DOCUMENT RESUME

ED 463 576 EA 031 575

AUTHOR Anderson, Lowell Bruce; Bracken, Jerome; Bracken, Marilyn C.
TITLE Review of Department of Defense Education Activity (DoDEA) Schools. Volume II: Quantitative Analysis of Educational Quality. IDA Paper.
INSTITUTION Institute for Defense Analyses, Alexandria, VA.
SPONS AGENCY Department of Defense, Washington, DC.
REPORT NO IDA-P-3544
PUB DATE 2000-10-00
NOTE 94p.; For volume I, see EA 031 574.
CONTRACT DASW01-98-C-0067
PUB TYPE Information Analyses (070) -- Numerical/Quantitative Data (110)
EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Educational Quality; Elementary Secondary Education; *Evaluation; Foreign Countries; Military Personnel
IDENTIFIERS *Dependents Schools; Europe

ABSTRACT

This volume compiles, and presents in integrated form, the Institute for Defense Analyses' (IDA) quantitative analysis of educational quality provided by the Department of Defense's dependent schools. It covers the quantitative aspects of volume 1 in greater detail and presents some analyses deemed too technical for that volume. The first task in conducting the study involved reviewing relevant literature to determine the most important factors on which data should be obtained and analysis performed. The second task was to become familiar with the data. The third task was to integrate the data and conduct the analysis. After the introduction (chapter 1), the following topics are covered in their own chapters: student achievement, educational resources, college attendance, teacher quality, measuring the Department of Defense Educational Activities (DoDEA) contribution to student achievement, and conclusions and recommendations. Recommendations include continuing to focus on and deliver good student achievement as evidenced in this study, concentrating on maintaining DoDEA's high level of intellectual capability of its teachers when recruiting and managing them, and introducing a system to measure the contribution of DoDEA to student achievement. The report concludes with 30 tables. (RT)

IDA

INSTITUTE FOR DEFENSE ANALYSES

Review of Department of Defense Education Activity (DODEA) Schools
Volume II: Quantitative Analysis of Educational Quality

Lowell Bruce Anderson
Jerome Bracken
Marilyn C. Bracken

October 2000

Approved for public release; distribution unlimited.

IDA Paper P-3544

Log: H 00-002449

This work was conducted under contract DASW01 98 C 0067, Task BE-6-1786, for the Under Secretary of Defense (Personnel and Readiness). The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-3544

Review of Department of Defense Education Activity (DODEA) Schools

Volume II: Quantitative Analysis of Educational Quality

Lowell Bruce Anderson
Jerome Bracken
Marilyn C. Bracken

CONTENTS

Tables

I. INTRODUCTION

This volume compiles, and presents in integrated form, IDA's quantitative analysis of educational quality provided by DoD's dependent schools. It covers the quantitative aspects of volume I in greater detail, and it presents some analyses deemed too technical for that volume. Topics addressed in this volume include student achievement, measuring the DoDEA contribution to student achievement, college attendance, teacher quality, and resources.

A. APPROACH

The first task of the approach to the quantitative analysis of educational quality was to read the literature to determine the most important factors on which data should be obtained and analysis performed.

The most enlightening sources found on how to approach the problem were as follows:

1. Eric A. Hanushek and Others, *Making Schools Work, Improving Performance and Controlling Costs*, The Brookings Institution, 1994.
2. Gary Burtless (Editor), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, The Brookings Institution, 1996.
3. Marci Kastoroom and Chester E. Finn, Jr. (Editors), *Better Teachers, Better Schools*, Thomas B. Fordham Foundation, 1999.
4. Dale Ballou and Michael Podgursky, *Teacher Pay and Teacher Quality*, W. E. Upjohn Institute, 1997.
5. Education Week, *Quality Counts 2000*, January 13, 2000.
6. James W. Pellegrino, Lee R. Jones, and Karen J. Mitchell (Editors), *Grading the Nation's Report Card*, National Research Council, National Academy Press, 1999.
7. Michael J. Feuer, Paul W. Holland, Bert F. Green, Merul W. Bertenthal, and F. Cadelle Hempjill (Editors), *Uncommon Measures: Equivalence and Linkage Among Educational Tests,* National Research Council, National Academy Press, 1999.

The first two sources deal with the relationships between student achievement and educational resources. The third and fourth sources focus on teachers. The fifth source postulates a wide variety of measures of educational quality. The sixth and seventh sources address issues of performance testing. In addition, a wide variety of books and journal articles, both general and specific, were consulted.

Interviews were conducted with a number of experts from the U.S. Department of Education, state Departments of Education, county Departments of Education, universities, and educational testing organizations.

The second task was to become familiar with the available data. The best source of data found was the National Center for Educational Statistics. In particular, the National Assessment of Educational Progress (NAEP) provided data by state (including DoDDS and DDESS as states) in a form such that DoDDS and DDESS could be compared with states with respect to results in tests for reading, writing, mathematics and science, and in educational inputs such as class size and teacher qualifications.

Data at a level lower than states, however, are also important. Comparing DoDDS and DDESS to school districts with similar characteristics provides insights not available at the state level, both on student performance and on resources.

DoDDS and DDESS administer the CTBS/Terra Nova test, which is also administered by a number of states. Some states report data by school district. The best data source found was Maryland, which has 23 districts. DoDDS and DDESS could thus be compared with all 23 districts. Several other states also have good comparable CTBS/Terra Nova data. We learned, however (from Dr. Steven Gorman of the National Center for Educational Statistics), of the "Lake Wobegon Effect"[1] discovered by John Jacob Cannell and subsequently investigated by many researchers. In a study of nationally-normed tests he found that all states were above average. The norms developed by the contractors insured that the test results for all of their customers would be above average. Thus the CTBS/Terra Nova test results need to be approached with some caution. They can be very useful for comparing one jurisdiction with another but should be viewed with some skepticism with respect to the absolute results.

Data were also available from the Washington, DC, metropolitan area Boards of Education in a document produced by the Fairfax, Virginia, County School System. Nine school systems that have large enrollments of children from military families provided a variety of data elements that were useful in the analysis.

The third task was to integrate the data and conduct the analysis.

---

1 John Jacob Cannell, "Nationally Normed Elementary Achievement Testing in America's Public Schools: How All 50 States Are Above the National Average," *Educational Measurement: Issues and Practice*, Volume 7, Number 2, 1988.

B. EDUCATIONAL RESOURCES AND STUDENT ACHIEVEMENT

In the first and second references in the previous section, Hanushek traces the history of educational inputs and outputs in the United States. Since the 1950s, test scores have not increased, while expenditures per student have increased greatly. Associated with increased expenditures per student, class sizes have decreased and teacher salaries have increased.

There is a long-standing disagreement among economists on the relationships among test scores, expenditures per student, and other input resources. This argument is elaborated on in the second reference above and in many articles in prominent academic journals. There seems to be almost unanimous agreement, however, that student achievement is related to teachers.

A large majority of teachers are women. In the last 20 years, the wages of women employees in teaching relative to the wages of women employees in the general economy have decreased from the well above average to well below average. Thus, if capability is related to compensation, the women teachers of today are from a less competitive pool of people.

Identification and investigation of the major educational inputs and outputs of DoDEA compared with those of other jurisdictions is the essence of the study.

C. SCOPE OF ANALYSIS

Chapter II, Student Achievement, is the most detailed inquiry. It compares DoDDS and DDESS with states for the NAEP in reading, writing, mathematics, and science, and it examines the performance of minorities, lower-scoring students, and higher-scoring students. It compares DoDDS and DDESS with national norms for the CTBS/Terra Nova in reading, language, mathematics, science, and social studies. While 23 states use some form of the CTBS/Terra Nova, data were available and analyzed by school district for three states. The analysis compares results on the CTBS/Terra Nova for several school districts near military installations and examines test results and comparative ethnicity data for Montgomery County, Maryland.

DoDDS and DDESS results on the SAT are compared with states. Included in this comparison is the application of several methods for correcting for participation rates. SAT results are compared with several school districts in the Washington metropolitan area. There are some inconsistencies between the NAEP and CTBS/Terra Nova results, which are relatively high, and the SAT results, which are average for DoDDS and below average for DDESS. These inconsistencies are analyzed in two sections.

