere, RS vs. andere), Sozialschichtniveau, Mädchenanteil sowie Sprachherkunft. Die Ergebnisse belegen, dass sich die Heterogenität *nicht* negativ auf den Leistungszuwachs auswirkt: die Pfadkoeffizienten betragen für den Zuwachs beim Hörverstehen $-.01$, beim Zuwachs der C-Testleistung $.01$ und sind beide nicht signifikant (ohne Abbildung).

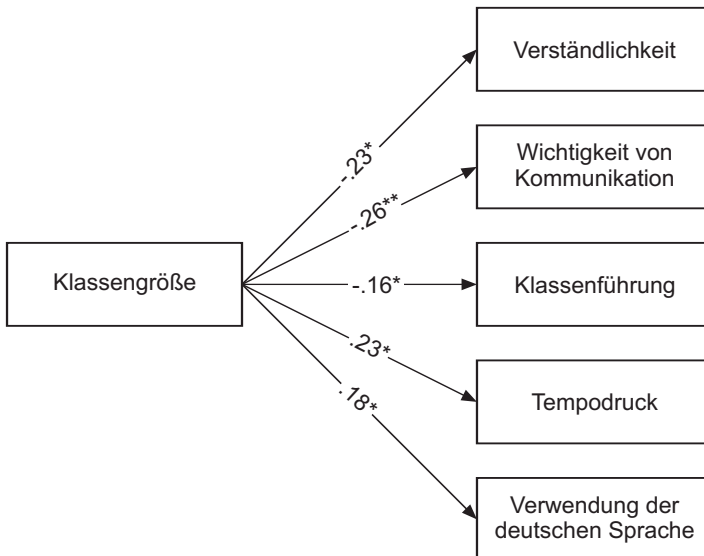


Abbildung 32.7: Klassengröße und Prozessqualität des Unterrichts (unter Kontrolle von Sozialstatus, kognitiver Grundfähigkeit, Mädchenanteil, Erstsprache, Bildungsgänge).

32.9 Ausblick

Wir haben in diesem Kapitel versucht, ein Bild der Wirksamkeit des Englischunterrichts zu zeichnen, und haben dabei drei Schülermerkmale in den Blick genommen, von denen Messungen zu Beginn und am Ende der neunten Jahrgangsstufe vorliegen, so dass Entwicklungen festgestellt werden können: Kompetenz im Bereich des *Hörverstehens* und der *Textrekonstruktion* (C-Test) sowie – als nicht-kognitives Merkmal – *Lerninteresse* im Fach Englisch. Die Ergebnisse haben gezeigt, dass die verschiedenen Zielkriterien mit unterschiedlichen Unterrichtsmustern verknüpft sind, dass es sich also gelohnt hat, anstelle eines Globalwertes „der“ Englischleistung nach unterschiedlichen Kompetenzen zu differenzieren. Zwar ist der Erklärungsbeitrag von Person- und Unterrichtsmerkmalen zahlenmäßig gering – angesichts der extrem hohen Stabilität der betrachteten Schülermerkmale jedoch beachtenswert. Insgesamt gesehen kommt Unterschieden zwischen Lehrpersonen und Unterschieden in der Qualität des Englischunterrichts eine erheblich höhere Bedeutung zu, als dies in den großen deutschen Leistungsstudien der 80er Jahre von Fend (1984) belegt werden konnte.

Wir betrachten die hier berichteten Ergebnisse als Basis, die durch weitere Analysen zunehmend ergänzt und verfeinert werden sollen.

a) Dies betrifft zunächst die Präzisierung durch die Ergebnisse der Englisch-Videostudie (vgl. Kapitel 29) und die Konkretisierung für den Vergleich von Schulen

mit bilinguaem Sachfachunterricht mit Schulen mit regulärem Sachfachunterricht (vgl. Kapitel 35).

b) Während der Schwerpunkt dieses Kapitels auf bildungsgangübergreifenden Ergebnissen lag, werden künftige Analysen bestimmten Fragen verstärkt bildungsgangspezifisch, z.B. nur für die Hauptschule, nachgehen.

d) Komplexe unterrichtliche Bereiche wie der Umgang mit *Schülerfehlern* aus Schüler- und Lehrersicht (sowie auch videobasiert), deren Berücksichtigung den Rahmen dieses Kapitels gesprengt hätten, sind späteren Publikationen vorbehalten.

c) Die Auswertungsstrategie sieht weiterhin komplexe Analysen vor, die hier aus Platzgründen nicht entfaltet werden konnten, beispielsweise um die Frage nach *Mediationsvariablen*, die die Wirksamkeit einzelner Unterrichtsmerkmale erklären, um die Vereinbarkeit unterschiedlicher *Zielkriterien* des Unterrichts, beispielsweise von Kompetenzzuwachs (von Klassen) und Verringerung der Streuung innerhalb der Klassen, um vertiefende Mehrebenenanalysen der *Schülerwahrnehmungen* des Unterrichts sowie um die Verfeinerung des Erklärungsmodells durch Einbezug weiterer Kontext- und Unterrichtsmerkmale, beispielsweise des hier nicht berücksichtigten *Klassenklimas*.

