Deutsches Institut für Internationale Pädagogische Forschung

Z 43

Akademie der Pädagogischen Wissenschaften
– Präsenzbibliothek –

MITTEILUNGEN UND NACHRICHTEN

SONDERHEFT

Die Leistungen im Englischunterricht in der Bundesrepublik im internationalen Vergleich

Von Prof. Dr. Walter Schultze

FRANKFURT AM MAIN

1975

DEUTSCHES INSTITUT FÜR INTERNATIONALE PÄDAGOGISCHE FORSCHUNG
UND GESELLSCHAFT ZUR FÖRDERUNG PÄDAGOGISCHER FORSCHUNG E.V.

MITTEILUNGEN UND NACHRICHTEN

SONDERHEFT

Die Leistungen im Englischunterricht in der Bundesrepublik im internationalen Vergleich

Von Prof. Dr. Walter Schultze

Frankfurt am Main
1975

92, 2390

Bibliothek für
Bildungsgeschichtliche
Forschung
Berlin

Herausgeber: Forschungskollegium des Deutschen Instituts
für Internationale Pädagogische Forschung

Redaktion : Dr. Rosemarie Klaus-Roeder, DIPF,
6000 Frankfurt a.M. 90, Schloßstr. 29

I N H A L T

		Seite
1.	Anlage und organisatorische Durchführung der Untersuchung	2
1.1	Allgemeines	2
1.2	Die Stichprobe	2
1.3	Die Untersuchungsinstrumente	4
2.	Die Bundesrepublik im internationalen Vergleich	10
2.1	Die Meßergebnisse	10
2.2	Das Verhältnis von Prädiktoren und Kriteriumsvariablen	13
2.2.1	Die Stellung des Englischunterrichts in den Schulen	14
2.2.2	Die Lehrervariablen	17
2.2.3	Die Schülervariablen	19
2.2.4	Zusammenfassung der Ergebnisse der Regressionsanalyse	20
2.3	Zusammenfassung der Ergebnisse durch Lewis	21
3.	Die Ergebnisse in der Bundesrepublik	23
3.1	Zur Stichprobe	23
3.2	Population II	25
3.3	Population IV	29
4.	Schluß	34
4.1	Zusammenfassung	34
4.2	Kritische Anmerkungen	35
5.	Anhang: Die wichtigsten Ergebnisse der Studie über den Französischunterricht in acht Ländern	37

1. Anlage und organisatorische Durchführung der Untersuchung

1.1 Allgemeines

In einem Sonderheft der "Mitteilungen und Nachrichten" des Deutschen Instituts für Internationale Pädagogische Forschung berichtete der Verfasser 1974 über die Leistungen im naturwissenschaftlichen Unterricht in der Bundesrepublik im internationalen Vergleich. Dort wurde zugleich (S. 2 ff.) eine Vorgeschichte der Untersuchungen der International Association for the Evaluation of Educational Achievement, I.E.A., gegeben und bereits darauf hingewiesen, daß Englisch als Fremdsprache ein Teil der zweiten Phase der I.E.A.-Untersuchungen, der Sechs-Fächer-Studie, war. Die Bundesrepublik hat sich mit den Populationen II (den 14jährigen) und IV (den Schülern der Abgangsklassen der Gymnasien) an dieser Untersuchung beteiligt.

Da diese Untersuchung ein Jahr später als die Erhebung im naturwissenschaftlichen Unterricht erfolgte, erschien auch die internationale Publikation über die Ergebnisse dieser Untersuchung ein Jahr später.

Im folgenden soll über die Ergebnisse der Untersuchung berichtet werden, wobei sich die Darstellung hinsichtlich des internationalen Vergleichs auf die Veröffentlichung von Lewis-Massad stützt (E.Glyn Lewis and Carolyn E. Massad: The Teaching of English as a Foreign Language in Ten Countries. Stockholm 1975). Zusätzlich werden die Veröffentlichungen von John D.Carroll: The Teaching of French as a Foreign Language in Eight Countries.Stockholm 1975, sowie Robert L.Thorndike: Reading Comprehension Education in Fifteen Countries. Stockholm 1973, herangezogen.

1.2 Die Stichprobe

Hinsichtlich der Anlage und organisatorischen Durchführung der Untersuchung gelten die im genannten Sonderheft gemachten Ausführungen auch für die Erhebungen zum Englischunterricht. Diese erstrecken sich allerdings nur auf die Populationen II und IV, da die Population I (die 10jährigen) verständlicherweise hierfür nicht in Betracht kam.

Die Zahl der Schulen, die aus verschiedenen Gründen ausfielen und nicht immer ersetzt werden konnten, war größer als bei der Studie über den naturwissenschaftlichen Unterricht, so daß die Zahl der beteiligten Schulen, Lehrer und Schüler insgesamt geringer war.

Tabelle 1: Zahl der in der BRD beteiligten Schulen, Lehrer und Schüler

Pop.II		Schüler	Schulen	Lehrer
	Hptsch.	402	25	
	Realsch.	415	14	
	Gymn.	255	11	
	Gesamt	1072	50	248
Pop.IV		1379	61	360
	Insgesamt	2451	111	608

Da die Population II wiederum nach Hauptschulen, Realschulen und Gymnasien untergegliedert werden muß, wird ohne weiteres ersichtlich, daß die Stichprobe von 1072 Schülern und 50 Schulen zu gering ist, um noch nach Ländern aufgeschlüsselt zu werden, so daß sie nicht als repräsentativ für die einzelnen Länder der Bundesrepublik angesehen werden kann, wenn auch der Versuch gemacht wurde, die Zahl der Schulen und Schüler in Relation zu den entsprechenden Zahlen der Länder zusammenzustellen. Hinsichtlich der Population IV gelten Einschränkungen lediglich für die kleineren Länder; wir kommen später darauf zurück.

Einen Hinweis auf die Repräsentanz der Stichprobe für die Bundesrepublik bietet die Tatsache, daß der in beiden Untersuchungen zusätzlich durchgeführte Wortschatztest praktisch die gleichen Ergebnisse zeigte.

Tabelle 2: Durchschnittswerte des Wortschatztests bei den Untersuchungen des naturwissenschaftlichen Unterrichts und des Englischunterrichts in der Bundesrepublik (Standardabweichungen in Klammer)

	Englisch	Naturw.
Pop.II	15.8 (9.8) [1)]	15.9 (9.8)
Pop.IV	24.0 (7.0)	24.5 (6.6)

1) gewichtet nach dem tatsächlichen Verhältnis der Schüler in den drei Schulformen.

1.3 Die Untersuchungsinstrumente

Die vollständige Erhebung sah vor, daß folgende Bereiche durch Tests gemessen wurden:
a) Verständiges Lesen,
b) Hörverstehen,
c) Sprechen,
d) Schreiben.
Während die Lesefähigkeit mit den üblichen Tests und Auswertungsverfahren ohne weiteres erfaßbar war, konnten die drei übrigen Bereiche nur unter Zuhilfenahme des Tonbandes und nicht-objektiver Auswertungsmethoden einbezogen werden. Da die Vervielfältigung der Tonbänder bei der relativ großen Zahl der beteiligten Untersuchungsgruppen nicht nur erhebliche Kosten verursacht hätte, sondern auch voraussetzte, daß in allen beteiligten Schulen hochwertige Tonbandgeräte zur Verfügung standen, da uns andererseits aus Untersuchungen in England über den Frühbeginn mit Französischunterricht bekannt war, daß viele Schüler rein akustisch Schwierigkeiten hatten, die Wiedergabe der Tests durch Tonbandgeräte aufzunehmen, kurz gesagt,-wegen der Umständlichkeit des Erhebungsverfahrens und der damit verbundenen weiteren Kosten - wurde in der Bundesrepublik wie auch in vielen anderen Ländern davon abgesehen, diese Bereiche in die Untersuchung einzubeziehen. Wir beschränkten uns somit auf das verständige Lesen. Wie die Analyse ergab, steht diese Leistung in hoher Korrelation mit dem Gesamttest, so daß diese Entscheidung nachträglich durch die Ergebnisse gerechtfertigt wurde. Außerdem hätte hinsichtlich der Sprech- und Schreibfähigkeit aus rein zeitlichen Gründen eine nur sehr viel kleinere Gruppe getestet werden können, so daß auch die Repräsentanz nicht mehr gewährleistet war. Wie aus der Tabelle 3 ersichtlich, hat kein Land bei beiden Populationen alle Teiltests durchgeführt.

Tabelle 3: Seite 5

Da die Bundesrepublik lediglich am Lesetest beteiligt war, werden im folgenden nur Angaben über diesen Test gemacht. Die Instrumente wurden grundsätzlich in der gleichen Weise entwickelt wie die Instrumente für den naturwissenschaftlichen Unterricht, d.h. die Tests, Fragebogen und Verhaltens- bzw. Einstellungsskalen wurden in Zusammenarbeit von dem internationalen Ausschuß mit den nationalen Zentren aufgestellt. Sie sollten gleichzeitig phonologische Lesefertigkeiten wie grundlegende grammatische Einsichten und ein bestimmtes Vokabular testen.

Darüber hinaus sollte die Einstellung der Schüler zu der englischen Sprache und zum englischen Volk durch

Tabelle 3: Größe der Stichprobe und beteiligte Länder an den Teiltests

	Land	Population	Schülerzahl
Hörverstehen	Chile	IV	231
Schreiben 1,2,3	Chile	IV	182
Schreiben 4	Chile	IV	157
	Finnland	II	605
	Finnland	IV	997
	Ungarn	IV	565
	Italien	II	582
Sprechen	Belgien (fr.)	II	162
	Belgien (fr.)	IV	192
	Finnland	II	200
	Finnland	IV	279
	Ungarn	IV	118
	Italien	II	70
	Italien	IV	28
	Schweden	II	218
	Schweden	IV	204

Nach Tab. 2.4 bei Lewis/Massad, S.59

zwei Einstellungsskalen erfaßt werden. Weiterhin wurde von den Lehrern das relative Gewicht von Hörverstehen, Lesen, Sprechen und Schreiben innerhalb der Ziele des Englischunterrichts erfragt.

Die nationalen Zentren lieferten über 1.500 Einzelaufgaben und 30 Texte verschiedenen Inhalts und verschiedenen Schwierigkeitsgrades als Unterlage für die Entwicklung der Vortests, die dann von den nationalen Zentren auf ihre Eignung überprüft und in einzelnen Ländern vorgetestet wurden. Um alle kulturellen Unterschiede zwischen den Ländern auszuschalten, wurden Inhalte gewählt, die in allen teilnehmenden Ländern vorausgesetzt werden konnten.

Der Lesetest bestand aus 60 Aufgaben, die in einer Stunde (60 Minuten) zu lösen waren. Die Aufgaben waren vom Vielfachauswahltyp und wurden in sechs Gruppen unterteilt. In den Tests für die Pop.II bestand die erste Gruppe aus sechs Aufgaben, in denen die Schüler das Antonym zu einem gegebenen Wort in einem kurzen Text unter fünf Möglichkeiten zu finden hatten.

Is your friend a fat boy? No, he is .....

A. round
B. short
C. thin
D. large
E. tall

Die zweite Gruppe bestand aus acht Aufgaben, in denen die Fähigkeit, die Aussprache geschriebener Worte zu erkennen, geprüft wurde. Die Schüler hatten das Wort zu unterstreichen, in dem der gleiche Laut vorkommt, der in dem Reizwort unterstrichen ist.

bread

A. head
B. teach
C. please
D. east
E. sea

Die dritte Gruppe bestand aus zehn Aufgaben, bei denen es darauf ankam, die richtige grammatische Struktur zu erkennen, indem von drei Auswahlantworten diejenige auszusuchen war, die einen Reizsatz richtig vollendete.

Mr. Smith is looking ..... his lost key.

A. for
B. of
C. to

In den 6 Aufgaben der vierten Gruppe wurde ebenfalls ein Reizsatz gegeben, in dem ein Wort unterstrichen war. Aus drei Kurzdefinitionen war die richtige Definition dieses Wortes zu wählen.

Please turn off the radio.

A. I want to listen in
B. I don't want to listen in
C. I want to see

Bei den Gruppen fünf und sechs handelte es sich um die üblichen Tests für verständiges Lesen. In Gruppe fünf wurden vierzehn einfache Sätze gegeben. Anschließend waren drei Behauptungen auf die richtige hin zu prüfen. Der Abschnitt sechs bestand aus längeren Texten. Die anschließenden 16 Aufgaben mußten durch die richtige Antwortwahl (vier Möglichkeiten) gelöst werden. (Von einem Beispiel wird abgesehen).

Die Tests für die Population IV wiesen grundsätzlich die gleiche Struktur auf. Die erste Gruppe bestand aus acht Aufgaben, bei denen es darauf ankam, die richtige Betonung mehrsilbiger Wörter zu beherrschen.

(i) how-ev-er
(ii) to-geth-er
(iii) gen-tle-man

A. Bei (i) und (ii) gleich
B. Bei (i) und (iii) gleich
C. Bei (ii) und (iii) gleich
D. Bei allen gleich
E. Bei allen unterschiedlich

In der zweiten Gruppe wurde die Fähigkeit getestet, ein gegebenes Eigenschaftswort in Beziehung zu Dingwörtern zu setzen und das Dingwort zu unterstreichen, dem es nicht zugeordnet werden kann (fünf Möglichkeiten, acht Aufgaben).

Raw
A. skin
B. wind
C. meat
D. sigh
E. food

In der dritten Gruppe -- bei ebenfalls acht Aufgaben - mußten Sätze vervollständigt werden, wobei die richtige Form eines Verbs, die richtige Wortordnung oder der Gebrauch des Artikels bekannt sein mußten.

We should like ..... a story.

A. you to tell us
B. you tell us
C. it that you told us

Die Gruppe vier mit elf Aufgaben prüfte das Leseverständnis relativ einfacher Sätze, während die fünfte Gruppe aus fünf zusammenhängenden Prosatexten bestand, die in zunehmendem Maße in ihrem Schwierigkeitsgrad ansteigen. Auf jeden Text folgten fünf oder sechs Fragen, denen vier Alternativantworten zugeordnet waren (25 Aufgaben). Es ist einleuchtend, daß mit diesen Aufgaben nicht nur die Beherrschung der englischen Sprache, sondern zugleich eine allgemeine intellektuelle Fähigkeit geprüft wird, die völlig unabhängig von der Beherrschung einer bestimmten Sprache ist.