Chapter III, Resources, contains comparative data by state from the National Center for Educational Statistics Common Core of Data on expenditures per student. It contains comparative data by state from the NAEP on class size and on teacher qualifications. It also contains comparative data for the Washington metropolitan area for a wide variety of educational resource inputs including expenditures per student and teacher salaries and benefits.

Chapter IV, College Attendance, examines the quality of the colleges attended by DoDEA graduates. The first analysis addresses college attendance at the high end—of those DoDEA graduates going on for higher education, what portion attend the nation's top universities and colleges? The second analysis addresses college attendance by all DoDEA students—how well are DoDEA graduates represented at all universities and colleges? In the first inquiry, all of the students attending top universities and colleges are considered. In the second inquiry, a sampling approach is used.

Chapter V, Teacher Quality, examines the quality of teachers in DoDEA compared with the quality of teachers in public schools and private schools. Research has shown that student achievement is strongly related to the intellectual capability of teachers. The chapter compares the selectivity of the undergraduate schools of public school teachers, private school teachers, and DoDEA teachers. The chapter also compares the SAT and ACT scores of the undergraduate schools attended by DoDEA teachers to those of the population in general and to the scores of traditionally teachers colleges in Pennsylvania.

Chapter VI, Measuring the DoDEA Contribution to Student Achievement, addresses the extremely important question of whether the performance of DoDEA students is worse than expected, about as expected, or greater than expected given the quality of the students. If these same students were in the public schools of the United States in general, would they score higher than average, as DoDEA compares to states on the NAEP? Or, if these students were in Montgomery County, Maryland, would they score lower than average, as DoDEA compares to Montgomery County on the CTBS/Terra Nova and the SAT? We attempted to run regressions with data from Maryland school districts to estimate whether children from military families tended to increase or decrease the CTBS/Terra Nova scores of the districts in which they reside but were unsuccessful in reaching robust results. We also tried to obtain and analyze similar data from other jurisdictions. A methodology developed by the IDA study team is presented which would provide DoDEA with a straightforward procedure for determining the impact of DoDEA on the student achievement.

Chapter VII provides conclusions and recommendations resulting from the analysis.

II. STUDENT ACHIEVEMENT

Three measures of student achievement are considered in this chapter. They are the National Assessment of Educational Progress (NAEP), the Comprehensive Test of Basic Skills (CTBS), and the Scholastic Aptitude Test (SAT).

A. NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP)

1. Background

The best measure for evaluating the student achievement of DoDDS and DDESS is the National Assessment of Educational Progress (NAEP). It is the nation's only ongoing survey of what students know and can do in various academic subjects. Authorized by Congress and administered by the National Center for Educational Statistics in the Department of Education, the NAEP regularly reports to the public on the educational progress of students in grades 4, 8 and 12.

In 1998, NAEP conducted a state-by-state reading assessment of fourth-grade and eighth-grade students and a state-by-state writing assessment of eighth-grade students.

In 1996, NAEP conducted a state-by-state mathematics assessment of fourth-grade students and eighth-grade students and a state-by-state science assessment of eighth-grade students.

The NAEP does not test each student. Rather, it uses statistical sampling techniques that are designed to result in an accurate score for each state. The District of Columbia, DoDDS and DDESS are treated the same as states.

All states do not always participate in the NAEP. For the examinations mentioned above, there were different numbers of participating entities.

The methodology of the NAEP has been reviewed by independent outside experts. In particular, see the generally positive report by the National Research Council, *Grading the Nation's Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress*, National Academy Press, 1999.

2. NAEP Scores

Tables II-1 through II-6, below, give NAEP scores by state.[1] Enrollments for 1999 are also given to allow the reader to compare the sizes of the states.[2] The District of Columbia, DoDDS and DDESS are considered as states. The data in each table are sorted from the highest-scoring state to the lowest-scoring state.

Only the District of Columbia and Wyoming, in addition to DoDDS and DDESS, have fewer than 100,000 students. Nine states have between 100,000 and 200,000 students: Alaska, Delaware, Hawaii, Montana, New Hampshire, North Dakota, Rhode Island, South Dakota, and Vermont. Thirteen states have a million or more students. In comparing DoDDS and DDESS with states, one should remember that the comparisons may be somewhat limited by the sizes and the heterogeneity of the entities.

Some statistics concerning these tables are given below, where the term "States Participating" includes the states that chose to participate and the District of Columbia, which always participated, but does not include DoDDS and DDESS. Note that the reading and writing performance of DoDDS and DDESS ranks high among states, particularly at the Grade 8 level. Mathematics performance is about average. Science performance is above average.

	States Participating	States Higher than DoDDS	States Higher than DDESS
Grade 4 Reading	40	6	10
Grade 8 Reading	36	3	3
Grade 8 Writing	36	1	1
Grade 4 Mathematics	44	19	22
Grade 8 Mathematics	41	17	24
Grade 8 Science	41	11	14

1 Sources: NAEP Reading: Report Card for the Nation and the States, National Center for Educational Statistics, March 1999; NAEP 1998 Writing: Report Card for the Nation and the States, National Center for Educational Statistics, September 1999; NAEP 1994 Mathematics: Report Card for the Nation and the States, National Center for Educational Statistics, February 1997; NAEP 1994 Science: Report Card for the Nation and the States, National Center for Educational Statistics, May 1997.

2 Source: Education Week, Quality Counts 2000, January 13, 2000. Data from "Early Estimates of Public Elementary and Secondary Education Statistics: School Year 1998-1999," National Center for Educational Statistics, January 1999

Table II-1. NAEP Grade 4 Reading (1998)

State	1999 Enrollment (000)	NAEP Score
Connecticut	545	232
Montana	161	226
New Hampshire	195	226
Maine	220	225
Massachusetts	964	225
Wisconsin	888	224
Iowa	503	223
DoDDS	**76**	**223**
Colorado	699	222
Kansas	470	222
Minnesota	858	222
Oklahoma	627	220
DDESS	**36**	**220**
Wyoming	94	219
Kentucky	646	218
Rhode Island	154	218
Virginia	1,100	218
Michigan	1,700	217
North Carolina	1,200	217
Texas	3,900	217
Washington	1,000	217
Missouri	921	216
New York	2,900	216
West Virginia	296	216
Maryland	837	215
Utah	447	215
Nation	**46,100**	**215**
Oregon	543	214
Delaware	113	212
Tennessee	909	212
Alabama	759	211
Georgia	1,400	210
South Carolina	644	210
Arkansas	456	209
Nevada	311	208
Arizona	829	207
Florida	2,300	207
New Mexico	329	206
Louisiana	754	204
Mississippi	502	204
California	5,800	202
Hawaii	187	200
District of Columbia	80	182

Table II-2. NAEP Grade 8 Reading (1998)

State	1999 Enrollment (000)	NAEP Score
Maine	220	273
Connecticut	545	272
Montana	161	270
Massachusetts	964	269
DoDDS	**76**	**269**
DDESS	**36**	**269**
Kansas	470	268
Minnesota	858	267
New York	2,900	266
Oregon	543	266
Virginia	1,100	266
Wisconsin	888	266
Oklahoma	627	265
Utah	447	265
Washington	1,000	265
Colorado	699	264
North Carolina	1,200	264
Missouri	921	263
Kentucky	646	262
Maryland	837	262
Rhode Island	154	262
Texas	3,900	262
West Virginia	296	262
Wyoming	94	262
Arizona	829	261
Nation	**46,100**	**261**
Tennessee	909	259
New Mexico	329	258
Georgia	1,400	257
Nevada	311	257
Arkansas	456	256
Delaware	113	256
Alabama	759	255
South Carolina	644	255
California	5,800	253
Florida	2,300	253
Louisiana	754	252
Mississippi	502	251
Hawaii	187	250
District of Columbia	80	236

Table II-3. NAEP Grade 8 Writing (1998)