Literatur

- Bausch, K. R./Christ, H./Krumm, H. J. (Hrsg.) (2002): Handbuch Fremdsprachenunterricht. Tübingen: Francke.
- Butzkamm, W. (1998): Zehn Prinzipien des Fremdsprachenlernens und -lehrens. In: Timm, J.-P. (Hrsg.): Englisch lernen und lehren. Berlin: Cornelsen, S. 45-52.
- Butzkamm, W. (2004): Lust zum Lehren, Lust zum Lernen. Eine neue Methodik für den Fremdsprachenunterricht. Tübingen: Francke Verlag.
- Ehrenreich, S. (2004): Auslandsaufenthalt und Fremdsprachenlehrerbildung (Münchener Arbeiten zur Fremdsprachen-Forschung). München: Langenscheidt.
- Fend, H. (1984): Determinanten von Schulleistungen. Wie wichtig sind die Lehrer? In: Unterrichtswissenschaft 12, H.1, S. 62-87.
- Fend, H. (1998). Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim: Juventa.
- Gruehn, S. (2000): Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Münster: Waxmann.
- Haenisch, H. (1979): Schulleistungsvergleiche zwischen Gesamtschulen und Schularten des gegliederten Schulsystems in Nordrhein-Westfalen am Ende des 6. Schuljahres. In: Haenisch, H./Lukesch, H./Klaghofer, R./Krüger-Haenisch, E.-M. (Hrsg.): Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik. Paderborn: Schöningh, S. 6-226.
- Helmke, A./Jäger, R. S. (Hrsg.) (2002): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik.
- Judd, E. L./Tan, L./Walberg, H. J. (2001): Teaching additional languages (Educational Practices Series, Vol. 6). Brussels: International Academy of Education & International Bureau of Education. URL: <http://www.ibe.unesco.org>.
- Klippel, F. (2004): Fachdidaktik Englisch. In: Blömeke, S./Reinhold, P. / Tulodziecki, G. /Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Braunschweig: Klinkhardt, S. 439-442.

- Köller, O./Trautwein, U. (2004): Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In: Köller, O./ Watermann, R./ Trautwein, U./Lüdtke, O. (Hrsg.): Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. Opladen: Leske + Budrich, S. 285-326.
- Lewis, E. G./Massad, C. E. (1975): The teaching of English as a foreign language in ten countries. New York, London, Sydney, Tokyo: John Wiley & Sons.
- Lukesch, H. (1979): Leistungsvergleich zwischen Gesamtschulen und herkömmlichen Schulen am Ende der Pflichtschulzeit in Nordrhein-Westfalen. In: Haenisch H./ Lukesch, H./Klaghofer, R./ Krüger-Haenisch, E.-M. (Hrsg.): Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik. Paderborn: Schöningh, S. 227-364.
- Muthén, L. K./Muthén, B. O. (2003): Mplus User's Guide. Los Angeles: Muthen & Muthen.
- Schulinspektorat des Kantons Graubünden (2004): Erhebung Pflichtfremdsprache Englisch 1. Klasse Oberstufe Schuljahr 2003/04. Schul- und Kindergarteninspektorat des Kantons Graubünden, Schweiz.
- Spada, N./Fröhlich, M. (1995): The communicative orientation of language teaching (COLT) observation scheme: coding conventions and applications. Sydney: National Centre for English Language Teaching and Research.
- Timm, J.-P. (Hrsg.) (1998): Englisch lernen und lehren. Berlin: Cornelsen.
- Waxman, H. C./Tharp, R. G./Hilberg, R. S. (Hrsg.) (2004): Observational Research in U.S. Classrooms. New Approaches for Understanding Cultural and Linguistic Diversity. Cambridge: Cambridge University Press.
- Wienold, G./Achtenhagen, F./Van Buer, J./Oldenbürger, H.-A./Rösner, H./Schluroff, M. (1985): Lehrerverhalten und Lernmaterial in institutionalisierten Lehr-Lern-Prozessen. Am Beispiel des Englischanfangsunterrichts (Band 9 – 2). Göttingen: Seminar für Wirtschaftspädagogik der Georg-August-Universität.