Im Lesetest waren 16 Aufgaben in beiden Populationen gleich, um evtl. einen Fortschritt in der Sprachbeherrschung feststellen zu können, der allerdings später bei der Analyse nicht ermittelt wurde, weil sich diese "Ankeraufgaben" aufgrund der Ergebnisse nicht als repräsentative Aufgabenreihe für den gesamten Lesetest erwiesen.

Die Zuverlässigkeit (reliability) des Lesetests (Kuder-Richardson, Formel 20) war in sämtlichen Ländern gut. Sie betrug für die Pop.II im internationalen Durchschnitt 0.93, für die Bundesrepublik desgleichen.
In der Pop.IV betrug sie international 0.84, für die Bundesrepublik ergab sie den geringsten aller für die Länder ermittelten Wert mit nur 0.68.
Die Validität des Tests konnte nur durch eine Beurteilung seitens der nationalen Zentren gesichert werden; eine statistische Evidenz kann mangels entsprechender anderer Messungen nicht gegeben werden.

Außer dem durch die kognitiven Tests erfaßten Leseverständnis wurden fünf Variablen, die im Zusammenhang mit dem Unterricht stehen, ermittelt. Davon wurde die "School-Umgebungsskala" von den weiteren Untersuchungen ausgeschlossen, da sie nur eine sehr niedrige Zuverlässigkeit hatte. Die Einstellung zur Schule, die in allen Untersuchungen mittels einer Einstellungsskala gemessen wurde, hatte nur eine bescheidene Zuverlässigkeit, genauso wie drei weitere Angaben, die den Englischunterricht angeben. Sie betreffen englische Aktivitäten außerhalb der Schule, Interesse am Englischen und eine Einschätzung des Nutzens, den man von der Beherrschung der englischen Sprache hat. Ihre Zuverlässigkeit beträgt international 0.52, 0.66 und 0.65, in der Bundesrepublik 0.40, 0.64 und 0.70.

Der Schwierigkeitsgrad des Tests insgesamt sowie der einzelnen Aufgaben wird bei Lewis/Massad lediglich nach der Zahl der Schüler, die den Test bis zum Ende durchgearbeitet haben, sowie nach der Zahl der ausgelassenen Aufgaben erörtert. Leider findet sich keine Angabe über den Schwierigkeitsgrad der einzelnen Testaufgaben (Prozentzahl der richtigen Lösungen in verschiedenen Ländern). Die Zahl der Schüler in Pop.II, die die letzten vier Aufgaben nicht erreichen, ist in der Bundesrepublik klein (5.8 bis 8.7 Prozent); in der Population IV noch geringer (3.2 bis 7.5 Prozent), so daß von daher der Schwierigkeitsgrad angemessen erscheint. Eine andere Möglichkeit, den Schwierigkeitsgrad des Tests insgesamt zu beurteilen, ist die durchschnittliche Zahl der richtigen Lösungen. Sie beträgt im internationalen Durchschnitt für die Pop.II 24.5 Punkte, für die Bundesrepublik 29.9 Punkte.[1] Bei etwa 30 richtigen von 60 möglichen Lösungen wäre ein angemessener Durchschnittswert erreicht. Insgesamt, d.h. international jedoch erweist sich der Test, insbesondere in Italien, Thailand und Finnland als zu schwer. Für die Population IV beträgt der internationale Durchschnitt 25.8 Punkte (von 60), der Wert für die Bundesrepublik beträgt 41.2 Punkte, womit sich der Test für die Bundesrepublik als recht leicht erweist. Der niedrige internationale Durchschnitt kommt auch dadurch zustande, daß Länder wie Chile und Thailand, aber auch Italien und Ungarn, relativ niedrige Werte erreichen. Wenn also auch der Schwierigkeitsgrad der einzelnen Aufgaben nicht bekannt ist, so darf man doch aufgrund des Gesamtergebnisses feststellen, daß die Tests für die Bundesrepublik in Pop. II angemessen, in Pop. IV leicht sind und eine entsprechende Verteilung aufweisen.[2]

Die Streuung der Tests ist allgemein groß. Die durchschnittliche Streuung in der Pop.II beträgt international 16.3, in der Bundesrepublik 16.2 Punkte; in Pop.IV ist die Streuung entsprechend der stärkeren Auslese der Schüler und der damit größeren Homogenität der Klassen geringer; sie beträgt international durchschnittlich 9.4 Punkte, in der Bundesrepublik ist sie mit 6.1 Punkten die geringste von allen Ländern.
Zusätzlich wurde wie bei allen Untersuchungen der Phase II ein Wortschatztest gegeben, der in allen Fällen gleich war.

---

[1] Gewichtet nach den tatsächlichen Schülerzahlen in den drei Schularten 32.4 Punkte.

[2] Es zeigt sich hier, daß es wenig sinnvoll ist, gleiche Tests in Ländern mit großen Leistungsunterschieden zu verwenden.

2. Die Bundesrepublik im internationalen Vergleich

2.1 Die Meßergebnisse

Da der prozentuelle Anteil der Schüler, der Englisch lernt, in den beteiligten Ländern unterschiedlich ist, können Vergleiche nur mit großer Vorsicht vorgenommen werden. Trotzdem vermögen die in Tabelle 4 gegebenen absoluten Durchschnittswerte ein erstes Bild zu geben.

Tabelle 4: Durchschnittswerte und Standardabweichungen (in Klammer) für den Lesetest sowie für den Wortschatztest

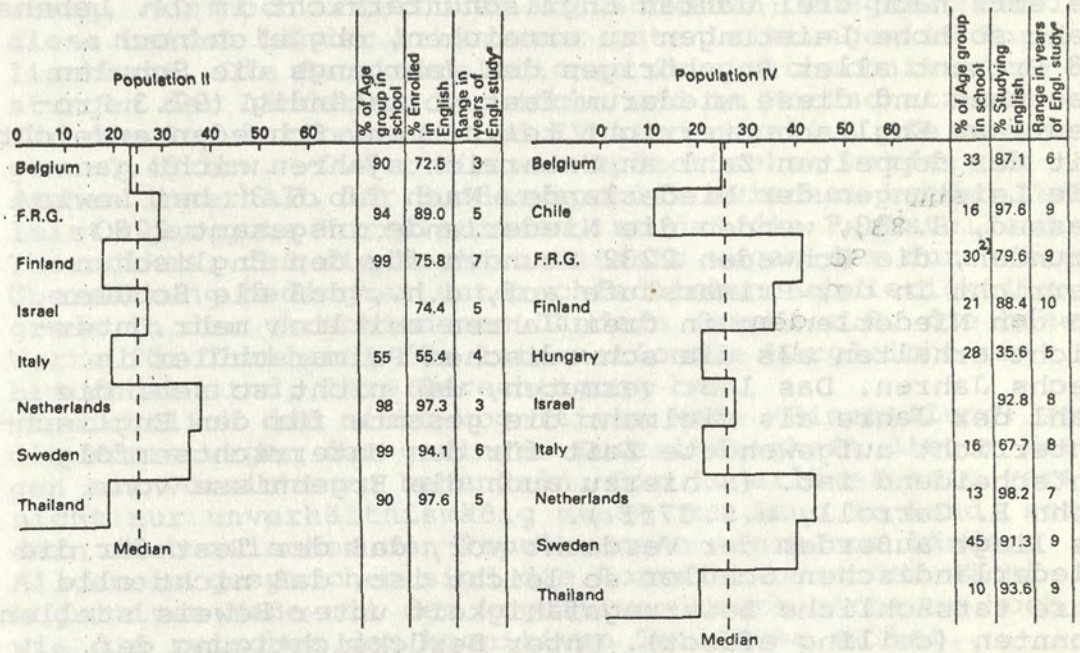
	Pop. II		Pop. IV	
Land	Lesetest	Wortschatzt.	Lesetest	Wortschatzt.
Belgien(fr.)	22.8(15.2)	24.1(9.7	24.2(10.8)	30.7(7.9)
BRD	29.9(16.2)	15.3(9.8)	41.2(6.1)	24.0(7.0)
Chile	- -	- -	10.6(9.8)	20.9(9.5)
Finnland	16.7(17.8)	15.5(10.1)	35.6(8.3)	14.8(6.9)
Ungarn	- -	- -	20.1(11.2)	34.3(6.4)
Israel	26.3(18.7)	20.5(11.6)	27.4(12.9)	18.7(7.8)
Italien	18.6(17.3)	25.9(10.2)	20.2(12.6)	32.9(6.7)
Niederlande	37.4(16.4)	21.1(9.8)	42.9(6.5)	25.6(7.6)
Schweden	35.0(14.9)	13.5(8.2)	39.5(9.0)	25.1(7.4)
Thailand	18.4(13.4)	30.3(8.8)	19.7(9.1)	34.2(6.6)
Insgesamt	24.5(16.3)	20.8(9.8)	25.8(9.4)	25.3(7.2)

[1] Nach Lewis/Massad, Tab. 4.1 und 4.2, S.102/103. Es wurde die im englischen Text gewählte alphabetische Wortfolge der Länder beibehalten.

Der für die Pop.II in der Bundesrepublik ermittelte Durchschnittswert von 29.9 Punkten ist falsch. Nach der vorliegenden Grundauszählung der Rohwerte beträgt er 36.0. Dieser Wert ist jedoch wiederum zu hoch, weil das Verhältnis von Hauptschülern zu Realschülern und Gymnasiasten ein Übergewicht der letzten beiden Gruppen in unserer Stichprobe hat; korrigiert entsprechend dem richtigen Verhältnis beträgt der Durchschnitt 32.4 Punkte. Es befinden sich jedoch bei Lewis keine Angaben darüber, ob es sich bei dem in Tab.4 gegebenen Durchschnittswert bereits um eine irgendwie gewichtete Rechnung handelt.

Zum Zwecke eines sinnvolleren Vergleichs wurden die absoluten Werte aller Tests in Standardwerte umgewandelt, indem sie auf den jeweiligen Medianwert des Landes bezogen wurden. Abbildung 1 gibt eine Darstellung der standardisierten Durchschnittswerte bei gleichzeitiger Angabe des Prozentsatzes der noch in den Schulen befindlichen Schüler, des Prozentsatzes dieser Schüler, die Englischunterricht erhalten, und der Zahl der Jahre, die Englischunterricht für diese Schüler erteilt worden ist, weil diese Variablen entscheidend für die Schulleistung sind.

Abbildung 1: Standardwerte der Länder im Lesetest[1)]



1) Unveränderte Übernahme von Lewis/Massad, Figur 4.1, S.85.

2) Die 30 Prozent beziehen sich auf sämtliche Schulen, nicht nur auf die Gymnasien.

Infolge der von uns durchgeführten Korrektur der Durchschnittswerte für die Pop.II würden sich auch die Standardwerte entsprechend erhöhen und die Bundesrepublik mit den Leistungen in der Pop.II noch näher an die Niederlande und Schweden heranrücken. Dadurch zeichnen sich diese drei Länder durch hohe Leistungen aus. Auffällig ist dabei, daß in den Niederlanden die Schüler der Pop.II nur drei Jahre Englischunterricht, entsprechend in Pop.IV nur sieben Jahre Englischunterricht gehabt haben, während andere Länder mit mehr Jahren an Englischunterricht in ihren Leistungen erheblich darunter liegen. Es würde sich zweifellos lohnen, im einzelnen den Ursachen nachzugehen, die es ermöglichen, in den Niederlanden bereits nach drei Jahren Englischunterricht im 14. Lebensjahr so hohe Leistungen zu erreichen, obgleich noch 98 Prozent aller Angehörigen des Jahrgangs die Schulen besuchen und diese wiederum fast vollständig (97.3 Prozent) am Englischunterricht teilnehmen. Schweden erreicht mit der doppelten Zahl an Unterrichtsjahren nicht ganz die Leistungen der Niederlande. Nach Tab.6.3 bei Lewis/ Massad, S.183, wenden die Niederlande insgesamt 2280 Stunden, die Schweden 2232 Stunden für den Englischunterricht in der Primarstufe auf, d.h., daß die Schüler in den Niederlanden in drei Jahren zeitlich mehr Unterricht erhalten als die schwedischen Primarschüler in sechs Jahren. Das läßt vermuten, daß nicht so sehr die Zahl der Jahre als vielmehr die gesamte für den Englischunterricht aufgewendete Zeit für den Unterrichtserfolg entscheidend ist. (S.hierzu auch die Ergebnisse von John B. Carroll, u.S.37ff.).
Es liegt außerdem der Verdacht vor, daß der Test für die niederländischen Schüler so leicht ist, daß nicht alle ihre tatsächliche Leistungsfähigkeit unter Beweis stellen konnten (ceiling-effect). Unter Berücksichtigung der Streuung (Tab.4) ergibt sich nämlich, daß 16 Prozent der Schüler 53.8 Punkte (bei 60 möglichen Lösungen) und mehr erreicht haben.

Für die Pop.IV sieht das Ergebnis folgendermaßen aus:

Die drei leistungsfähigsten Länder sind die gleichen wie in Pop.II. Die Bundesrepublik folgt mit 41.2 Punkten den Niederlanden (42.9 Punkte) und liegt vor Schweden (39.5 Punkte).