State	1999 Enrollment (000)	NAEP Score
Connecticut	545	165
DDESS	**36**	**160**
DoDDS	**76**	**156**
Maine	220	155
Massachusetts	964	155
Texas	3,900	154
Virginia	1,100	153
Wisconsin	888	153
Oklahoma	627	152
Colorado	699	151
Montana	161	150
North Carolina	1,200	150
Oregon	543	149
Minnesota	858	148
Rhode Island	154	148
Tennessee	909	148
Washington	1,000	148
Nation	**46,100**	**148**
Maryland	837	147
Georgia	1,400	146
Kentucky	646	146
New York	2,900	146
Wyoming	94	146
Alabama	759	144
Delaware	113	144
West Virginia	296	144
Arizona	829	143
Utah	447	143
Florida	2,300	142
Missouri	921	142
California	5,800	141
New Mexico	329	141
Nevada	311	140
South Carolina	644	140
Arkansas	456	137
Louisiana	754	136
Hawaii	187	135
Mississippi	502	134
District of Columbia	80	126

Table II-4. NAEP Grade 4 Mathematics (1996)

State	1999 Enrollment (000)	NAEP Score
Connecticut	545	232
Maine	220	232
Minnesota	858	232
North Dakota	114	231
Wisconsin	888	231
Indiana	989	229
Iowa	503	229
Massachusetts	964	229
Texas	3,900	229
Montana	161	228
Nebraska	291	228
New Jersey	1,300	227
Utah	447	227
Colorado	699	226
Michigan	1,700	226
Pennsylvania	1,800	226
Missouri	921	225
Vermont	105	225
Washington	1,000	225
Alaska	134	224
North Carolina	1,200	224
Oregon	543	224
DoDDS	**76**	**224**
New York	2,900	223
Virginia	1,100	223
West Virginia	296	223
Wyoming	94	223
DDESS	**36**	**223**
Nation	**46,100**	**222**
Maryland	837	221
Kentucky	646	220
Rhode Island	154	220
Tennessee	909	219
Arizona	829	218
Nevada	311	218
Arkansas	456	216
Florida	2,300	216
Delaware	113	215
Georgia	1,400	215
Hawaii	187	215
New Mexico	329	214
South Carolina	644	213
Alabama	759	212
California	5,800	209
Louisiana	754	209
Mississippi	502	208
District of Columbia	80	187

Table II-5. NAEP Grade 8 Mathematics (1996)

State	1999 Enrollment (000)	NAEP Score
Iowa	503	284
Maine	220	284
Minnesota	858	284
North Dakota	114	284
Montana	161	283
Nebraska	291	283
Wisconsin	888	283
Connecticut	545	280
Vermont	105	279
Alaska	134	278
Massachusetts	964	278
Michigan	1,700	277
Utah	447	277
Colorado	699	276
Indiana	989	276
Oregon	543	276
Washington	1,000	276
Wyoming	94	275
DoDDS	**76**	**275**
Missouri	921	273
Nation	**46,100**	**271**
Maryland	837	270
New York	2,900	270
Texas	3,900	270
Virginia	1,100	270
Rhode Island	154	269
DDESS	**36**	**269**
Arizona	829	268
North Carolina	1,200	268
Delaware	113	267
Kentucky	646	267
West Virginia	296	265
Florida	2,300	264
California	5,800	263
Tennessee	909	263
Arkansas	456	262
Georgia	1,400	262
Hawaii	187	262
New Mexico	329	262
South Carolina	644	261
Alabama	759	257
Louisiana	754	252
Mississippi	502	250
District of Columbia	80	233

Table II-6. NAEP Grade 8 Science (1996)

State	1999 Enrollment (000)	NAEP Score
Maine	220	163
Montana	161	162
North Dakota	114	162
Wisconsin	888	160
Minnesota	858	159
Iowa	503	158
Wyoming	94	158
Massachusetts	964	157
Nebraska	291	157
Vermont	105	157
Utah	447	156
Colorado	699	155
Connecticut	545	155
Oregon	543	155
DoDDS	**76**	**155**
Alaska	134	153
Indiana	989	153
Michigan	1,700	153
DDESS	**36**	**153**
Missouri	921	151
Washington	1,000	150
Virginia	1,100	149
Nation	**46,100**	**148**
Kentucky	646	147
North Carolina	1,200	147
Rhode Island	154	147
West Virginia	296	147
New York	2,900	146
Arizona	829	145
Maryland	837	145
Texas	3,900	145
Arkansas	456	144
Tennessee	909	143
Delaware	113	142
Florida	2,300	142
Georgia	1,400	142
New Mexico	329	141
Alabama	759	139
South Carolina	644	139
California	5,800	138
Hawaii	187	135
Mississippi	502	133
Louisiana	754	132
District of Columbia	80	113

Average test scores comparing DoDDS and DDESS with the nation are as follows:

	DoDDS	DDESS	Nation
Grade 4 Reading	223	220	215
Grade 8 Reading	269	269	261
Grade 8 Writing	156	160	148
Grade 4 Mathematics	224	223	222
Grade 8 Mathematics	275	269	270
Grade 8 Science	155	153	148

In reading and writing, DoDDS and DDESS scores are significantly higher than the national average. In Grade 4 Mathematics, DoDDS and DDESS are slightly higher. In Grade 8 Mathematics, DoDDS is higher and DDESS is lower. In Grade 8 Science, DoDDS and DDESS are higher.

3. Performance of Minorities

Data for the performance of minority students on the NAEP are available from the National Center for Educational Statistics.[3]

The following data show the comparative average performance of African-American students in DoDDS, DDESS and the nation on the 1998 tests.

1998 Test	DoDDS	DDESS	Nation
Grade 4 Reading	212	209	193
Grade 8 Reading	259	253	241
Grade 8 Writing	148	150	130

The following data show the comparative average performance of Hispanic students in DoDDS, DDESS and the nation on the 1998 tests.

1998 Test	DoDDS	DDESS	Nation
Grade 4 Reading	216	211	195
Grade 8 Reading	263	268	243
Grade 8 Writing	153	153	129

Average achievement of minority students in DoDDS and DDESS is significantly better than that of minority students in the nation.

---

3 Source: National Center for Educational Statistics web site www.nces.ed.gov.

4. Performance of Lower-Scoring Students

Data for the performance of students by percentile on the NAEP are available from the National Center for Educational statistics.[4]

The following data show the average performance of students in the 10^{th} percentile of DoDDS, the 10^{th} percentile of DDESS and the 10^{th} percentile of the nation.

1998 Test	DoDDS	DDESS	Nation
Grade 4 Reading	181	173	160
Grade 8 Reading	228	224	215
Grade 8 Writing	113	108	102

The following data show the average performance of students in the 25^{th} percentile of DoDDS, the 25^{th} percentile of DDESS and the 25^{th} percentile of the nation.

1998 Test	DoDDS	DDESS	Nation
Grade 4 Reading	203	197	192
Grade 8 Reading	249	246	239
Grade 8 Writing	135	131	124

These data demonstrate that, on average, the lower-scoring students in DoDDS and DDESS outperform the lower-scoring students in the nation.

5. Performance of Higher-Scoring Students

The following data show the average performance of students in the 90^{th} percentile of DoDDS, the 90^{th} percentile of DDESS and the 90^{th} percentile of the nation.

1998 Test	DoDDS	DDESS	Nation
Grade 4 Reading	265	265	261
Grade 8 Reading	308	313	304
Grade 8 Writing	199	212	192

4 Source: National Center for Educational Statistics web site www.nces.ed.gov .

The following data show the average performance of students in the 75th percentile of DoDDS, the 75th percentile of DDESS and the 75th percentile of the nation.

1998 Test	DoDDS	DDESS	Nation
Grade 4 Reading	246	245	242
Grade 8 Reading	290	292	286
Grade 8 Writing	179	188	172

These data demonstrate that, on average, the higher-scoring students in DoDDS and DDESS outperform the higher-scoring students in the nation.

The relative performance advantage of DoDDS and DDESS is not as great with higher-scoring students as with lower-scoring students. Since the NAEP is a zero-based test, it is permissible to make comparisons based on ratios of scores. Lower-scoring DoDDS and DDESS students score on the average seven percent better than the nation while higher-scoring DoDDS and DDESS students score on the average three percent better than the nation.

It is also interesting to note that, with respect to lower-scoring students, DoDDS usually scores higher than DDESS, while with respect to higher-scoring students, DDESS usually scores higher than DoDDS.

6. Conclusions

On a national basis, considering the best comparable data available, DoDDS and DDESS perform better than most states in Grade 4 reading and better than almost all states in Grade 8 reading and writing. DoDDS and DDESS perform about average in mathematics and better than average in science.

DoDDS and DDESS minority students perform much better than minority students in the nation.

DoDDS and DDESS lowest-scoring students perform better than lowest-scoring students in the nation.

DoDDS and DDESS highest-scoring students perform slightly better than highest-scoring students in the nation.

B. CTBS/TERRA NOVA

1. Introduction

DoDDS and DDESS administer the CTBS/Terra Nova test to all students in grades 3 through 11. There are five subjects: reading, language arts, mathematics, science and social studies.

The CTBS/Terra Nova is a nationally norm-referenced test. A norm-referenced test allows a comparison of student performance against a nationally representative sample of students (a norm group.) A national percentile score of 50 is equivalent to performance at the national median. A national percentile score of 65 is equivalent to performance at the level of the top 35 percent of the nation.

This assumes that the national norm is correct. A well-known study by John Jacob Cannell proposed the "Lake Wobegon Effect."[5] In a study of all of the states using a nationally-normed test, none scored below average. There are many subsequent documents and papers addressing this important issue. The hypothesis of the Cannell study is that the testing contractor developed a national norm that was too low and hence all of the contractor's customers performed at least at the average level. Consequently, in evaluating performance against such a national norm one must retain some skepticism that a score of 50 may not really be the score of the average child in the nation—the score may be inflated due to the norm group being lower-qualified than the average child. In examining the available data for a wide variety of states, we find many examples of school districts with scores lower than 50, and some states with scores near 50, but we cannot be sure that the average score across all states would be 50.

Table II-7 presents DoDDS and DDESS scores for 1999.[6] To facilitate summary comparisons, an average is also given. DoDDS consistently scores in the middle to high 60s and DDESS consistently scores in the low to middle 60s on the CTBS/Terra Nova.

---

5 John Jacob Cannell, "Nationally Normed Elementary Achievement Testing in America's Public Schools: How All 50 States Are Above the National Average," Educational Measurement: Issues and Practice, Volume 7, Number 2, 1988.

6 Source: DoDEA web site www.odedodea.edu.

Table II-7. DoDDS and DDESS 1999 CTBS Scores

Grade	Population	Read	Lang	Math	Science	Soc Stud	Average
DoDDS							
3	6733	58	61	58	57	52	57
4	6242	68	66	64	65	67	66
5	5953	69	65	63	67	66	66
6	5483	65	69	65	69	63	66
7	4966	69	65	62	63	64	65
8	4676	65	69	64	67	66	66
9	4180	71	71	66	62	70	68
10	3541	72	72	70	68	74	71
11	2989	71	69	69	68	74	70
DDESS							
3	3078	62	62	63	64	54	61
4	2851	66	63	64	67	63	65
5	2539	66	63	62	66	63	64
6	2344	62	66	64	68	61	64
7	1642	65	61	58	59	59	60
8	1376	62	66	59	67	61	63
9	1031	67	67	59	60	63	63
10	714	68	66	64	64	68	66
11	560	67	64	61	63	70	65

Some variant of the CTBS is administered in 22 states. Many of the states do not use the same test. Many of the states have hundreds of school districts, and in many cases data are not available at the state level.

Here we analyze data developed for three states by school district. Complete data are available for Maryland (except for Calvert County) and for Nevada. Data are included for all school districts with more than 5,000 students for New Mexico.

Assuming that the tests are identical, this allows a direct comparison of DoDDS and DDESS with school districts in these three states. Some of the school districts have more students than DoDDS and DDESS.

2. Results

a. Data

The enrollments of Maryland, New Mexico and Nevada are 842,000, 325,000 and 297,000, respectively. Test data are available for 1999, 1998 and 1999, respectively.[7]

Tables II-8, II-9 and II-10 give data on the school districts in these states in a format that allows comparisons with DoDDS and DDESS.

In the Maryland comparison, school district, DoDDS and DDESS 1999 data are available for Grades 4 and 8 for the three subjects of Reading, Language Arts and Mathematics.

In the New Mexico comparison, school district, DoDDS and DDESS 1998 data are available for Grades 4, 6 and 8 for the five subjects of Reading, Language Arts, Mathematics, Science and Social Studies.

In the Nevada comparison, school district, DoDDS and DDESS 1999 data are available for Grades 4, 8 and 10 for the three subjects of Reading, Language Arts and Science.

To facilitate comparisons, averages are taken of the available data.

b. Interpretation of the Data

In Maryland, two districts score higher than DoDDS and five districts score higher than DDESS. Eight districts score in the 60s or better. Five of the eight districts have more than 25,000 students. The state average score is 50.8. Two large districts have very low scores.

In New Mexico, DoDDS and DDESS score better than all of the districts and only one district has a score in the 60s. The state average is 49.9. The largest districts, Las Cruces and Albuquerque, score 55.8 and 55.1, respectively.

---

7 Sources: Data for Maryland are from web site www.msp.msde.state.md.us Data for New Mexico are from New Mexico State Department of Education, The Accountability Report, November 1999. Data for Nevada are from web site www.nsn.k12.nv.us.

Table II-8. Maryland School District, DoDDS and DDESS 1999 CTBS Scores

District	Enroll-ment	Gr. 4 Read	Gr. 4 Lang	Gr. 4 Math	Gr. 6 Read	Gr. 6 Lang	Gr. 6 Math	Aver-age
Montgomery	127,933	65	67	72	72	69	81	71
Howard	41,858	69	72	73	68	68	72	70
DoDDS	**76,000**	**68**	**66**	**64**	**65**	**69**	**65**	**66**
Garrett	5,082	71	66	62	66	66	62	66
Carroll	27,224	61	64	65	66	64	68	65
Harford	38,909	68	66	62	62	64	65	65
DDESS	**36,000**	**66**	**63**	**64**	**62**	**66**	**64**	**64**
Queen Anne's	6,888	64	67	68	59	59	64	64
Frederick	35,383	56	65	61	62	61	69	62
Kent	2,891	56	61	62	64	60	62	61
Talbot	4,590	56	58	60	60	54	69	60
Worcester	6,916	58	54	64	52	51	61	57
Charles	22,263	64	56	55	54	55	51	56
Anne Arundel	74,079	54	56	58	53	55	58	56
Baltimore County	105,914	55	60	56	56	54	52	56
Washington	20,159	53	56	55	52	50	57	54
Wicomico	14,330	50	56	46	56	56	54	53
Allegany	10,978	51	51	50	56	54	55	53
Cecil	15,550	49	56	53	50	50	51	52
State-Wide	**841,671**	**50**	**54**	**49**	**51**	**50**	**51**	**51**
Caroline	5,685	46	51	50	51	49	48	49
St. Mary's	14,743	48	51	42	48	49	49	48
Dorchester	5,143	48	50	37	44	45	36	43
Somerset	3,113	51	49	40	41	40	35	43
Prince George's	130,259	40	40	33	40	41	39	39
Baltimore City	106,540	34	33	24	26	25	19	27

Table II-9. New Mexico School District, DoDDS and DDESS 1998 CTBS Scores

District	Enrollment	Gr. 4 Avg.	Gr. 6 Avg.	Gr. 8 Avg.	Average
DoDDS	**76,000**	**66.0**	**66.2**	**66.2**	**66.1**
DDESS	**36,000**	**64.6**	**64.2**	**63.0**	**63.9**
Alamagordo	8,075	62.8	58.9	61.3	61.0
Hobbs	8,208	66.5	55.9	55.6	59.3
Rio Rancho	9,719	59.3	53.4	60.1	57.6
Clovis	8,712	59.4	58.3	53.9	57.2
Las Cruces	22,403	58.1	52.8	56.5	55.8
Carlsbad	6,728	57.0	51.7	58.1	55.6
Farmington	10,421	56.2	56.3	53.2	55.2
Albuquerque	85,847	60.6	48.7	55.9	55.1
Roswell	10,673	50.4	54.6	54.5	53.2
Santa Fe	14,712	52.4	52.5	47.5	50.8
State-Wide	**325,000**	**52.9**	**47.4**	**49.3**	**49.9**
Los Lunas	8,571	54.1	42.7	47.5	48.1
Central Consolidated	7,488	48.0	26.2	36.2	36.8
Deming	5,569	40.7	32.5	36.6	36.6
Gadsden	12,666	31.9	38.7	26.1	32.2
Espanola	5,165	34.7	32.7	27.8	31.7
Gallup	14,261	26.3	26.7	27.8	26.9