Dieses günstige Bild vom Englischunterricht in der Bundesrepublik (eine Aufschlüsselung nach Schularten erfolgt unten in Kapitel 3) muß im Hinblick auf die nur 9 Prozent der Schüler, die in der Population noch die letzte Klasse unserer Gymnasien besuchen, relativiert werden.
Die Strenge der Auslese spiegelt sich in der Streuung wider. Die Standardabweichung beträgt für die Bundesrepublik 6.1 Punkte, die Niederlande folgen mit 6.5 Punkten dicht auf, während die Standardabweichung in Schweden 9.0 Punkte beträgt, ein Spiegelbild der größeren

Schülerzahl, die sich in den erfaßten Abschlußklassen befindet.
Da im Unterschied zur Untersuchung in den Naturwissenschaften keine Berechnung der durchschnittlichen Leistungen der jeweils besten 1, 5 und 9 Prozent leistungsstärksten Schüler erfolgt ist, ist ein echter Vergleich der Leistungen in der Bundesrepublik mit den Leistungen in anderen Ländern nicht möglich.
Was die weiteren Ergebnisse betrifft, so ist bemerkenswert, daß der Wortschatztest, dessen Ergebnisse ebenfalls in Tabelle 4 wiedergegeben sind, international schlecht mit der fremdsprachlichen Leistung korreliert. Die Bundesrepublik liegt hier am Ende der unteren Leistung, nur noch unterboten von Schweden; ausgerechnet diese beiden Länder haben sehr gute Leistungen im Englischtest. Die Rangkorrelationen der Durchschnittsleistungen der Länder im Englisch- und im Wortschatztest für die Populationen II und IV sind negativ, sie betragen -0.45 bzw. -0.41. Wie wir später sehen werden, bestehen innerhalb der Bundesrepublik zwischen beiden Testleistungen gute Korrelationen, so daß die Feststellung Thorndikes, daß die Wortschatztests wegen der durch die Übersetzung bedingten unterschiedlichen Schwierigkeitsgrade in den verschiedenen Ländern (siehe z.B. die hohen Werte für Thailand) für internationale Vergleiche unbrauchbar sind, ihre Berechtigung hat.
Hinsichtlich der nicht-kognitiven Maße (Einstellung) sind die Werte wegen der geringen Zuverlässigkeit dieser Messungen mit Vorsicht zu behandeln. Die Zuverlässigkeit variiert nicht nur unverhältnismäßig stark von Land zu Land, sondern auch zwischen den Populationen einzelner Länder. Allgemein gesprochen sind die Korrelationen der nicht-kognitiven Maße mit den kognitiven Tests nicht so hoch wie die zwischen den kognitiven Teiltests selbst.
Die Interessenskala und die geschätzte Nützlichkeit einer Beherrschung der englischen Sprache korrelieren in allen Ländern und allen Populationen verhältnismäßig hoch. Der innere Zusammenhang zwischen diesen beiden Einstellungsskalen dürfte ohne weiteres einsichtig sein.

2.2 Das Verhältnis von Prädiktoren und Kriteriumsvariablen

Um die zahlreichen unabhängigen Variablen (Prädiktoren) in den Griff zu bekommen, werden sie zum Zwecke der weiteren statistischen Verarbeitung zu Gruppen (Blöcken) zusammengefaßt. Leider sind diese Zusammenfassungen nicht in allen Erhebungen der zweiten Phase in paralleler Form erfolgt. Lewis hat außerdem unabhängig von der von Carolyn E. Massad durchgeführten Regressionsanalyse in seinem Teil der Darstellung Kategorien benutzt, die sich zum Teil mit denen der Regressionsanalyse überschneiden. Lewis Darstellung beschränkt sich ferner weitgehend auf Beschreibungen, weniger auf statistische Auswertungen. Da er außerdem seine Ergebnisse in einem doppelten Durchgang darstellt, ist es nicht ganz leicht, diese zusammen-

zufassen, und Wiederholungen sind nicht immer vermeidlich. Wir folgen seiner Hauptgliederung und beziehen die Ergebnisse der Regressionsanalyse von C. Massad in diese Gliederung ein.

2.2.1 Die Stellung des Englischunterrichts in den Schulen

In diesem Kapitel faßt Lewis sowohl die Angaben zum sozio-ökonomischen Schulindex als aber auch den allgemeinen kulturellen Hintergrund für die Arbeit in der Schule zusammen. Gleichzeitig behandelt er die für den Unterricht aufgewendete Zeit , für die in der Regressionsanalyse ein besonderer Block (III) gebildet wird. Die eigentliche Unterrichtspraxis wird von ihm unter dem Kapitel "Charakteristika der Lehrer" behandelt, wahrscheinlich weil diese Angaben auf den Lehrerfragebogen zurückgehen.
Die Gründe, aus denen in zehn Ländern Englisch unterrichtet wird, sind verschieden, verschieden ist auch der Grad der industriellen Entwicklung sowie ihre ökonomische Lage. Es besteht einige Evidenz, daß die Leistungen in den zehn Ländern mit der besonderen Form, in der diese sozietalen Variablen wirksam sind, korrelieren. Linguistische Minoritäten in den Ländern scheinen ohne Einfluß auf die Leistung zu sein. Linguistisch homogene Länder wie die Bundesrepublik und Schweden zeigen genauso gute Leistungen wie linguistisch heterogene Länder wie Israel und Belgien. Auch der linguistische Abstand der Muttersprache vom Englischen scheint in keiner eindeutigen Relation zur Leistung im Englischunterricht zu stehen. Allerdings macht sich in den Bevölkerungsgruppen, in denen ein Unterschied zwischen der Sprache in der Schule und der Sprache im Hause der Schüler besteht, ein negativer Einfluß auf die Leistung des Fremdsprachenunterrichts bemerkbar. (Dieser Einfluß beschränkt sich wahrscheinlich nicht nur auf den Fremdsprachenunterricht.)

Eine gewisse Rolle scheint der infolge der ökonomischen Entwicklung auf dem Gebiet des Handels bestehende Kontakt mit englischsprechenden Ländern zu spielen. In Ländern wie Israel und Schweden vollzieht sich deshalb das Erlernen der englischen Sprache unter ganz anderen Voraussetzungen als in Ungarn, Thailand oder Chile. Doch sollte man annehmen, daß dieser Faktor in den Einschätzungen der Nützlichkeit des Spracherwerbs zum Ausdruck kommt, was aber nicht der Fall ist. In allen Ländern beträgt die diesbezügliche Korrelation in den Populationen IV 0.83 bis 0.90. Bezieht man in die Voraussetzungen, die in den Schulen bezüglich des Erlernens einer Sprache vorhanden sind, auch den Wortschatz als Maß der schulischen Lernfähigkeit der Schüler ein, so erweist sich dieser als von großem Gewicht in den einzelnen Ländern. Der Zusammenhang zwischen dem Wortschatz und der Englischleistung ist innerhalb sämtlicher beteiligter Länder

gleichsinnig positiv, so daß sich der Wortschatz als der beste Prädiktor für die Leistung im Englischunterricht erweist. Der Zusammenhang beträgt in der Bundesrepublik in Pop.II r=0.65, für Pop.IV r=0.43, ähnlich hoch ist er in Schweden. In der Regressionsanalyse wurde der Wortschatz als gesonderter Block VI behandelt und erwies sich auch dort als guter Prädiktor,besonders für die Leseleistung und für das Hörverstehen. In der Bundesrepublik erklärt er in der Pop.IV weit mehr als in den meisten anderen Ländern einen großen Teil der Varianz und erreicht mit 10.4 Prozent trotz seiner Stellung am Schluß der Analyse den größten Anteil aller Blöcke. In Pop.II beträgt die zusätzlich erklärte Varianz allerdings nur 2.5 Prozent, wahrscheinlich deshalb, weil durch die Variablen der Blöcke I bis V bereits 76 Prozent aller Varianz aufgeklärt werden.
Durch den sozio-kulturellen Schulindex und das Geschlecht werden in der Pop.II in der Bundesrepublik 15.1 Prozent der Varianz, in der Pop.IV allerdings nur 1 Prozent der gesamten Varianz erklärt.
Nach den Angaben bei Lewis (Tabellen 8.1 und 8.2, S. 247/248/249) ergeben sich hinsichtlich der Eltern in der Bundesrepublik folgende Tatsachen:

Tabelle 5: Ausmaß der Schulbildung der Eltern in den Populationen II und IV (in Jahren)

Jahre	Pop.II	Pop.IV
Vater	9.0(2.8)	11.7(4.1)
Mutter	8.5(2.3)	10.0(3.3)

Wenn schon hieraus eindeutig die unterschiedliche kulturelle Situation der Elternhäuser zwischen den Populationen II und IV zum Ausdruck kommt, so noch stärker in den Berufen der Väter:

Tabelle 6: Berufe der Väter in den Populationen II und IV

Kategorie	Pop.II	Pop.IV
1	0.2	0.0
2	59.0	16.9
3	37.6	62.0
4	3.2	21.2

Hieraus wird also eindeutig die Verschiebung der beruflichen Zusammensetzung der Eltern zwischen der Pop.II

und der Pop.IV infolge der Selektion einsichtig: Während sich in der Pop.II, die praktisch noch sämtliche Schüler eines Jahrgangs enthält, 40.8 Prozent der Schüler in den Kategorien 3 und 4 befinden, sind es in der Pop.IV 83.2 Prozent. Diesbezüglich wird die Bundesrepublik von keinem anderen Land übertroffen. Die entsprechenden Zahlen betragen für die Niederlande 82.0, für Ungarn 55.9 (nur Kategorie 3), für Schweden 52.3 und für Italien 45.2 Prozent. Auffällig ist der hohe Prozentsatz in Ungarn.

Erwähnenswert ist in diesem Zusammenhang noch, daß in der Bundesrepublik nach Angaben der Schüler in der Pop.IV 49.3 Prozent der Väter und 48.4 Prozent der Mütter Englisch gelernt haben, wobei allerdings über den Grad der Sprachbeherrschung keine Aussagen vorliegen. Interessant jedoch ist, daß in allen Ländern, die hohe Leistungen aufzuweisen haben, diese Zahlen entsprechend groß sind. So haben in Pop.IV in den Niederlanden 60.4 Prozent der Väter und 49.7 Prozent der Mütter Englisch gelernt, in Schweden sind es 48.3 bzw. 51.0 Prozent. (Lewis, a.a.O). Darin kommt die bereits erwähnte Tatsache zum Ausdruck, daß ganz offensichtlich für die Leistungsunterschiede zwischen den Ländern eine gewisse Tradition des Fremdsprachenunterrichts in den Schulen eine Rolle spielt. Alle anderen Variablen, die sich auf die Voraussetzung für einen erfolgreichen Unterricht beziehen, haben demgegenüber ein sehr geringes Gewicht. Zum Teil stellen die Relationen sicherlich nur zufällige Schwankungen um den Nullwert dar, was allerdings ein Hinweis darauf wäre, daß die zugrunde liegenden Variablen durch den Fragebogen nur unzureichend erfaßt sein dürften, es sei denn, daß diese Variablen in verschiedenen Ländern entgegengesetzte Wirkungen haben, so daß sie sich gegenseitig aufheben.

Die Angaben über die häusliche Umgebung - insbesondere über ihren möglichen kulturellen Einfluß durch Bücher, Film, Fernsehen, Gewohnheiten usw. - sind sehr unterschiedlich, doch scheint dieser Einfluß in allen Ländern gleichmäßig gering zu sein, genauso wie die Zeit, die die Schüler in englischsprechenden Ländern verbracht haben. Gerade von dieser Variablen hätte man einen relativ großen Einfluß auf die Schulleistung erwartet.
Die Unterschiede in den Schulsystemen scheinen keinen Einfluß auf die Höhe der Leistung zu haben. Schweden mit einem hohen Prozentsatz von Schülern in der Pop.IV hat ausgesprochen gute Leistungen, die Niederlande und die Bundesrepublik mit ihren relativ stark selektiven Systemen haben ebenfalls hohe Leistungen. Dabei muß allerdings gesagt werden, daß die Leistungen der besten 1, 5 und 9 Prozent nicht berechnet worden sind, so daß er bezüglich dieser Aussage auf unsicherem Grund steht.
Die Größe des Kollegiums und die Beschäftigung von Schulassistenten mit Englisch als Muttersprache, der

finanzielle Beitrag, den die Schulen bzw. Behörden für den Englischunterricht auswerfen, wie eine Reihe anderer offensichtlich nicht präzis genug erfaßter äußerer Schuldaten erweisen sich als von nicht eindeutigem Einfluß. Aus allen diesen Tatsachen zieht Lewis den Schluß, daß der Einfluß des Unterrichts auf den Englischunterricht gegenüber allen außerschulischen Einflüssen im Gegensatz zu anderen Schulfächern von einmaliger Bedeutung sei. Er hat sicherlich insoweit recht, als die in den Blöcken II bis IV[1)] zusammengefaßten "schulischen Variablen" in der Pop.II in der Bundesrepublik 57.7 Prozent der Varianz aufklären, in der Pop. IV sind es zwar insgesamt nur 16.9 Prozent, doch sind dies immerhin fast 50 Prozent der insgesamt aufgeklärten Varianz, die für die Pop.IV in der Bundesrepublik nur 34 Prozent beträgt.

2.2.2 Die Lehrervariablen

Die Korrelationen der Leistungen mit den Lehrervariablen sind niedrig bis mittelgroß und wechseln von Land zu Land. Kein eindeutiger Bezug besteht zwischen den Leistungen und dem Geschlecht der Lehrer sowie ihrem Alter. Während in einigen Ländern eine bessere Leistung mit männlichen Lehrern höher korreliert, besteht insgesamt die Tendenz, daß die höhere Leistung mit weiblichen Lehrern verbunden ist. Von positivem Einfluß ist die Dauer der Ausbildung, die sich vor allen Dingen auch in der Selbsteinschätzung der Lehrer hinsichtlich ihrer sprachlichen Kompetenz niederschlägt.
In der Pop.II in der Bundesrepublik besteht eine besonders auffällige Relation zwischen den Leistungen der Schüler und dieser Selbsteinschätzung der Lehrer, wahrscheinlich weil diesbezüglich große Unterschiede zwischen den

---

1) Block I umfaßt die Variablen "sozio-kultureller Schulindex" und das Geschlecht der Schüler;

Block II setzt sich aus den Variablen Schultyp bzw. Schulzweig zusammen;

Block III besteht aus drei Variablen, die die Zeit betreffen: Die Jahre Englischunterricht bis zum Zeitpunkt der Testung, die jetzige Klasse sowie die Klasse und das Alter zu Beginn des Englischunterrichts;

Block IV enthält die auf den Unterricht bezogenen Variablen;

Block V faßt verschiedene Variablen zusammen, die im Schülerfragebogen erfaßt worden sind bzw. aus dem Meinungsfragebogen gewonnen wurden (Stunden der Hausarbeit, Bildungsziele des Schülers, sein Verhältnis zur Schule, seine außerschulischen Aktivitäten, Interesse usw.);

Block VI betrifft die Ergebnisse des Wortschatztests.