Table II-10. Nevada School District, DoDDS and DDESS 1999 CTBS Scores

District	Enroll-ment	Gr 4 Rd	Gr 4 Ma	Gr 4 Lg	Gr. 4 Sc	Gr 8 Rd	Gr 8 Ma	Gr 8 Lg	Gr 8 Sc	Gr 10 Rd	Gr 10 Ma	Gr 10 Lg	Gr 10 Sc	Av
Eureka	378	62	64	68	73	82	77	79	78	77	59	74	75	72
DoDDS	**76,000**	**68**	**64**	**66**	**65**	**65**	**64**	**69**	**67**	**72**	**70**	**72**	**68**	**68**
DDESS	**36,000**	**66**	**64**	**63**	**67**	**62**	**59**	**66**	**67**	**68**	**64**	**66**	**64**	**65**
Douglas	7,302	58	60	55	62	60	64	56	62	65	64	63	68	61
Washoe	50,948	53	50	52	60	59	48	57	58	59	57	61	63	56
Carson City	8,288	49	42	47	55	58	59	51	60	64	64	60	67	56
Lander	1,857	54	49	56	59	52	47	57	56	60	58	61	64	56
Storey	532	55	52	48	59	68	55	55	63	58	47	47	66	56
Churchill	4,766	49	52	44	56	56	51	51	54	61	53	58	64	54
Elko	10,586	50	44	51	56	53	51	49	57	53	51	55	61	53
State-Wide	**296,621**	**49**	**53**	**52**	**53**	**53**	**49**	**51**	**52**	**53**	**52**	**56**	**58**	**53**
White Pine	1,836	53	49	52	60	49	42	45	60	53	50	54	60	52
Lincoln	1,081	52	49	40	54	44	38	40	50	69	56	62	69	52
Clark	190,822	48	56	53	50	52	49	49	49	50	51	55	56	52
Lyon	6,154	48	48	48	55	49	48	43	54	55	49	54	60	51
Humboldt	4,257	50	52	46	58	50	45	48	55	47	42	48	56	50
Nye	5,227	48	41	42	53	49	34	46	52	50	44	49	59	47
Pershing	999	48	34	43	52	44	32	44	53	49	40	48	55	45
Mineral	1,046	35	28	33	47	46	36	46	50	37	37	44	55	41
Esme-realda	114	29	33	23	27	39	50	37	47					36

In Nevada, one very small district scores higher than DoDDS and DDESS and one other small district has a score in the 60s. The state average is 52.6. The largest districts, Washoe and Clark, score 56.4 and 51.5, respectively.

3. Conclusions

It is not possible to interpret the performance of DoDDS and DDESS on the CTBS/Terra Nova in a comparison across all states because not all states take the test. We can, however, obtain and present data for school districts within a few states.

The comparisons made here reveal that, unlike the comparisons with states using NAEP, where for reading and writing DoDDS and DDESS are near the top, there are a sizeable number of school districts on a par with DoDDS and DDESS.

It should be noted that in Maryland, Montgomery County, which has higher scores than DoDDS and DDESS, though relatively wealthy, is also ethnically diverse. Comparative ethnic percentages follow:[8]

Ethnicity	Montgomery County	DoDDS	DDESS
White	53.4	47	47
African American	20.3	18	26
Hispanic	13.2	7	4
Asian	12.7	8	4
Native American	.4	1	1
Other		19	18

A caveat in interpreting CTBS/Terra Nova data is that although the DoDDS and DDESS percentiles are high, the norm reference to the 50th percentile is suspect due to the Lake Wobegon Effect. However, this should not affect the comparative results with the school districts in Maryland, New Mexico and Nevada.

C. SCHOLASTIC APTITUDE TEST (SAT)

1. Background

The SAT program consists of the SAT I: Reasoning Test and the SAT II: Subject Tests and related products and services. The SAT I is a three-hour, primarily multiple-choice test that measures developed verbal and mathematical reasoning abilities related to successful performance in college. The SAT I test is designed to supplement the secondary school record and other information about the student in assessing readiness for college-level work. The SAT II tests are designed to measure knowledge in specific subject areas and the student's ability to apply this knowledge.

The SAT I: Reasoning Test is what is usually referred to as the SAT. It has two parts, verbal and mathematical. Scores on each part are from 200 (lowest) to 800 (highest).

---

8 Sources: Montgomery County data are from FY 2000 Metropolitan Area Boards of Education Guide, Produced by Fairfax County Public Schools, October 1999. DoDDS and DDESS data are from DoDEA.

2. DoDDS and DDESS 1999 SAT Results

Table II-11 presents 1999 SAT scores for DoDDS and DDESS as well as for the 50 states, the District of Columbia, and the nation. Also given are the participation rates.[9]

The states and other jurisdictions in the table are sorted from highest to lowest total score. Note, however, that the total SAT scores presented here can be off by one due to rounding.

Note that, for the first seventeen states (through Mississippi), participation rates were low and scores were high. The participation rates are low in these states because many colleges in them prefer or require applicants to take the ACT assessment test instead of the SAT. Accordingly, virtually the only the students in those states who take the SAT are those who are considering applying to a (likely out-of-state) college that requires the SAT.

In general, it is reasonable to believe that the smarter a student is, the more likely it is that that student will take the SAT. This means that one would expect that the average SAT scores across jurisdiction would usually vary inversely with participation rates. Thus, a jurisdiction with a lower SAT and higher participation rate cannot be directly compared to one with a higher SAT and lower participation rate.

Sections 3 and 4 below present two ways to compare the SAT scores and participation rate of a single jurisdiction (such as DoDDS or DDESS) with those of a collection of other jurisdictions (such as the states of the nation) considering this inverse relationship between scores and participation rates. Section 5 then presents a logical consistency check on the results of these two approaches.

3. Adjustment for Participation Rates—Strict Comparability Procedure

a. DoDDS

The DoDDS 1999 SAT participation rate was 63 percent. The DoDDS 1999 SAT total score was 1007.

Of the 51 states, counting the District of Columbia as a state, eight states (New Jersey, New Hampshire, Massachusetts, Vermont, Connecticut, Maryland, Maine, and Virginia) had a participation rate that was the same as or higher than DoDDS and also had a total score that was the same as or higher than DoDDS. Of these 51 states, six states (Florida, Hawaii, Indiana, Texas, North Carolina, and South Carolina) had a participation rate that was the same as or lower than DoDDS and also had a total score that was the same as or lower than DoDDS.

---

9 Source: College Board web site: www.collegeboard.org.

Table II-11. Average 1999 SAT Scores

State	Enrollment (000)	Verbal Score	Math Score	Total Score	% Participation
North Dakota	114	594	605	1199	5
Iowa	503	594	598	1192	5
Minnesota	858	586	598	1184	9
Wisconsin	888	584	595	1179	7
South Dakota	142	585	588	1173	4
Illinois	2,000	569	585	1154	12
Kansas	470	578	576	1154	9
Missouri	921	572	572	1144	8
Nebraska	291	568	571	1139	8
Utah	447	570	568	1138	5
Oklahoma	627	567	560	1127	8
Michigan	1,700	557	565	1122	11
Arkansas	456	563	556	1119	6
Louisiana	754	561	558	1119	8
Alabama	759	561	555	1116	9
Tennessee	909	559	553	1112	13
Mississippi	502	563	548	1111	4
New Jersey	1,300	598	510	1108	80
Ohio	1,800	534	568	1102	25
Wyoming	94	546	551	1097	10
Kentucky	646	547	547	1094	12
Montana	161	547	546	1093	21
New Mexico	329	549	542	1091	12
Idaho	245	542	540	1082	16
Colorado	699	536	540	1076	32
Washington	1,000	525	526	1051	52
Oregon	543	525	525	1050	53
Arizona	829	524	525	1049	34
West Virginia	296	527	512	1039	8
New Hampshire	195	520	518	1038	72
Alaska	134	516	514	1030	50
Nevada	311	512	517	1029	34
Massachusetts	964	511	511	1022	78
Vermont	105	514	506	1020	70
Connecticut	545	510	509	1019	80
Nation	**46,161**	**505**	**511**	**1016**	**43**
Maryland	837	507	507	1014	65
California	5,800	497	514	1011	49
Maine	220	507	503	1010	68
Virginia	1,100	508	499	1007	65
DoDDS	**76**	**506**	**501**	**1007**	**63**
Rhode Island	154	504	499	1003	70
Delaware	113	503	497	1000	67
Florida	2,300	499	498	997	53
New York	2,900	495	502	997	76
Hawaii	187	482	513	995	52
Indiana	989	496	498	994	60
Pennsylvania	1,800	498	495	993	70
Texas	3,900	494	499	993	50
North Carolina	1,200	493	493	986	61
Dist. of Columbia	80	494	478	972	77
Georgia	1,400	487	482	969	63
DDESS	**36**	**483**	**474**	**957**	**34**
South Carolina	644	479	475	954	61

No state had both the same participation rate and the same total score as DoDDS.

Thus, there are 14 states that are strictly comparable with DoDDS according to the 1999 SAT participation rates and total scores. DoDDS had both an equal or better participation rate and an equal or better total score than 6 of these 14 comparable states, and DoDDS had both an equal or worse participation rate and an equal or worse total score than 8 of these 14 comparable states. The remaining 37 states are not strictly comparable to DoDDS according to the 1999 SAT participation rates and total scores.

b. DDESS

The DDESS 1999 SAT participation rate was 34 percent. The DDESS 1999 SAT total score was 957.

Of the 51 states, counting the District of Columbia as a state, 25 states had a participation rate that was the same as or higher than DDESS and also had a total score that was the same as or higher than DDESS. Of these 51 states, no state had a participation rate that was the same as or lower than DDESS and also had a total score that was the same as or lower than DDESS.

Accordingly, no state had both the same participation rate and the same total score as DDESS.

Thus, there are 25 states that are strictly comparable with DDESS according to the 1999 SAT participation rates and total scores. DDESS had both an equal or better participation rate and an equal or better total score than none of these 25 comparable states, and DDESS had both an equal or worse participation rate and an equal or worse total score than all 25 of these 25 comparable states. The remaining 26 states are not strictly comparable to DDESS according to the 1999 SAT participation rates and total scores.

c. Conclusions

DoDDS does better than 6, and worse than 8, of the 14 states that it can be compared with according to this measure. This is about average.

DDESS does worse than all 25 of the 25 states that it can be compared with according to this measure. This is far below average.

4. Adjustment for Participation Rates—Elimination Procedure

The elimination procedure applied in this section attempts to preserve data from more states in the comparison of DoDDS and DDESS with the nation than does the strict comparability method of the previous section

a. Starting Point

Table II-12 gives the starting point for this SAT comparison. Each state's participation percentage times its fraction of the total enrollment gives the percent of the total enrollment participating in the SAT from that state. This percent of total enrollment participating is proportional to the number of students from that state who took the SAT. Accordingly, this percent of total enrollment participating times the state's average SAT score per test-taking student gives a relative measure of the contribution of that state to multi-state averages. In particular, the average SAT score of any subset of states equals the sum of these contributions from the states in that subset divided by the sum of the percent of the total enrollment participating from those states. For example, the nation's average total SAT score is the sum over all states of the state contributions, 43173, divided by the sum over all states of the state percent of US enrollment participating, 42.3, which is 1021. (Note that these values lack some precision due to rounding in the table entries.)

b. DoDDS

For DoDDS, taking as the starting point all of the states in the nation—with a participation rate of 43 percent—the procedure eliminates states with lower participation rates (beginning with the state with the lowest participation rate). It continues until the remaining states have a participation rate of 63 percent, identical to DoDDS. It then compares the average SAT score of those remaining states to that of DoDDS

Table II-13 gives the DoDDS SAT comparison. When the nation is modified to yield a 63 percent participation rate, the average SAT is (63724 / 63.25) = 1007.

c. DDESS

For DDESS, taking as the starting point all of the states in the nation—with a participation rate of 43 percent—the procedure eliminates states with higher participation rates (beginning with the state with the highest participation rate). It continues until the remaining states have a participation rate of 34 percent, identical to DDESS. It then compares the average SAT of the remaining states to that of DDESS.

Table II-14 gives the DDESS SAT comparison. When the nation is modified to yield a 34 percent participation rate, the average SAT is (34837 / 34.13) = 1021.

Table II-12. Initial SAT Data for the Elimination Procedure

State	% Participation	Enrollment (000)	Fraction of Total Enrollment	% of Total Enrollment Participating	Total SAT Score	Contribution
New Jersey	80	1,300	0.02816	2.2530	1108	2496.31
Connecticut	80	545	0.01181	0.9445	1019	962.47
Massachusetts	78	964	0.02088	1.6289	1022	1664.74
Dist. of Columbia	77	80	0.00173	0.1334	972	129.71
New York	76	2,900	0.06282	4.7746	997	4760.27
New Hampshire	72	195	0.00422	0.3042	1038	315.71
Vermont	70	105	0.00227	0.1592	1020	162.41
Rhode Island	70	154	0.00334	0.2335	1003	234.23
Pennsylvania	70	1,800	0.03899	2.7296	993	2710.47
Maine	68	220	0.00477	0.3241	1010	327.32
Delaware	67	113	0.00245	0.1640	1000	164.01
Maryland	65	837	0.01813	1.1786	1014	1195.09
Virginia	65	1,100	0.02383	1.5489	1007	1559.77
Georgia	63	1,400	0.03033	1.9107	969	1851.47
North Carolina	61	1,200	0.026	1.5858	986	1563.55
South Carolina	61	644	0.01395	0.8510	954	811.87
Indiana	60	989	0.02143	1.2855	994	1277.79
Oregon	53	543	0.01176	0.6234	1050	654.62
Florida	53	2,300	0.04983	2.6408	997	2632.84
Washington	52	1,000	0.02166	1.1265	1051	1183.94
Hawaii	52	187	0.00405	0.2107	995	209.60
Alaska	50	134	0.0029	0.1451	1030	149.50
Texas	50	3,900	0.08449	4.2243	993	4194.77
California	49	5,800	0.12565	6.1567	1011	6224.44
Arizona	34	829	0.01796	0.6106	1049	640.52
Nevada	34	311	0.00674	0.2291	1029	235.71
Colorado	32	699	0.01514	0.4846	1076	521.39
Ohio	25	1,800	0.03899	0.9748	1102	1074.28
Montana	21	161	0.00349	0.0732	1093	80.06
Idaho	16	245	0.00531	0.0849	1082	91.88
Tennessee	13	909	0.01969	0.2560	1112	284.67
Illinois	12	2,000	0.04333	0.5199	1154	599.99
Kentucky	12	646	0.01399	0.1679	1094	183.72
New Mexico	12	329	0.00713	0.0855	1091	93.31
Michigan	11	1,700	0.03683	0.4051	1122	454.53
Wyoming	10	94	0.00204	0.0204	1097	22.34
Minnesota	9	858	0.01859	0.1673	1184	198.06
Kansas	9	470	0.01018	0.0916	1154	105.75
Alabama	9	759	0.01644	0.1480	1116	165.15
Missouri	8	921	0.01995	0.1596	1144	182.60
Nebraska	8	291	0.0063	0.0504	1139	57.44
Oklahoma	8	627	0.01358	0.1087	1127	122.46
Louisiana	8	754	0.01633	0.1307	1119	146.22
West Virginia	8	296	0.00641	0.0513	1039	53.30
Wisconsin	7	888	0.01924	0.1347	1179	158.76
Arkansas	6	456	0.00988	0.0593	1119	66.32
North Dakota	5	114	0.00247	0.0123	1199	14.81
Iowa	5	503	0.0109	0.0545	1192	64.94
Utah	5	447	0.00968	0.0484	1138	55.10
South Dakota	4	142	0.00308	0.0123	1173	14.43
Mississippi	4	502	0.01087	0.0435	1111	48.33
Total		46,161	1	42.3217		43172.99