Lehrern in den Hauptschulen und den Gymnasiallehrern bestehen.
Kein Unterschied besteht hinsichtlich der Spezialisierung, d.h. Lehrer, die - wie in einigen Ländern üblich - lediglich Englischunterricht erteilen, erreichen keine bessere Leistung als Lehrer, die in mehreren Fächern unterrichten. Eine positive Beziehung besteht zwischen dem Umfang, den die Lehrer auf die Vorbereitung ihres Unterrichts verwenden, und der Schülerleistung. Erstaunlich unterschiedlich sind die Zahlen, die von Lehrern für ihre Vorbereitungszeit angegeben werden. In der Bundesrepublik behaupten 46.5 Prozent aller Lehrer in der Pop.II, mehr als zehn Stunden für ihre Vorbereitungsarbeit auf den Englischunterricht zu benutzen. Die schwedischen Lehrer fallen nur zu 30.2 Prozent in diese Gruppe. In der Pop.IV wenden 68.4 Prozent der Lehrer in der Bundesrepublik mehr als zehn Stunden für ihre Unterrichtsvorbereitung auf, während es in Schweden nur noch 43.3 Prozent sind. In allen anderen Ländern widmen kaum mehr als 25 Prozent der Lehrer mehr als zehn Stunden der Vorbereitungsarbeit. Dabei ist allerdings zu berücksichtigen, daß in diese von Lewis gegebene Übersicht nicht die Zahl der von Lehrern tatsächlich gegebenen Englischstunden eingeht, so daß von daher ein Unsicherheitsfaktor in diesen Zahlen enthalten ist.
Merkwürdig ist, daß die Dauer des Auslandsaufenthalts der Englischlehrer ohne Einfluß auf die Leistung ihrer Schüler zu sein scheint.
Auf die Frage, welchen Einfluß verschiedene Faktoren auf die Ziele und Inhalte des Englischunterrichts haben, ergibt sich für die Bundesrepublik der höchste Prozentsatz an Lehrern, die Prüfungsanforderungen als den wesentlichsten Faktor ansehen. (In Pop. II 74.7 Prozent, in Pop. IV 77.8 Prozent). Von den Lehrern in Finnland und Schweden werden vor allem nachschulische Bedürfnisse des Schülers als maßgeblich für den Englischunterricht angeführt.
Was die methodische Gestaltung des Unterrichts anbelangt, so vermißt Lewis eine einheitliche Methode, die diesbezüglichen Unterschiede zu erfassen. "Inkonsistenz scheint der Trend für alle diese Variablen zu sein" (S.99). Betont wird die allgemeine Neigung der Lehrer, zu Beginn des Englischunterrichts dem Hören ein größeres Gewicht als dem Lesen und Schreiben zu verleihen, und in der allgemeinen Zielsetzung, das Leseverständnis in den Vordergrund zu rücken. Was die Benutzung der Muttersprache anbelangt, so glaubt er, daß der Umfang ihres Gebrauchs mit dem Alter zusammenhängt und daß hier Untersuchungen eine Klärung über das güngstigste Ausmaß auf den verschiedenen Stufen bringen könnten. Der Zeitfaktor erwies sich als der durchschlagendste Faktor im Zusammenhang mit den Leistungen. Aber nach Lewis scheint selbst die Zahl der aufgewendeten Unterrichtsstunden in beiden Populationen nur einen losen Zusammenhang mit der Lei-

stung zu haben, während sich die Zahl der Unterrichtsjahre erst mit wachsender Zahl derselben positiv geltend macht.
Immerhin erklärt der Block III, in dem die Zeitfaktoren zusammengefaßt sind, für die Pop. II in der Bundesrepublik 14 Prozent, für die Pop.IV 7.5 Prozent der Varianz. Differenzierte Angaben über die Wirksamkeit der verschiedenen Zeitfaktoren gehen jedoch aus der Analyse nicht hervor. Als unmittelbar einsichtiges Ergebnis kommt heraus, daß die Leistung umso besser ist, je mehr Jahre Englischunterricht ein Schüler gehabt hat. Die Länge der Unterrichtsstunden, die zwischen 40 und 58 Minuten schwankt, scheint ohne Einfluß auf die Leistung zu sein. Lewis glaubt, daß ein eindeutiges Verhältnis zwischen der Zeit, die dem Englischunterricht gewidmet ist, und der Größe der Englischklassen insofern besteht, als weniger Zeitaufwand mit kleinen Klassen und mehr Zeit mit größeren Klassen verbunden ist. In der Bundesrepublik haben in Pop.II 44.8 Prozent der Schüler Unterricht in Klassen mit 25 und weniger Schülern, während das in Belgien zu 91 Prozent, in Schweden zu 73.2 Prozent der Fall ist. In der Pop.IV ändern sich diese Verhältnisse jedoch erheblich. 91.4 Prozent der Klassen haben höchstens 25 Schüler, in Schweden hingegen nur 75.7 Prozent.

2.2.3 Die Schülervariablen

Während Pop.II ein relativ einheitliches Alter von 14.5 bis 14.8 Jahren in sämtlichen Ländern hat, variiert das Alter der Schüler in Pop.IV wegen der unterschiedlichen Schuldauer stärker und schwankt zwischen 17.8 und 19.2 Jahren. Hinsichtlich der häuslichen Situation ist interessant, daß in der Bundesrepublik nach den Angaben der Schüler das Interesse der Eltern der Pop.II an den Schulleistungen im umgekehrten Verhältnis zu den Angaben über das Ausmaß an Hilfen bei den Hausaufgaben steht. Wahrscheinlich läßt diese Tatsache Rückschlüsse auf die Zuverlässigkeit der Angaben im Schülerfragebogen zu. Deshalb muß man sicherlich ein Fragezeichen hinter die von den Schülern gemachten Angaben über das Ausmaß der Hausaufgaben und den Umfang der elterlichen Hilfe setzen. In Schweden geben 46 Prozent der Schüler an, wöchentlich weniger als zwei Stunden Hausarbeit für sämtliche Unterrichtsfächer aufzuwenden, was verständlich wird, wenn man berücksichtigt, daß sie Ganztagsschulen besuchen. Wegen dieser unterschiedlichen Situation der Schüler in den beteiligten Ländern wird verständlich, daß die Angaben der Schüler über das Ausmaß der Hausaufgaben in keiner klaren Relation zu den Unterschieden in den Leistungen stehen. Interessant ist, daß die Länder, die offenbar den geringsten Anteil an Hausarbeit von ihren Schülern erwarten, den höchsten Prozentsatz von Schülern in kleinen Klassen haben. So werden in Belgien in der Pop.IV z.B. mehr als ein Drittel der Schüler in Klassen

mit weniger als 15 Schülern unterrichtet. In Israel haben hingegen mehr als 45 Prozent der Klassen mehr als 25 Schüler.

Allgemein zeigt das Bild beträchtliche Unterschiede zwischen den Ländern und eine ganz geringe Konstanz hinsichtlich des Verhältnisses zwischen den Variablen und der Leistung. Während sich nach Lewis selbst in bezug auf die Korrelation zwischen dem Interesse am Englischunterricht und der Leistung keine durchgehend gleichsinnige Tendenz ergibt, die Erwartungen der Schüler in bezug auf ihren Bildungsgang sowie die Selbsteinschätzung ihrer Englischleistung jedoch in den meisten Ländern eine mittlere bis hohe Korrelation mit der tatsächlichen Leistung zeigen, stellt C.Massad hingegen fest, daß sich die Variablen "Interesse" am Englischunterricht und "Bildungserwartung" als gleichsinnig wirksame Variablen durch alle Populationen hindurch erwiesen haben.

Alle anderen Schülervariablen sind jedoch ungleichmäßig hinsichtlich ihres Gewichts und ihrer Richtung für die verschiedenen Populationen und Länder. Die meisten Korrelationen sind niedrig, ja zeigen zum Teil gegenteilige Tendenz, so daß die Ergebnisse im Hinblick auf das eigentliche Ziel der Untersuchung auch hier als relativ mager bezeichnet werden können. Es ist schließlich keine große Einsicht, wenn es in dem Beitrag von C.Massad heißt: "Das Interesse am Englischen scheint der einzige relativ stetige Prädiktor für englisches Lesen und Hörverstehen zu sein. Das bedeutet, daß Motivation ein Sammelfaktor beim Englischlernen sein könnte" (S.285).

2.2.4 Zusammenfassung der Ergebnisse der Regressionsanalyse

In der Tabelle 7 geben wir noch einmal eine Übersicht über das gesamte Ergebnis der Regressionsanalyse (nach C.Massad Tab.9.1 und Tab. 9.3, S.288/291).

Tabelle 7: Prozentsätze der durch die Blöcke I-VI erklärten Varianz in der Bundesrepublik

B l o c k	Pop.II	Pop.IV
I Schulindex und Geschlecht	15.1	1.0
II Schultyp bzw. Schulzweig	30.7	0.5
III Zeitfaktoren	14.0	7.5
IV Unterrichtsvariablen	13.0	8.9
V Verschiedene Variablen (Interesse usw.)	3.2	5.7
VI Wortschatz	2.5	10.4
Insgesamt	78.5	34.0

Es zeigt sich, daß auch in dieser Untersuchung durch die erfaßten Variablen für Pop.II ein höherer Anteil der Varianz erklärt wird als für die Pop.IV. Das wird aus dem großen Einfluß der verschiedenen Schulformen auf die Leistungsunterschiede verständlich. International betrug der Anteil der erklärten Varianz für die Pop.II 55.7 und für die Pop.IV 42.0 Prozent. Wenn der Wert für die Bundesrepublik in Pop. IV niedriger liegt, so ist das verständlich, weil der Grad der Streuung bei der Ermittlung der Höhe der Varianz eine große Rolle spielt und diese Streuung in der Bundesrepublik die geringste aller beteiligten Länder ist. Wenn also in der Pop.II 30.7 Prozent der Varianz durch den Block II erklärt wird, so besagt das weiter nichts, als daß in der Bundesrepublik die Schulart (Haupt-, Realschule und Gymnasium) eine deutliche Voraussage hinsichtlich der Leistung der sie besuchenden 14jährigen Schüler zuläßt. Über den Beitrag der schulischen Variablen insgesamt haben wir uns bereits im vorigen Abschnitt geäußert. Die Tatsache, daß die Schulvariablen insgesamt ein erhebliches Gewicht haben, daß jedoch die Variablen im einzelnen betrachtet eine große Ungleichmäßigkeit zeigen, ist ein weiterer Hinweis dafür, daß die Meßinstrumente zur detaillierten Erfassung der Variablen nicht ausreichten bzw. durch die Übersetzung in verschiedene Sprachen an Eindeutigkeit verloren haben. Wenn die Autorin dieses Kapitels, Carolyn E. Massad, glaubt, daß "die Ergebnisse, über die in diesem Kapitel berichtet wird, nur als Anfang einer Serie von Analysen und weiterer Forschung in den hier untersuchten Gebieten betrachtet werden sollen", und im übrigen auf eine weitere Analyse der in der Datenbank gesammelten Fakten und auf zusätzliche Studien vertraut, dürfte kaum zu erwarten sein, daß bessere, aussagekräftigere Ergebnisse aus den in dieser Untersuchung erfaßten Daten herauskommen. Die beste Analyse kann nichts herausbringen, wenn die durch die Meßinstrumente hereingeholten Daten unzureichend sind.

2.3 Zusammenfassung der Ergebnisse durch Lewis

Der Verfasser glaubt, daß die Meßinstrumente und Fragebogen ganz allgemein eine hinreichende Validität für die Länder, in denen sie angewendet wurden, aufweisen. Die Ergebnisse der Untersuchung belegen jedoch eigentlich das Gegenteil. Lewis stellt weiterhin fest: "Die Faktoren, die über die Länder hinweg sich als klare Prädiktoren für den Erfolg im Englischunterricht erweisen, scheinen die Zeit, die Unterrichtspraxis in der Klasse sowie die Haltung und Erwartung der Schüler zu sein." Eine solche Aussage hätte allerdings nicht einer so umfassenden, nicht ganz einfachen und darüber hinaus kostspieligen Studie bedurft. Und da sich die Zeit als wichtiger Faktor für die erreichte Leistung erwiesen hat, wird

empfohlen, zunächst eindeutig die Leistungsanforderungen, die zu bestimmten Zeitpunkten im Englischunterricht erreicht werden sollen, festzulegen. Die Ziele sollten nach den verschiedenen Erwartungen der Schüler unterschiedlich sein; denn die Bedürfnisse der Schüler sollten sich im Lehrplan niederschlagen. So sollten Schüler, die nur eine Lesefertigkeit im Englischen erwerben wollen (reading ability) einem anderen Lehrgang folgen als Schüler, die gleichzeitig sprechen bzw. schreiben wollen.

Nachdem die Leistungsanforderungen festgelegt worden sind, wäre das Alter für den Beginn des Englischunterrichts sowie seine Dauer für die verschiedenen Ziele zu bestimmen. Er gibt jedoch zu, daß die vorliegende Untersuchung dafür keine hinreichende Information liefert, und betont, daß dafür eingehendere Untersuchungen mit weiteren Instrumenten, wenn möglich auch Beobachtungen in der Klasse, erforderlich sind. Selbstverständlich ist seine Folgerung, daß in der Lehrerbildung alles getan werden sollte, um die Leistungen und damit das Bewußtsein des Lehrers von seiner Leistungsfähigkeit zu steigern.
Hinsichtlich der Arbeit in der Klasse kommt es darauf an, im Unterricht ein ausgewogenes Verhältnis zwischen dem Gebrauch der Muttersprache und dem Gebrauch des Englischen im Unterricht zu erreichen. Aber über eine richtige Relation auf den verschiedenen Altersstufen kann auch die vorliegende Untersuchung keine Auskunft geben. Bedauert wird, daß kein einheitliches Instrument besteht, um die Unterrichtsmethoden zu beschreiben und ein Methodenprofil zu erarbeiten, das zu den Unterschieden der Leistung in Beziehung gesetzt werden kann. Außerdem wäre ein einfaches Instrument zu entwickeln, um einheitlich die Aktivitätsformen der Schüler innerhalb der Klasse zu beschreiben. Das allerdings ist, wie jeder weiß, der einschlägig gearbeitet hat, ein frommer Wunsch, der seit langem die Forschung bewegt.