Table II-13. States Whose Combined Average SAT Participation Rate Is 63%

State	% Participatio	Enrollment (000)	Fraction of Total Enrollment	% of Total Enrollment Participating	Total SAT Score	Contribution
New Jersey	80	1,300	0.0575	4.5997	1108	5096.51
Connecticut	80	545	0.0241	1.9284	1019	1964.99
Massachusetts	78	964	0.04264	3.3256	1022	3398.77
Dist. of Columbia	77	80	0.00354	0.2724	972	264.82
New York	76	2,900	0.12826	9.7479	997	9718.66
New Hampshire	72	195	0.00862	0.6210	1038	644.56
Vermont	70	105	0.00464	0.3251	1020	331.58
Rhode Island	70	154	0.00681	0.4768	1003	478.21
Pennsylvania	70	1,800	0.07961	5.5728	993	5533.75
Maine	68	220	0.00973	0.6617	1010	668.27
Delaware	67	113	0.005	0.3349	1000	334.85
Maryland	65	837	0.03702	2.4062	1014	2439.92
Virginia	65	1,100	0.04865	3.1623	1007	3184.45
Georgia	63	1,400	0.06192	3.9009	969	3780.00
North Carolina	61	1,200	0.05307	3.2375	986	3192.18
South Carolina	61	644	0.02848	1.7375	954	1657.54
Indiana	60	989	0.04374	2.6245	994	2608.76
Oregon	53	543	0.02402	1.2728	1050	1336.49
Florida	53	2,300	0.10172	5.3914	997	5375.25
Washington	52	1,000	0.04423	2.2999	1051	2417.16
Hawaii	52	187	0.00827	0.4301	995	427.92
Alaska	50	134	0.00593	0.2963	1030	305.22
Texas	50	3,900	0.17249	8.6245	993	8564.13
Total		22,610	1	63.2501		63723.98

d. Conclusions

Adjusting the national data by deleting states until the remaining states match the DoDDS participation rate of 63 percent results in those remaining states having an average total SAT score of 1007. This is the same as the score of DoDDS. Hence, DoDDS performance is average for the nation according to this method of considering participation rates.

Adjusting the national data by deleting states until the remaining states match the DDESS participation rate of 34 percent results in those remaining states having an average total SAT score of 1021. This is far higher than the score of DDESS. Hence, DDESS performance is far below average for the nation according to this method of considering participation rates.

Table II-14. States Whose Combined Average SAT Participation Rate Is 34%

State	% Participation	Enrollment (000)	Fraction of Total Enrollment	% of Total Enrollment Participating	Total SAT Score	Contribution
Maryland	65	837	0.02282	1.4830	1014	1503.79
Georgia	63	1,400	0.03816	2.4043	969	2329.72
North Carolina	61	1,200	0.03271	1.9954	986	1967.43
South Carolina	61	644	0.01755	1.0708	954	1021.59
Indiana	60	989	0.02696	1.6176	994	1607.85
Oregon	53	543	0.0148	0.7845	1050	823.71
Florida	53	2,300	0.0627	3.3229	997	3312.92
Washington	52	1,000	0.02726	1.4175	1051	1489.76
Hawaii	52	187	0.0051	0.2651	995	263.74
Alaska	50	134	0.00365	0.1826	1030	188.12
Texas	50	3,900	0.10631	5.3155	993	5278.32
California	49	5,800	0.1581	7.7470	1011	7832.25
Arizona	34	829	0.0226	0.7683	1049	805.97
Nevada	34	311	0.00848	0.2882	1029	296.60
Colorado	32	699	0.01905	0.6097	1076	656.07
Ohio	25	1,800	0.04907	1.2267	1102	1351.78
Montana	21	161	0.00439	0.0922	1093	100.73
Idaho	16	245	0.00668	0.1069	1082	115.62
Tennessee	13	909	0.02478	0.3221	1112	358.20
Illinois	12	2,000	0.05452	0.6542	1154	754.97
Kentucky	12	646	0.01761	0.2113	1094	231.18
New Mexico	12	329	0.00897	0.1076	1091	117.41
Michigan	11	1,700	0.04634	0.5097	1122	571.93
Wyoming	10	94	0.00256	0.0256	1097	28.11
Minnesota	9	858	0.02339	0.2105	1184	249.23
Kansas	9	470	0.01281	0.1153	1154	133.06
Alabama	9	759	0.02069	0.1862	1116	207.81
Missouri	8	921	0.02511	0.2008	1144	229.77
Nebraska	8	291	0.00793	0.0635	1139	72.28
Oklahoma	8	627	0.01709	0.1367	1127	154.10
Louisiana	8	754	0.02055	0.1644	1119	183.99
West Virginia	8	296	0.00807	0.0645	1039	67.07
Wisconsin	7	888	0.02421	0.1694	1179	199.77
Arkansas	6	456	0.01243	0.0746	1119	83.46
North Dakota	5	114	0.00311	0.0155	1199	18.63
Iowa	5	503	0.01371	0.0686	1192	81.72
Utah	5	447	0.01218	0.0609	1138	69.33
South Dakota	4	142	0.00387	0.0155	1173	18.16
Mississippi	4	502	0.01368	0.0547	1111	60.81
Total		36,685	1	34.1301		34836.95

5. Adjustment for Participation Rates—Hypothetical Score Projection Procedure

As stated in Section C.2 above, the basic reason for considering participation rates here is the belief that, the smarter a student is, the more likely it is that the student will take the SAT. Thus, jurisdictions with higher percentages of their students taking the SAT can expect to have somewhat lower average SAT scores due to this effect. If this belief is valid, then there must be some function, which may depend on the school system involved, that relates expected SAT scores of students to the propensity of students to take the SAT. Furthermore, this function should have certain, well-defined properties.

The purpose of this section is to demonstrate that it is possible for such a function to exist and to be consistent with the data given above. This is done by producing such a function for each of the two comparisons being made. The argument is not that the functions produced here are necessarily correct, or even close to being correct. Instead, the argument is that, if no such function could be found, then the underlying premise of the procedures above would be called into question. Conversely, exhibiting these functions provides a theoretical check on the logical consistency of the two data-driven approaches presented above.

a. Formulas

Throughout this section, "score" will mean either a math or a verbal SAT score minus 200. Thus, all scores run from a minimum of 0 to a maximum of 600. In particular, let $m = 600$.

For any given jurisdiction and either math or verbal SAT, let:

B = the (presumed known and positive) total number of students in that jurisdiction who are eligible to take the SAT.

$T(s)$ = the (presumed known) number of students who took the SAT and scored s or higher.

$p = T(0)/B$ = the fraction of eligible students who took the SAT. It is assumed throughout that $0 < p < 1$.

$t(s) = T(s)/B$ = the fraction of eligible students who took the SAT and scored s or higher, so

$$1 > p = t(0) \geq t(s) \geq t(s') \geq t(m) \geq 0$$

for all s and s' such that $0 \leq s \leq s' \leq m$.

$u(s) = t(s)/p = T(s)/pB$ = the fraction of SAT-taking students who scored s or higher, so

$$1 = u(0) \geq u(s) \geq u(s') \geq u(m) \geq 0$$

for all s and s' such that $0 \leq s \leq s' \leq m$.

$A(s,p,q)$ = the (unknown for $q > p$) number of students who would score s or higher if qB students were to take the SAT, where $p \le q \le 1$.

$a(s,p,q) = A(s,p,q)/B$ = the (unknown for $q > p$) fraction of eligible students who would score s or higher if qB students were to take the SAT, where $p \le q \le 1$.

$r(s,y)$ = the (unknown) marginal fractional rate at which students would score s or higher at the point at which yB students were taking the SAT. That is, for small fractional z, if $(y + z)B$ students were to take the SAT, then the number of students who score would s or higher would be about $A(s,p,y) + r(s,y)zB$. Therefore,

$$A(s,p,q) = T(s) + B\int_p^q r(s,y)dy,$$

and so

$$a(s,p,q) = t(s) + \int_p^q r(s,y)dy \tag{5.1}$$

for $p \le q \le 1$.

With this notation, the statement that "the smarter a student is, the more likely it is that that student will take the SAT" can be formalized as

$$u(s) \ge r(s,p) \ge r(s,y) \ge r(s,q) \ge 0 \tag{5.2}$$

for all relevant s and all y and q such that $p \le y \le q \le 1$. Clearly, $r(s,y)$ must also satisfy the property that

$$1 = u(0) = r(0,y) \ge r(s,y) \ge r(s',y) \ge r(m,y) \ge 0 \tag{5.3}$$

for all relevant y and all s and s' such that $0 \le s \le s' \le m$.