3. Die Ergebnisse in der Bundesrepublik

Im folgenden geben wir eine Analyse des uns zugegangenen Grundmaterials für die Ergebnisse innerhalb der Bundesrepublik. Für diese Analyse gilt das, was bereits als Vorbemerkung zu unserem Bericht über den naturwissenschaftlichen Unterricht gesagt worden ist: Es können nur die Zahlen zugrunde gelegt werden, die uns als Computerauswurf zugegangen sind. Für weitere Fragestellungen wäre vielleicht ein Rückgriff auf die Magnetbänder der Datenbank nützlich gewesen. Er hätte jedoch den finanziellen und arbeitsmäßigen Aufwand aus den schon genannten Gründen nicht gelohnt. Außerdem sind seit Beginn der Studie bis zur Veröffentlichung des internationalen Berichts neun Jahre vergangen. Selbst seit Durchführung der Testung (1971) ist eine so lange Zeit vergangen, daß Rückfragen bei den beteiligten Schulen kaum noch Sinn gehabt hätten. Aus dem gleichen Grunde wurden den Schulen Ergebnisse nicht mehr mitgeteilt, da sich gerade in den letzten Jahren die Verhältnisse in den Schulen durch organisatorische Umwandlungen, die sich auf den Unterricht, die Zusammensetzung der Lehrkörper und der Schülerschaft auswirkten, vielfach geändert haben.

3.1 Zur Stichprobe

Tabelle 1 gibt einen Überblick über die Stichprobe. Die Gründe für die geringere Zahl von Schulen, Schülern und Lehrern, die dieser Stichprobe gegenüber der Untersuchung des naturwissenschaftlichen Unterrichts zugrunde gelegt wurde, sind bereits erwähnt worden. Trotzdem können wir die Stichprobe für die Bundesrepublik als repräsentativ ansehen, wofür ein Vergleich der erreichten Leistungen im Wortschatztest für die beiden Stichproben im naturwissenschaftlichen Unterricht und im Englischunterricht einen weiteren Hinweis gab (s.S.3).
Eine weitere Stütze erfährt diese Annahme durch einen Vergleich der Berufsgruppen der Väter in den beiden Untersuchungen (s. S.24).

Zu Tab. 1: Die geringe Zahl der Schüler in der Pop.II gestattet keine Aufschlüsselung nach Ländern, wenn man gleichzeitig noch eine Untergliederung nach Hauptschulen, Realschulen und Gymnasien vornimmt. Es dürfte auffallen, daß das Verhältnis von Hauptschulen zu Realschulen und zu Gymnasien nicht den tatsächlichen statistischen Verhältnissen entspricht. Deshalb wurden für die weitere Analyse die Zahlen entsprechend gewichtet. Gemäß dem Zwischenbericht der Bund-Länder-Kommission verteilten sich die 13jährigen Schüler 1969 folgendermaßen:
Auf je 100 Hauptschüler fallen 32 Realschüler und 36 Gymnasiasten.
Diese Zahlen gelten jedoch nur für die Bundesrepublik

und dürfen nicht auf die Länder übertragen werden, da die Verhältnisse sehr unterschiedlich sind. Auch aus diesem Grunde verbietet sich eine Aufschlüsselung der Ergebnisse nach Ländern, da die diesbezüglichen Unterschiede zwischen den Ländern bei der Aufstellung der Stichprobe nicht berücksichtigt werden konnten.
Anders liegen die Verhältnisse für Pop.IV.

Tabelle 8: Zahl der an der Untersuchung beteiligten Schulen und Schüler der Pop.IV, aufgeschlüsselt nach Bundesländern

L ä n d e r	Schulen	Schüler
Baden-Württemberg	10	211
Bayern	8	149
Hessen	7	165
Niedersachsen	10	207
Nordrhein-Westfalen	13	312
Rheinland-Pfalz	6	145
Saar	3	74
Schleswig-Holstein	4	116
S u m m e	61	1379

In den 61 Schulen waren insgesamt 360 Lehrer beteiligt.

Die Zusammensetzung unserer Stichproben nach den Berufen der Väter würde einen Vergleich mit den Stichproben "naturwissenschaftlicher Unterricht" ermöglichen. Da, wie bereits gesagt, die Erfassung der Väter nach anderen Kategorien erfolgte als in der Studie über den naturwissenschaftlichen Unterricht, läßt sich ein solcher Vergleich nur bedingt durchführen. In den Tabellen 8.1 und 8.2 gibt Lewis die Zusammensetzung der Stichprobe nach den von ihm gewählten Kategorien, wobei die Kategorien 1-4 eine grobe Erfassung ermöglichen, wobei Kategorie 1 die niedrigste, Kategorie 4 die höchste Berufsgruppe umfaßt. Fassen wir zum Zwecke des Vergleichs die Kategorien 1 und 2 sowie 3 und 4 zusammen, so können wir ihnen entsprechende Kategorien der naturwissenschaftlichen Studie zuordnen. Das ergibt dann folgendes Bild:

Tabelle 9: Die Populationen nach den Berufsgruppen der Väter (in Prozent)[1)]

	P o p. II		P o p. IV	
Berufsgruppen	Naturw.	Engl.	Naturw.	Engl.
untere	46	59	8	17
obere	46	41	87	83

[1)] Die Summe ergibt bei der naturwissenschaftlichen Untersuchung deshalb nicht 100, weil eine Kategorie 9 "ohne Angaben" vorhanden ist.

Im großen und ganzen kann man wohl von einer befriedigenden Übereinstimmung sprechen, wenngleich der Anteil der Schüler aus den unteren Berufsgruppen in der Stichprobe "Englisch" in beiden Populationen etwas größer zu sein scheint. Auf die mit der Selektion bis zur letzten Klasse der Sekundarstufe II verbundene "soziale Auslese" sei trotz ständiger Wiederholung ausdrücklich hingewiesen.

3.2 Population II

Der internationale Vergleich zeigte bereits, daß die Leistungen der Pop.II in der Bundesrepublik insgesamt über dem internationalen Durchschnitt liegen. Die durchschnittliche Leistung der Population in der Bundesrepublik beträgt nach Lewis 29.9 Punkte bei einer Standardabweichung von 16.2 Punkten. Diese Zahl ist jedoch falsch, weil bei der Ermittlung des Durchschnitts Schüler mitgezählt wurden, deren Werte unbekannt sind und deshalb vom Computer nicht in die Durchschnittsrechnung mit einbezogen werden konnten. Aufgrund der uns vorliegenden Unterlagen beträgt der ungewichtete Durchschnitt für die Pop.II 36.0 Punkte, der gewichtete Durchschnitt 32.4. Er liegt also höher als bei Lewis angegeben. Entsprechend müßte auch die Abbildung 1 geändert werden. In Tab 10 schlüsseln wir diese Werte auf nach Hauptschulen, Realschulen und Gymnasien.

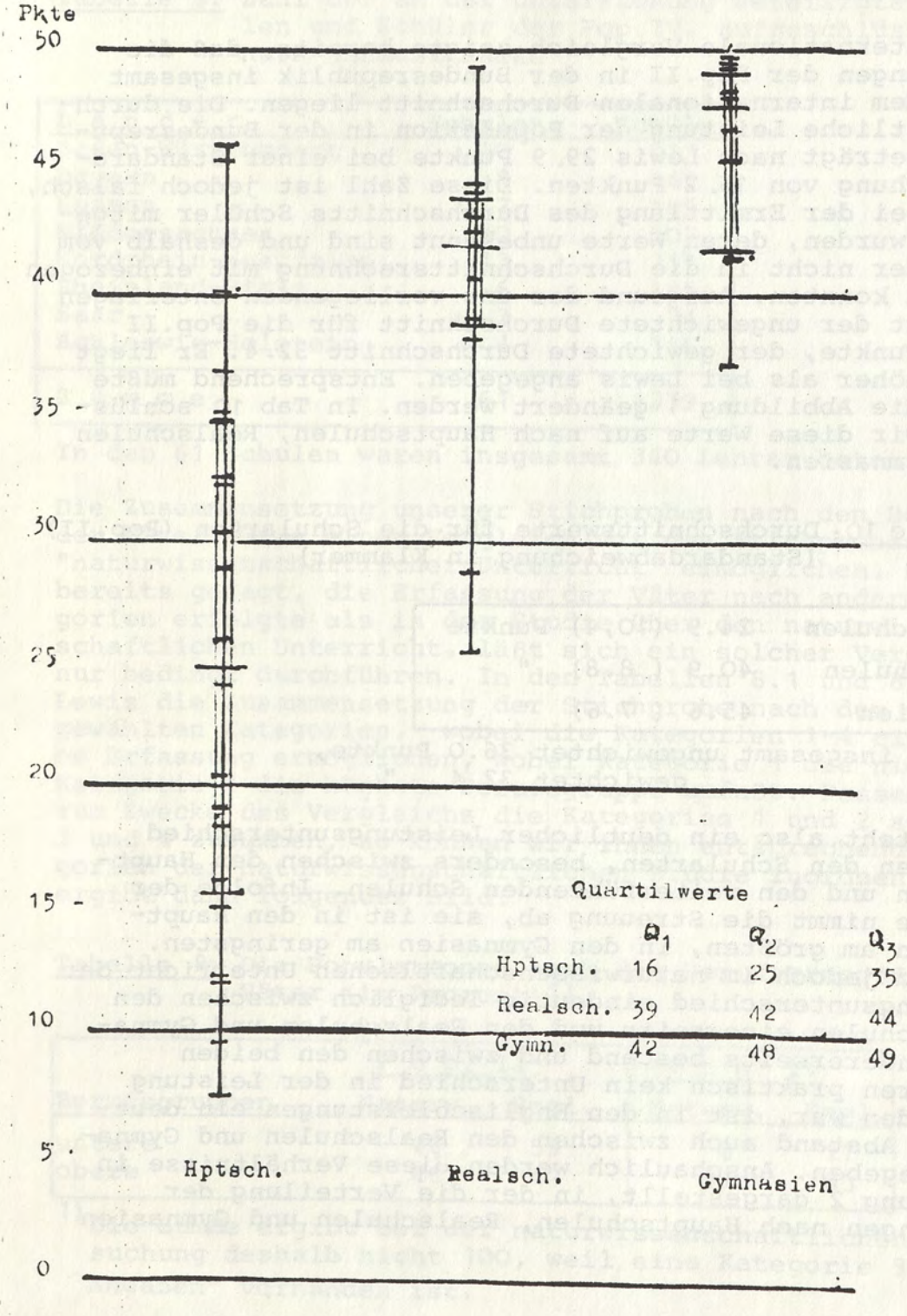
Tabelle 10: Durchschnittswerte für die Schularten (Pop.II)
(Standardabweichung in Klammer)

Hauptschulen	24.9 (10,4)	Punkte
Realschulen	40.9 (8.8)	"
Gymnasien	45,6 (7.6)	"

Pop.II insgesamt ungewichtet 36.0 Punkte,
gewichtet 32.4 " .

Es besteht also ein deutlicher Leistungsunterschied zwischen den Schularten, besonders zwischen den Hauptschulen und den weiterführenden Schulen. Infolge der Auslese nimmt die Streuung ab, sie ist in den Hauptschulen am größten, in den Gymnasien am geringsten. Während jedoch im naturwissenschaftlichen Unterricht der Leistungsunterschied eindeutig lediglich zwischen den Hauptschulen einerseits und den Realschulen und Gymnasien andererseits bestand und zwischen den beiden letzteren praktisch kein Unterschied in der Leistung vorhanden war, ist in den Englischleistungen ein deutlicher Abstand auch zwischen den Realschulen und Gymnasien gegeben. Anschaulich werden diese Verhältnisse in Abbildung 2 dargestellt, in der die Verteilung der Leistungen nach Hauptschulen, Realschulen und Gymnasien

Abbldg.2 Verteilung der Leistungen in Pop.II nach Hauptschulen, Realschulen und Gymnasien



Quartilwerte

	Q_1	Q_2	Q_3
Hptsch.	16	25	35
Realsch.	39	42	44
Gymn.	42	48	49

veranschaulicht wird. Die Durchschnittswerte der einzelnen Schulen werden dabei im Stabdiagramm dargestellt, wobei sich allerdings infolge der von Schule zu Schule unterschiedlichen Schülerzahlen ein wenig andere charakteristische Werte ergeben. Der Medianwert der Realschulen beträgt 42, der Medianwert der Gymnasien 48, während der Medianwert der Volksschulen sich praktisch nicht von der mittleren Standardabweichung unterscheidet. Weiterhin ist aus der Abbildung 2 zu entnehmen, daß die Streuung zwischen den Schulen im Bereich der Hauptschule noch sehr viel größer ist,als in der durchschnittlichen Standardabweichung der Schüler von 10.4 zum Ausdruck kommt. Die Streuung bei den Realschulen ist im wesentlichen auf wenige Schulen, die stark nach oben oder unten vom Durchschnitt abweichen, zurückzuführen, während die Q_1- und Q_3-Werte enger zusammenliegen als bei den Gymnasien. Schließlich bleibt festzustellen, daß es Hauptschulen gibt, deren Durchschnittsleistungen in den Leistungsbereich der Realschulen und Gymnasien hineinfallen, wie es andererseits eine noch größere Zahl von Realschulen gibt, die im Leistungsbereich der Gymnasien liegen.
Wie wir oben gesehen haben (s.S. 17), sind 57.7 Prozent der Varianz der Schülerleistungen in Pop.II auf schulische Variablen zurückzuführen. Ihnen im einzelnen nachzugehen, ist aufgrund der vorhandenen Unterlagen kaum möglich. Es bedürfte hier einer eingehenden Analyse, die sich insbesondere auf die konkrete Situation in den einzelnen Schulen mit ausgesprochen hohen und ausgesprochen niedrigen Leistungen in allen drei Schularten zu erstrekken hätte. Wir haben die Leistungen der einzelnen Schulen in Relation zu den Ergebnissen im Wortschatztest gestellt, weil angenommen werden darf, daß die im letzteren erfaßliche sprachliche Begabung sich in den Englischleistungen niederschlägt. Wenn sich auch international der Wortschatztest als ungeeignetes Werkzeug erwiesen hat, so hat sich seine Brauchbarkeit für Analysen innerhalb der Länder durchaus bestätigt.
Der gewichtete Durchschnitt im Wortschatztest beträgt für die Pop.II 15.8 Punkte. Für die einzelnen Schularten ergeben sich folgende Werte:
Für die Hauptschulen 11.8,
für die Realschulen 18.7 und
für die Gymnasien 24.4 Punkte.
Der Abstand von Schulform zu Schulform ist also deutlich. Ein gutes Bild dieser Beziehung ergibt auch eine Analyse der Extremfälle. Wählen wir die Hauptschulen, die um mehr als $\pm$ 5 Punkte vom Durchschnittswert 25 abweichen, so ergibt sich die Tabelle 11.

Tabelle 11: Die Leistungen der im Englischtest um mehr als ± 5 Punkte vom Durchschnitt abweichenden Hauptschulen der Pop.II sowie die Leistungen im Wortschatztest (Streuung in Klammer)

+ 5			- 5		
Schule	Engl.	Wortsch.	Schule	Engl.	Wortsch.
407	32.5(7.9)	15.1	403	12.1(7.9)	2.3
435	40.0(8.8)	14.4	422	12.4(10.3)	12.6
437	35.1(6.7)	17.0	458	7.6(9.1)	7.0
440	45.5(8.1)	14.6	467	19.4(5.9)	12.7
442	30.7(12.2)	13.5	470	15.2(7.7)	14.7
444	37.0(6.8)	21.3	471	19.0(18.8)	28.3
447	35.3(6.8)	10.9	479	18.7(6.6)	14.4
448	33.4(10.4)	11.1	486	13.8(11.0)	11.3
449	40.0(9.2)	16.5	488	9.8(8.5)	7.4
481	45.9(7.9)	27.6	496	16.3(13.0)	9.6
N = 10			N = 10		
D =	37.1(8.7)	15.1	D =	12.7(10.4)	9.5
	r = 0.52			r = 0.77	

Es handelt sich in beiden Fällen um je zehn Hauptschulen. Ein erster Blick auf die Tabelle zeigt, daß die Streuung unter den Schulen dieselbe Tendenz zeigt,wie der Vergleich der Streuung zwischen den Haupt-, Realschulen und Gymnasien ergeben hat. Durchschnittliche Standardabweichung der positiv abweichenden Schulen beträgt 8.7, der negativ abweichenden Schulen 10.4 Punkte. Die Unterschiede in den Durchschnittsleistungen dieser beiden Gruppen sind erstaunlich und geben ein deutliches Bild von den Leistungsunterschieden gerade innerhalb der Hauptschulen. Die Korrelation der Englischleistung mit dem Wortschatztest ergibt r = 0.52 bzw. r = 0.77. Wenn wir andererseits aus Untersuchungen wissen, daß der Wortschatztest in hohem Grade mit dem sozio-ökonomischen Schulindex korreliert, so dürfen wir annehmen, daß die Leistungsunterschiede zwischen den Schulen zu einem großen Teil auf Unterschiede im sprachlichen Leistungspotential ihrer Schüler und in der zum Teil damit zusammenhängenden sozio-kulturellen Situation des Elternhauses zurückgeführt werden können (s.hierzu auch die Ergebnisse der Regressionsanalyse in bezug auf das Gewicht der Blöcke I und VI in Pop.II).

Bei den Realschulen ergibt sich folgendes Bild: Die positiv abweichenden Schüler haben einen Leistungsdurchschnitt von 46.1 (6.6), die schlechten Schulen einen Schülerdurchschnitt von 33.0 (10.5). Das Bild ist also ebenso eindeutig, was die Streuung anbelangt. Wegen der geringen Zahl von Schulen ist es statistisch kaum gerechtfertigt, die Korrelation zwischen den Englischleistungen und dem Wortschatz auszurechnen. Das Gleiche gilt für die Gymnasien, obgleich auch hier die Korrela-

tion in allen Fällen ausgesprochen positiv, wobei die Korrelationen bei den negativ herausfallenden Schulen stets sehr viel höher sind als die Korrelationen bei den positiv herausragenden Schulen. Doch wäre es verfrüht, hieraus Schlüsse zu ziehen, da sich diese Erscheinung in Pop.IV nicht bestätigt. Die Durchschnittsleistungen der Schüler betragen bei den positiv herausragenden Gymnasien 49.5 Punkte (5.2), bei den negativ herausfallenden Schulen 41.7 (7.8) Punkte. Wieweit die Tatsache, daß in allen drei Schularten schlechte Durchschnittsleistungen immer zugleich mit größerer Streuung der Schülerleistungen verbunden sind, die Vermutung zuläßt, daß in diesen Schulen entweder stärker individualisiert wird, was jedoch kaum der Fall sein dürfte, oder aber, daß in ihnen eine besonders schwierige Unterrichtssituation infolge einer größeren Begabungsbreite der Schülerschaft besteht, ist nicht mit Sicherheit zu sagen. Allerdings hätten nachgehende Untersuchungen vielleicht doch zur Aufklärung dieses Tatbestandes beitragen können. Diese sind jedoch vier Jahre nach der Prüfung nicht mehr möglich.

3.3 Population IV

Wie bereits im internationalen Vergleich dargestellt, ist das Durchschnittsergebnis für die Bundesrepublik insgesamt mit 41.2 Punkten und einer Standardabweichung von 6.1 außerordentlich günstig. Wir haben auch bereits darauf hingewiesen, daß das Ergebnis zum Teil auf die im Vergleich zu anderen Ländern strenge Auslese des Abiturientenjahrgangs zurückzuführen sein dürfte.
Bei der Durchsicht der Ergebnisse fällt zunächst einmal auf, daß es einzelne Schulen mit nur ganz geringen Streuungswerten von 2.2 (Nr.627), 3.2 (Nr.668/669) und 3.3 Punkten (Nr.622) gibt, während sich andererseits Schulen mit relativ hohen Werten von 7.3 (Nr.644), 7.4 (Nr.649), 8.6 (Nr.642) und 8.7 (Nr.661) finden. Dabei verteilen sich die Schulen mit großen bzw. kleinen Streuungen unregelmäßig über sämtliche Länder der Bundesrepublik. Die gute Verteilung der Durchschnittsleistung in den Schulen ist aus dem linken Stabdiagramm, Abbildung 3, ersichtlich. Aus ihm ergibt sich aber auch, daß zwischen den Schulen erhebliche Unterschiede in der Leistung bestehen. Besteht doch zwischen der niedrigsten Durchschnittsleistung einer Schule und der höchsten Durchschnittsleistung der besten Schule ein Unterschied von 15.8 Punkten.
Greifen wir - wie wir es für Pop.II getan haben - Schulen mit positiv herausragenden und Schulen mit negativ herausfallenden Leistungen heraus, indem wir Schulen auswählen, die mit $\pm$ 3 vom Medianwert 41 abweichen, so ergibt sich das Bild der Tabelle 12.

Tabelle 12: Die Leistungen der positiv und negativ im Englischtest um mehr als 3 Punkte vom Mittelwert abweichenden Schulen der Pop.IV und ihre Leistungen im Wortschatztest

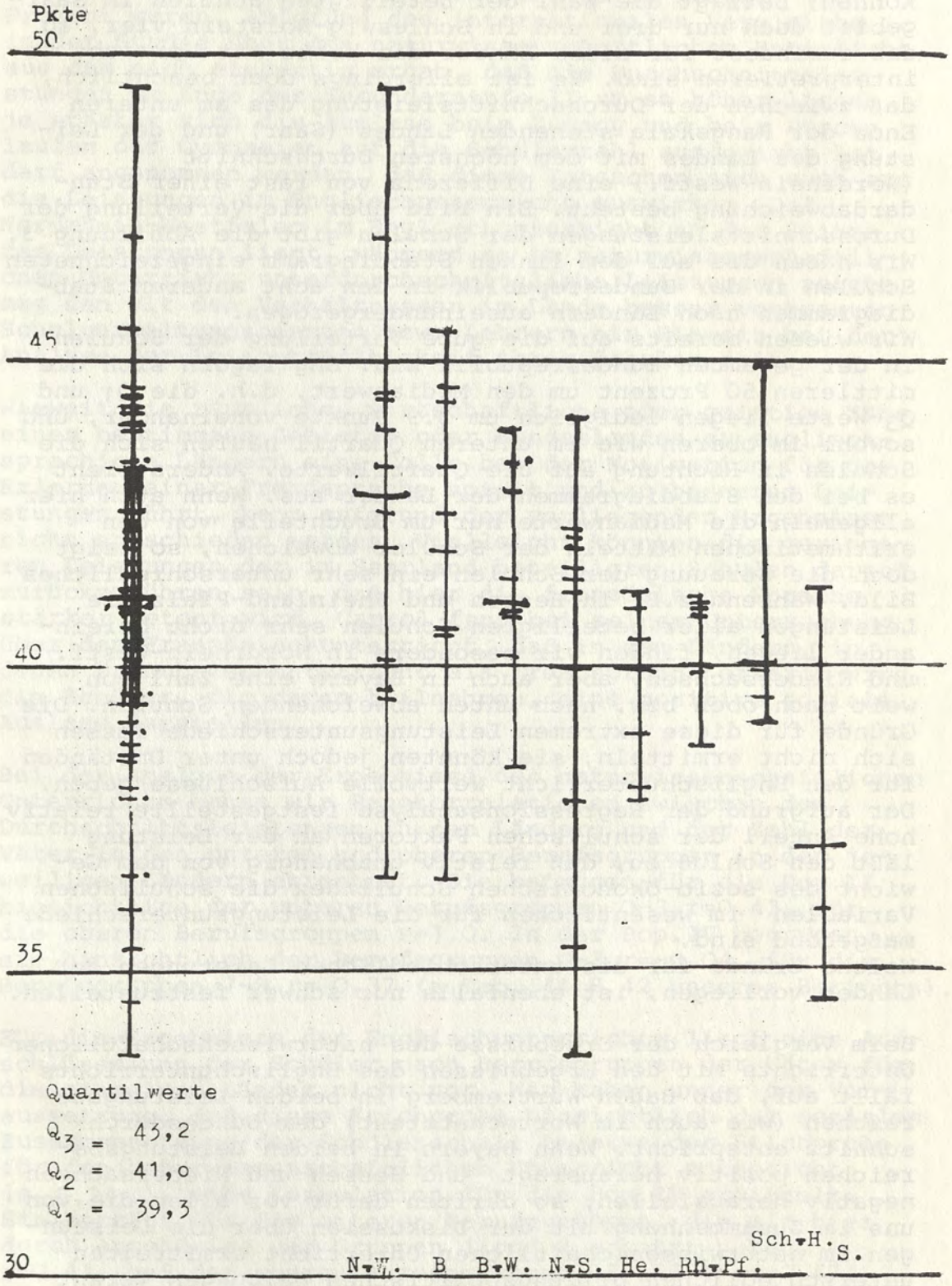
+ 3			- 3		
Schule	Engl.	Wortsch.	Schule	Engl.	Wortsch.
618	44.9 (3.9)	31.3	624	36.5 (5.4)	17.1
620	47.0 (5.5)	29.4	642	36.5 (8.6)	23.4
622	44.4 (3.3)	26.7	660	37.8 (7.8)	20.2
623	44.3 (4.0)	24.1	661	33.6 (8.7)	18.0
627	49.4 (2.2)	32.9	670	35.4 (8.2)	19.1
630	44.3 (4.9)	31.7	675	37.8 (4.4)	20.7
640	45.5 (4.7)	30.4	694	36.0 (7.2)	28.0
643	44.5 (5.5)	27.7	695	34.5 (6.7)	21.9
668	44.1 (3.2)	27.9			
689	44.9 (5.6)	28.2			
N=10; D=45.2(4.4)		28.7	N=8; D=36.0(7.2)		21.4
	r=0.50			r=0.05	

Zunächst einmal zeigt sich auch hier das bereits in der Pop.II in allen drei Schulformen einheitliche Bild, daß die höheren Leistungen mit niedrigeren Streuungen verbunden sind. Die durchschnittliche Streuung der 10 positiv herausragenden Schulen beträgt 4.4, die der 8 negativ herausfallenden Schulen 7.2 Punkte; die Korrelationen sind hier im Gegensatz zur Pop.II hoch, bei den positiv herausragenden Schulen (r=0.50), bei den negativ herausfallenden Schulen ist sie praktisch gleich null (r=0.05). Doch kann man eindeutig feststellen, daß mit einer größeren Zahl von Ausnahmen die Durchschnittsleistungen im Wortschatztest im allgemeinen den schlechteren bzw. höheren Leistungen im Englischtest entsprechen. Die durchschnittlichen Leistungen im Wortschatztest betragen 28.7 bzw. 21.4 Punkte.
Tab.13 gibt eine Aufschlüsselung der durchschnittlichen Leistung im Englischtest nach Bundesländern.

Tabelle 13: Durchschnittliche Leistungen der Schüler in Pop.IV nach Bundesländern (Standardabweichung in Klammer)

Baden-Württemberg	41.0 (5.9)	Punkte
Bayern	41.8 (6.2)	"
Hessen	39.5 (5.5)	"
Niedersachsen	39.8 (7.2)	"
Nordrhein-Westfalen	40.8 (5.6)	"
Rheinland-Pfalz	40.4 (5.3)	"
Saar	37.1 (6.9)	"
Schleswig-Holstein	40.6 (6.0)	"
	41.2 (6.1)	"

Abbldg. 3 Die Leistungen der einzelnen Schulen in Pop.IV in der Bundesrepublik und nach Ländern



Wir weisen ausdrücklich darauf hin,daß diese Zahlen nicht für alle Bundesländer als repräsentativ angesehen werden können, beträgt die Zahl der beteiligten Schulen im Saargebiet doch nur drei und in Schleswig-Holstein vier, so daß zumindest für diese Länder die Werte mit Vorsicht zu interpretieren sind. Es ist allerdings doch beachtlich, daß zwischen der Durchschnittsleistung des am unteren Ende der Rangskala stehenden Landes (Saar) und der Leistung des Landes mit dem höchsten Durchschnitt (Nordrhein-Westf.) eine Differenz von fast einer Standardabweichung besteht. Ein Bild über die Verteilung der Durchschnittsleistungen der Schulen gibt die Abbildung 3. Wir haben die auf dem linken Stabdiagramm eingezeichneten Schulen in der Bundesrepublik in den acht anderen Stabdiagrammen nach Ländern auseinandergezogen.
Wir wiesen bereits auf die gute Verteilung der Schulen in der gesamten Bundesrepublik hin. Eng lagern sich die mittleren 50 Prozent um den Medianwert, d.h. die Q_1 und Q_3-Werte liegen lediglich um 3.9 Punkte voneinander, und sowohl im oberen wie im unteren Quartil häufen sich die Schulen in Richtung auf die Quartilwerte. Anders sieht es bei den Stabdiagrammen der Länder aus. Wenn auch hier allgemein die Medianwerte nur um Bruchteile von den arithmetischen Mitteln der Schüler abweichen, so zeigt doch die Streuung der Schulen ein sehr unterschiedliches Bild. Während z.B. in Hessen und Rheinland-Pfalz die Leistungen aller beteiligten Schulen sehr dicht beieinander liegen, finden wir besonders in Nordrhein-Westf. und Niedersachsen, aber auch in Bayern eine Zahl von weit nach oben bzw. nach unten abweichenden Schulen. Die Gründe für diese extremen Leistungsunterschiede lassen sich nicht ermitteln, sie könnten jedoch unter Umständen für den Englischunterricht wertvolle Aufschlüsse geben. Der aufgrund der Regressionsanalyse festgestellte relativ hohe Anteil der schulischen Faktoren an der Leistung läßt den Schluß zu, daß relativ unabhängig von dem Gewicht des sozio-ökonomischen Schulindex die schulischen Variablen im wesentlichen für die Leistungsunterschiede maßgebend sind.
Welche Gründe für die unterschiedlichen Leistungen der Länder vorliegen, ist ebenfalls nur schwer festzustellen.

Beim Vergleich der Ergebnisse des naturwissenschaftlichen Unterrichts mit den Ergebnissen des Englischunterrichts fällt auf, daß Baden-Württemberg in beiden Leistungsbereichen (wie auch im Wortschatztest) dem Bundesdurchschnitt entspricht. Wenn Bayern in beiden Leistungsbereichen positiv herausragt und Hessen und Niedersachsen negativ herausfallen, so dürften dafür vor allem die von uns im Zusammenhang mit der Diskussion über die Leistungen im naturwissenschaftlichen Unterricht ermittelten unterschiedlichen bildungspolitischen Haltungen verantwortlich sein. In Bayern lag der Prozentsatz eines Altersjahrgangs, der zur Zeit der Untersuchung die Abiturklassen

besuchte, niedriger als in Hessen. In Bayern wurden z.B. nur 8.6 Prozent der im 4.Schuljahr 1961 vorhandenen Schüler zum Abitur geführt, während es in Hessen 14.6 Prozent waren. Aufgrund des internationalen Vergleichs in der Studie über den naturwissenschaftlichen Unterricht, aus dem sich eindeutig ergab, daß die Durchschnittsleistungen am Ende der Sekundarstufe II um so höher liegen je stärker sich die Auslese beim Zugang und beim Durchlaufen der Gymnasien auf die Schülerzahl ausgewirkt hat, darf angenommen werden, daß diese Tatsachen sich auch auf die Leistungen im Englischunterricht auswirken. Daß Nordrhein-Westfalen im Englischunterricht an der Spitze der Leistungen liegt, während es im naturwissenschaftlichen Unterricht unterdurchschnittliche Leistungen zeigte, mag den mit den Verhältnissen im Lande besser vertrauten Schulverwaltungsbeamten bzw. Lehrern ein Hinweis bei der Analyse der Leistungsfähigkeit ihrer Schulen sein.

Wieweit die räumliche, wirtschaftliche oder geistige Nähe eines bestimmten Gebietes oder Bundeslandes zu englischsprachigen Ländern eine Rolle bei der Motivation für das Erlernen einer Fremdsprache spielt und zu besseren Leistungen führt, kann aufgrund der vorliegenden Ergebnisse nicht entschieden werden. Vielleicht könnten die geringeren Leistungen der im Saarland beteiligten Schulen darauf zurückzuführen sein, daß hier die französische Sprache stärker betont wird. Carrol fand bei seiner Untersuchung über den Französischunterricht, daß in den Ländern, in denen Französischunterricht als Fremdsprache erteilt wird, die Schüler, die daran teilnehmen, eine positive soziale Auslese darstellen.

Bei der Analyse der Ergebnisse des naturwissenschaftlichen Unterrichts haben wir Rangkorrelationen zwischen den Durchschnittsleistungen in den Ländern und der Zahl der Väter in den unteren und oberen Berufsgruppen in den jeweiligen Ländern errechnet. Sie betrugen für die Pop.II hinsichtlich der unteren Berufsgruppen O-3 r=O.43, für die oberen Berufsgruppen r=1.O. In der Pop.IV betrugen sie hinsichtlich der Berufsgruppen O-3 r=-O.O4, für die Berufsgruppen 7-9 r= O.37 (s.Tab. 14 S.42 unseres Berichts).

Für die Ergebnisse des Englischunterrichts liegt eine Aufschlüsselung der Schüler nach Berufsgruppen der Väter für die einzelnen Länder nicht vor. Wir haben unter der Voraussetzung, daß diese Stichprobe hinsichtlich der sozialen Zusammensetzung der Schülerschaft in etwa der Stichprobe für den naturwissenschaftlichen Unterricht entspricht (s. S.24), eine Korrelation für die Pop.IV ermittelt. Sie beträgt für die unteren Berufsgruppen, die wir hier durch Angehörige der Gruppen 1 und 2 ersetzt haben, r=O.11, bei den oberen Berufsgruppen (3 und 4) r=O.47. Damit wird unsere Annahme, die wir aufgrund der Korrelationen in der genannten Studie S.42 gemacht haben,

bestätigt: Der Einfluß der Sozialstruktur der Elternschaft auf die Schulleistungen wird stärker durch den Anteil der Kinder aus Familien mit gehobenen Berufen als durch die Kinder aus Familien unterer Berufsgruppen bestimmt.

4. Schluß

4.1 Zusammenfassung

Ein Großteil der aus dem internationalen Vergleich gewonnenen Ergebnisse dieser Untersuchung stimmt mit den Ergebnissen der Untersuchung des naturwissenschaftlichen Unterrichts überein. Das Gewicht der außerschulischen, d.h. vor allem der sozio-ökonomischen Faktoren ist zwar nicht ganz so groß wie im naturwissenschaftlichen Unterricht, wirkt sich aber deutlich in den Leistungen aus. Die Bedeutung der schulischen Faktoren für den Erfolg hingegen ist größer als im naturwissenschaftlichen Unterricht. Unter ihnen erwiesen sich als wesentlich der Grad der Selektivität, der Zeitfaktor, d.h. die Zahl der insgesamt für den Englischunterricht aufgewendeten Zeit an Wochenstunden und Schuljahren, bis zu einem gewissen Grade das Ausmaß der Hausaufgaben und die Qualität der Lehrerbildung, die sich auch in der Selbsteinschätzung der Lehrer ausdrückt.

Es erwies sich abermals, daß die Klassengröße nicht in positiver Korrelation mit der Leistung steht, daß auch der Umfang der Gruppenarbeit (man mag darunter verstehen, was immer man will) ohne Einfluß auf die Leistung ist. Für die Bundesrepublik ergab sich darüber hinaus, daß in der Pop.II ein sehr manifester Leistungsunterschied zwischen den drei Schulformen besteht, daß außerdem eine durchgehende Tendenz in beiden Populationen und in allen Schulformen ersichtlich ist, die besagt, daß Schulen mit geringeren Durchschnittsleistungen durchweg größere Streuungen der Schülerleistungen aufweisen als Schulen mit besseren Leistungen. Darüber, ob hochleistungsfähige Schüler in homogenen Klassen besser gefördert werden als in heterogenen Klassen, lassen die Ergebnisse keine Aussage zu.
Die Leistungsunterschiede zwischen den Schulen sind in beiden Populationen auch im Englischunterricht groß. Es besteht die Vermutung, daß dafür nicht allein außerschulische Faktoren wie sie im sozio-ökonomischen Schulindex zum Ausdruck kommen, maßgebend sind, sondern daß sich hier spezifische unterrichtliche Faktoren auswirken. Ihnen nachzugehen, war in einer solchen breit angelegten Untersuchung nicht möglich, weil sich die Untersuchungsinstrumente für solche Feinanalysen als zu grob erwiesen haben.Solchen Unterschieden nachzugehen, wäre in der Tat besonderer Untersuchungen wert. Sie können jedoch nur auf nationaler Ebene durchgeführt wer-

den, da dafür spezifische Untersuchungsinstrumente und Untersuchungsverfahren entwickelt werden müssen.[1] Daß die Erwartung, gut Englisch zu lernen sowie der Grad der Motivation eine große Rolle für den Unterrichtserfolg spielen, dürfte selbstverständlich sein.

4.2 Kritische Anmerkungen

Wenn in der Zusammenfassung bei Lewis/Massad als zentrale Feststellung herauskommt, daß Schulorganisation und Curriculum einen großen Einfluß auf die Leistungen im Englischunterricht haben, so wäre eine so globale Aussage wahrscheinlich auch ohne eine so kostspielige, zeitraubende und umfassende internationale Untersuchung unmittelbar einsichtig gewesen. Erwartet werden durfte, daß darüber hinaus gesicherte Ergebnisse gewonnen würden, die neue Einsicht in das Zusammenwirken der am Unterrichtserfolg beteiligten Faktoren bringen würden und brauchbare Anwendungen ermöglichten. Es muß jedoch zusammenfassend festgestellt werden, daß nur wenige greifbare Antworten auf die einleitend entwickelten Fragestellungen gegeben werden können. Für die Erklärung der Leistungsunterschiede reichen die Ergebnisse zumeist nicht aus, weil zum Teil die Fragestellung infolge der Vielfalt der Situationen in den verschiedenen Ländern nicht durch eine eindeutige Terminologie erfolgt ist, und andererseits die Instrumente nicht fein genug waren, um die innerhalb der einzelnen Nationen bestehenden Unterschiede so zu erfassen, daß die Ursachen dafür sichtbar gemacht werden können.

So wird das mehr als magere Ergebnis der Analyse auch von den Autoren selbst festgestellt, wenn sie immer wieder auf die mangelnde Aussagekraft der Fakten hinweisen und zu vorsichtigen Schlüssen raten. Immer wieder - insbesondere was die schulischen Variablen betrifft - betonen sie, welche uneinheitlichen Tendenzen von Land zu Land und von Population zu Population im gleichen Land auftreten und wie wenig im Grunde Einzelprobleme der Unterrichtsorganisation und des Curriculum durch die Untersuchung beantwortet werden können. Wir könnten viele solcher Stellen aus der Untersuchung anführen, wollen aber nur an einigen wenigen deutlich machen, wo die eigentlichen Mängel der Untersuchung nach der Aussage der Autoren selbst liegen:

---

[1] Daß ein abgeschlossenes Forschungsprojekt, dessen Ergebnisse vor der Veröffentlichung stehen, detailliertere Einsichten in das Zusammenwirken der die Leistung bestimmenden Faktoren erbringt, ist zu hoffen: Frank Achtenhagen und Götz Wienold, Lehren und Lernen im Fremdsprachenunterricht, Bd. 1: Didaktische Analysen und Transformationsgrammatik als Instrumente der Curriculumforschung. Koesel-Verlag 1975.

S.48: "We should not be too ready to stress the power of the conclusions arrived in this study." (Wegen der Möglichkeit, daß die Stichprobe zu klein oder die Instrumente zu unbestimmt (inaccurate) sind.

S.93: "Caution should be exercised in interpreting these correlation-coefficients since the reliability of some measures were not computed and are in any case likely to vary considerably." Dabei wird auf Unregelmäßigkeiten (inconsistensies) sowohl von Land zu Land als zwischen den Populationen in den einzelnen Ländern hingewiesen.

S.97: Im Hinblick auf unzureichende Korrelationen zwischen Prädiktoren und Kriteriumsvariablen: "This may be an effect of the testing instruments, but it may be ...."

"For this reason..., considerable caution should be exercised in interpreting the data. It is an area where more research is required."

S.98/99: Im Hinblick auf die Qualifikation der Lehrer: "Here again because of the degree of inconsistency and the low correlations any interpretation must be tentative."

S.99: In bezug auf die Schul- und Klassenraumvariablen: "Inconsistencies seem to be the trend for all these variables."

S.175: Hinsichtlich des Einflusses des Geschlechts und der Koedukation: "But without more information about how the countries interprete coeducation... it would be premature to draw any definitive conclusion."

S.297/298: In bezug auf den Beginn und die Dauer des Englischunterrichts wird festgestellt, daß die Analyse dieses Berichts keine Unterlagen für solche Entscheidungen liefert, vielmehr sei eine weitere Analyse der Daten, die in der I.E.A.-Datenbank gespeichert sind, notwendig.

S.287: Als Schlußabsatz der Zusammenfassung: "The results reported in this chapter should be looked at as only a beginning in a series of analysis and further research in the areas on the study here. Reports of the National Centers, further analyses of the data collected in the Data-Bank as well as additional research studies will certainly contribute more to providing guidance for the improvement of the educational enterprise."

Das letztere erscheint mir aufgrund der immer wieder aufgezeigten Mängel der Gesamtuntersuchung mehr als zweifel-

haft; es ist nicht einzusehen, warum die Autoren keine effektivere Auswertung vorgenommen haben, wo ihnen doch alle Möglichkeiten einer statistischen Analyse zur Verfügung gestanden haben. Wenn schon international durch ausgeklügelte statistische Methoden , für deren Entwicklung und Anwendung ein Experte während der Gesamtuntersuchung zur Verfügung stand, keine hinreichenden Antworten auf die meisten Fragestellungen gegeben werden konnten, dürfte es kaum möglich sein, durch Analysen bei den nationalen Zentren zu weitergehenden allgemeinen, d.h. international aufschlußreichen Ergebnissen und Folgerungen zu kommen.

Schließlich reicht eine solche Stichprobe, die international stimmig und groß genug sein mag, für detaillierte nationale Analysen nicht mehr aus. Eine Heranziehung aller Schulen und weiterer Schulformen aus dem nicht allgemeinbildenden Schulwesen, wie es z.B. in Italien und Schweden geschehen ist, ist notwendig, wenn auf nationaler Ebene wirklich konkrete Aussagen über die Relation der unterschiedlichsten Prädiktoren mit den Leistungen im Englischunterricht gemacht werden sollen. Gegen den Anspruch einer solchen internationalen Erhebung, maßgebliche Ergebnisse für die Entwicklung der nationalen Curricula bereitstellen zu können, spricht auch die zu lange Dauer der Untersuchung(bei der Englischuntersuchung sind vom Beginn der Vorüberlegungen bis zum Erscheinen des Buches etwa neun Jahre vergangen).
Es ist erstaunlich, daß im Gegensatz zu den Mitarbeitern der Untersuchung zum Englischunterricht, die an der Studie über den Französischunterricht[1] beteiligten Mitarbeiter glauben, aufgrund ihrer Ergebnisse sehr viel positivere Aussagen machen zu können. Wir geben infolgedessen im Anhang eine kurze Übersicht über die Ergebnisse dieser Studie und die Schlußfolgerungen, die John B.Carroll daraus zieht.

5. Anhang: Die wichtigsten Ergebnisse der Studie über den Französischunterricht in acht Ländern[2]

(1.) International: a) Als wichtigste Variable erwies sich der Zeitfaktor.

b) Die Art der Muttersprache erwies sich nicht als wesentlicher Faktor.

---

[1] Diese Untersuchung lief ebenfalls in der Phase 2 der I.E.A.-Untersuchungen.

[2] John B. Carroll, The Teaching of French as a Foreign Language in Eight Countries, Stockholm 1975.

c) Das Interesse korrelierte nicht in dem Maße gleichsinnig wie in der Englischstudie. Doch glaubt Carroll, daß dieser Faktor immerhin für einen gewissen Prozentsatz der Varianz maßgeblich sein könnte.

d) Andere Variablen erwiesen sich sowohl zwischen den Ländern wie in den einzelnen Ländern als weitgehend einheitlich in ihrer positiven Tendenz: Die Selbsteinschätzung des Lehrers bezüglich seiner Sprachkompetenz, der Gebrauch der französischen Sprache im Unterricht, das Ausmaß der Hausarbeiten sowie die Einschätzung des Schwierigkeitsgrades des Französischunterrichts durch die Schüler.

(2.) Innerhalb der Länder:[1)]
Der Block I (außerschulische Faktoren) erwies sich als signifikant nur in einigen Ländern, im wesentlichen durch den sozio-ökonomischen Schulindex. Das Geschlecht war selten von erheblichem Gewicht.
Der Block II (Schultypen und -formen) erwies sich ebenfalls als signifikant nur in bestimmten Ländern und dann auch nur bei der Pop.II. Da in einigen Ländern keine Differenzierung der Schüler auf dieser Stufe stattfindet und in anderen Französisch nur in akademischen Zügen unterrichtet wird, liegt dies Ergebnis nahe.
Block III (die Zeitfaktoren) erwies sich von massiver Wirkung. Selbstverständlich war dieser Beitrag zur Erklärung der Varianz in den Ländern am größten, in denen die Streuung der für den Unterricht aufgewandten Zeit am größten war. Wenn auch der Zeitfaktor sich immer wieder als durchschlagend erwies, so konnte keine Klarheit darüber gewonnen werden, ob es einen Unterschied macht, mit welcher Klasse bzw. in welchem Alter der Französischunterricht beginnt. (Es besteht für Carroll eine gewisse Wahrscheinlichkeit, daß die übliche Annahme, der Frühbe-

---

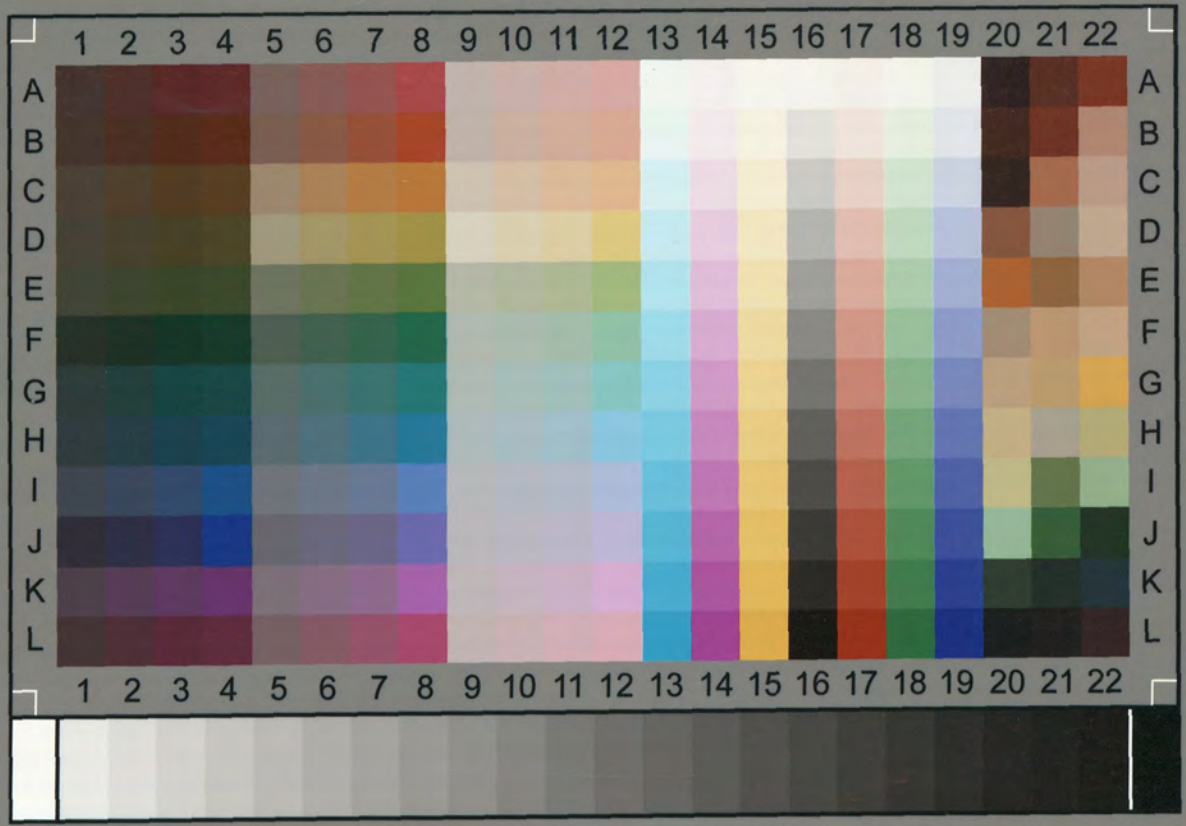
[1)] Die Analyse bediente sich ebenfalls der Regressionsanalyse, doch weist Carroll auf Nachteile dieser Analyse hin. Die Ergebnisse werden beeinflußt:
(a) durch den Grad der Varianz in den Prädiktoren und Kriterien-Variablen,
(b) durch den Grad des Zusammenhangs der Prädiktoren und der Kriterienvariablen und schließlich
(c) durch Stichprobenfehler sowie durch die entstehenden Schwierigkeiten, wenn Variable zu Blocks zusammengefaßt werden.

ginn des Unterrichts sei von Vorteil, nicht stimmt, ohne daß er diese Feststellung belegt).
Der Block IV (unterrichtliche Faktoren) erwies sich als wichtige Faktorengruppe in den meisten Ländern und Populationen, vor allem da, wo die Lehrerkompetenz große Streuung zeigte. Folgende Lehrervariablen stehen eindeutig in positiver Relation zur Leistung ihrer Schüler: Das Geschlecht, die Dauer der Universitätsausbildung, die Lehrerfahrung, vor allem jedoch die Selbsteinschätzung ihrer Kompetenz. Das Ausmaß des formalen Grammatikunterrichts zeigte keine eindeutige Tendenz; doch glaubt er feststellen zu können, daß ein gewisser Hinweis dafür (a fairly good indication) vorhanden ist, daß der Gebrauch der Muttersprache im Unterricht einen negativen Prädiktor darstellt.
Der Block V (Neigung und Interesse) hatte ein relativ hohes Gewicht in den meisten Ländern. Der Grad, in dem der Schüler Französisch als leicht oder schwer empfindet, erwies sich als einer der stärksten und signifikantesten Prädiktoren; wahrscheinlich, weil in dieser Variablen die sprachliche Begabung des Schülers durchschlägt.
Der Block VI schließlich, der den Wortschatztest einbezog, ergab einen zwar relativ kleinen, aber doch signifikanten Prozentsatz in allen Ländern und in allen Populationen, so daß er sich als guter Prädiktor vor allem der Leseleistung erweist.

Seine Ergebnisse wertet Carroll hinsichtlich der Organisation und der Durchführung des Französischunterrichts folgendermaßen aus:
Das Ausmaß der Zeit ist von größtem Einfluß. Wenn die Schüler der Pop.IV die international besten Leistungen erreichen sollen, muß der Unterricht wenigstens sechs Jahre dauern, muß also spätestens in Klasse 6 beginnen, wobei drei bis fünf Wochenstunden vorausgesetzt werden. Da jedoch auch andere Variablen von Einfluß auf die Leistung sind, kann das zeitliche Ausmaß variieren. Wollen Schüler schneller vorankommen, so gehören dazu bestmögliche Unterrichtsbedingungen, mehr als durchschnittliche Sprachbegabung und hohe Anforderungen des Schülers an die eigene Leistung, eine hohe Kompetenz des Lehrers, und es sollte im Unterricht möglichst viel Französisch gesprochen und die Muttersprache möglichst wenig verwendet werden. Das Ausmaß der Hausarbeit ist bedeutend für die Lesefertigkeit, weniger für das Hörverstehen.

Was das Curriculum anbelangt, so sollte der Unterricht auf unterschiedliche Lernziele gerichtet sein. Lesefähigkeit, Hörverstehen oder aber auch schriftlicher Ausdruck erfordern unterschiedliche Zielsetzungen.

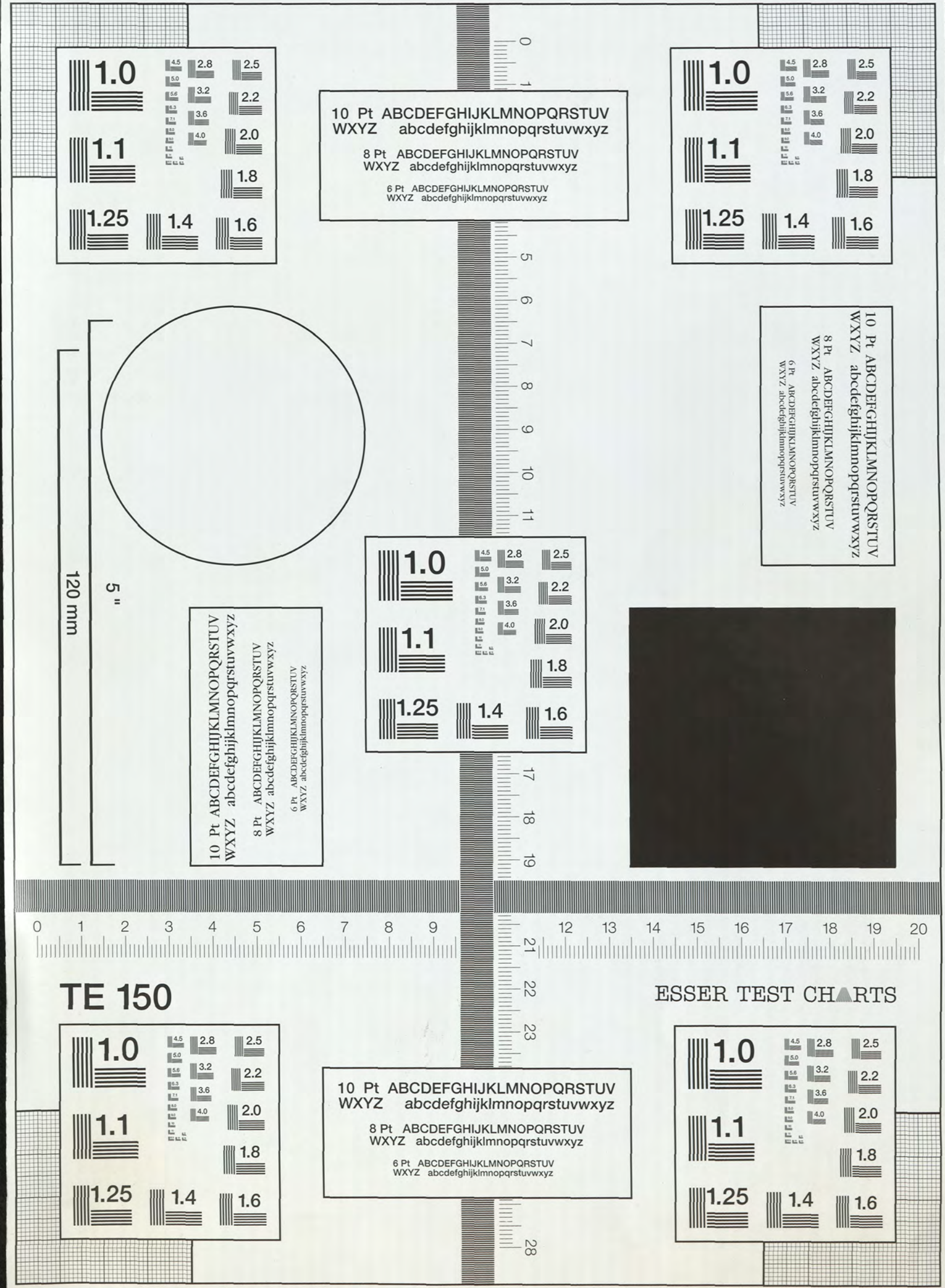
Was diese Darstellung von der Studie über den Englischunterricht unterscheidet, ist vor allem ihre durchsichtigere Anlage, die eindeutigere Ausdeutung ihrer Ergebnisse und der größere Mut des Autors, unterrichtliche Konsequenzen zu ziehen. Inhaltlich stimmen die Ergebnisse beider Studien bis auf Einzelheiten - wie z.B. der Gebrauch der Muttersprache im Unterricht - weitgehend überein.



IT8.7/2-1993
2013:02

Printed on Kodak Professional Paper - Made by Wolf Faust (www.coloraid.de)

Charge: R130227


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
4.5
5.0
5.6
6.3
7.1
8.0
9.0
10 Pt ABCDEFGHIJKLMNOPQRSTUV
WXYZ abcdefghijklmnopqrstuvwxyz
8 Pt ABCDEFGHIJKLMNOPQRSTUV
WXYZ abcdefghijklmnopqrstuvwxyz
6 Pt ABCDEFGHIJKLMNOPQRSTUV
WXYZ abcdefghijklmnopqrstuvwxyz
120 mm
5 "
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 28
TE 150
ESSER TEST CHARTS