The goal of this section is to demonstrate that there exists at least one functional form for $r(s,y)$ that satisfies (5.2) and (5.3), and that, when plugged into (5.1), reproduces the results of the two data-driven approaches presented above.

A simple function that satisfies (5.2) and (5.3) is constructed as follows. First, note that $r(s,y)$ must, in general, be defined over the rectangle whose corners are

$$(s = 0, y = p),\ (s = m, y = p)\ (s = m, y = 1)\ (s = 0, y = 1).$$

Second, note that, by (5.2), $u(s) \ge r(s,p) \ge r(s,y)$ for all relevant s and y. Thus, $u(s)$ is an upper bound on $r(s,p)$, which is an upper bound on $r(s,y)$ for all relevant s and y. Accordingly, setting these upper bounds equal gives $r(s,p) = u(s)$ for all relevant s. This provides a reasonable specification for $r(s,y)$ over one side (the $y = p$ side) of that rectangle.

Third, note that, by (5.3), $r(0,1)$ must equal $u(0)$.

Fourth, note that $r(m,1)$ is essentially the likelihood that, if all but one eligible student were to take the SAT, then the one student who did not take the SAT would have scored the maximum had he or she done so. Accordingly, $r(m,1)$ is quite reasonably set equal to zero.

Fifth, given that $r(0,1) = 1\times u(0)$, that $r(m,1) = 0 = 0\times u(1)$, and that $r(s,1)$ must not exceed $u(s)$ for any s, a first-order estimate for $r(s,1)$ between $s = 0$ and $s = m$ is $u(s)$ times a term that is linear in s between 1 at s = 0 and 0 at $s = m$. This yields $r(s,1) = ((m - s)/m)u(s)$ for $0 \le s \le m$, which defines $r(s,y)$ over the $y = 1$ side of the rectangle.

Sixth, given that $r(s,y)$ is defined over two opposite sides of a rectangle, it is easy to define it anywhere inside that rectangle by linear interpolation. By the second step above, this interpolation should give that $r(s,y) = u(s)$ when $y = p$. By the fifth step above, this interpolation should give that $r(s,y) = u(s)((m - s)/m)$ when $y = 1$. Linearly interpolating over y between $y = p$ and $y = 1$ gives that

$$r(s,y) = u(s)\left(\frac{1-y}{1-p} + \frac{y-p}{1-p}\,\frac{m-s}{m}\right) \tag{5.4}$$

for $0 \le s \le m$ and $p \le y \le 1$. Note that this specification of $r(s,y)$ satisfies (5.2) and (5.3) above.

Plugging (5.4) into (5.1) gives that

$$a(s,p,q) = t(s) + \int_p^q \frac{u(s)}{1-p}\left(1 - y + \frac{(m-s)y-(m-s)p}{m}\right)dy$$

$$= t(s)\left(1 + \frac{1}{p(1-p)}\int_p^q (1-(1-s/m)p-(s/m)y)dy\right)$$

$$= t(s)\left(1 + \frac{1}{p(1-p)}\left((1-(1-s/m)p)(q-p) - (s/m)(q^2-p^2)/2\right)\right)$$

$$= t(s)\left(1 + \frac{(q-p)}{p(1-p)}\left((1-(1-s/m)p) - (s/m)(q+p)/2\right)\right),$$

so

$$a(s,p,q) = t(s)\left(\frac{q}{p} - \frac{(q-p)^2 s}{2p(1-p)m}\right). \tag{5.5}$$

b. Using Equation (5.5) to Compare DoDDS's and National 1999 SAT Scores

As discussed above, the 1999 national SAT participation rate was 43% and the 1999 DoDDS SAT participation rate was 63%. Thus, equation (5.5) can be used to project the nation's scores at 43% participation to hypothetical scores that the nation might have achieved at

63% participation. If these projected national scores are comparable with DoDDS's scores, then this provides a theoretical explanation for the data-driven results of Sections 3 and 4 above.

Equation (5.5) projects scores from a 43% participation rate to a 63% participation rate by setting $p = 0.43$ and $q = 0.63$. Doing this, along with setting $m = 600$, gives

$$a(s,.43,.63) = t(s)(1.4651 - 0.000136s). \tag{5.6}$$

This equation is used to calculate values for Tables II-15 and II-16 as follows.

Consider Table II-15. The first column of that table is just a list of possible "math-SAT-minus-200" scores, s_i, running from $s_1 = 600$ down to $s_{61} = 0$. The i^{th} row of the second column gives $u(s_i)$, which, here, is the fraction of students who scored s_i or higher on the math SAT out of all of the students in the nation who took the 1999 SAT. These data were obtained on May 31, 2000, from the College Board Internet site at

http://www.collegeboard.org/index_this/sat/cbsenior/stats/stat001b.html

Let $s_0 = 0$. Then the i^{th} row of the third column gives the value of $s_i(u(s_{i-1}) - u(s_i))$. Therefore, except for rounding in the second column, the average national math SAT score (minus 200) would equal the sum of the terms in the third column. This sum is 311.68 and the national average (minus 200) is 311, so this rounding had a very small impact.

The i^{th} row of the fourth column gives $t(s_i) = pu(s_i) = 0.43u(s_i)$, which, here, is the fraction of students who took the SAT and scored s_i or higher on the math portion, out of all of the students in the nation eligible to take the 1999 SAT. The i^{th} row of the fifth column gives the value of $a(s_i,.43,.63)$ as computed by equation (5.6). The i^{th} row of the sixth column gives the value of $a(s_i,.43,.63)/q = a(s_i,.43,.63)/.63$, which, here, is the projected fraction of students who would score s_i or higher on the math SAT out of all of the students in the nation who would have taken the 1999 SAT had the 1999 national participation rate been 63%.

The i^{th} row of the seventh column gives the value of

$$s_i(a(s_{i-1},.43,.63) - a(s_i,.43,.63))/.63.$$

Thus, the projected average national math SAT score (minus 200) equals the sum of the terms in the seventh column. This sum is 306.42, and so the projection from 43% participation to 63% participation gives that the average national math SAT score would drop from 511.68 to 506.42.

The last two columns of Table II-15 give corresponding data for DoDDS. The i^{th} row of the eighth column gives the DoDDS value for $u(s_i)$, which is the fraction of students who scored s_i or higher on the math SAT out of all of the DoDDS students who took the 1999 SAT. These

Table II-15. Distribution and Projection of 1999 Math SAT Scores for the Nation and DoDDS

	Nation						DoDDS	
s_i	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$	$t(s_i)$	$a(s_i,p,q)$	$a(s_i,p,q)/q$	$(s_i/q) \times a((s_{i-1},p,q) - a(s_i,p,q))$	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$
600	0.01	6.0	0.0043	0.005949	0.009443	5.665762	0.001	0.60
590	0.01	0.0	0.0043	0.005955	0.009452	0.005477	0.003	1.18
580	0.01	0.0	0.0043	0.005961	0.009462	0.005384	0.005	1.16
570	0.01	0.0	0.0043	0.005967	0.009471	0.005291	0.007	1.14
560	0.02	5.6	0.0086	0.011945	0.018960	5.314036	0.009	1.12
550	0.02	0.0	0.0086	0.011957	0.018979	0.010211	0.010	0.55
540	0.02	0.0	0.0086	0.011968	0.018997	0.010025	0.011	0.54
530	0.03	5.3	0.0129	0.017970	0.028524	5.049034	0.017	3.18
520	0.04	5.2	0.0172	0.023983	0.038069	4.963423	0.023	3.12
510	0.05	5.1	0.0215	0.030008	0.047632	4.877441	0.027	2.04
500	0.06	5.0	0.0258	0.036045	0.057215	4.791087	0.037	5.00
490	0.07	4.9	0.0301	0.042094	0.066815	4.704362	0.045	3.92
480	0.08	4.8	0.0344	0.048154	0.076435	4.617266	0.052	3.36
470	0.09	4.7	0.0387	0.054226	0.086072	4.529799	0.061	4.23
460	0.11	9.2	0.0473	0.066340	0.105302	8.845490	0.082	9.66
450	0.13	9.0	0.0559	0.078478	0.124568	8.669906	0.092	4.50
440	0.15	8.8	0.0645	0.090639	0.143872	8.493578	0.109	7.48
430	0.17	8.6	0.0731	0.102824	0.163213	8.316509	0.121	5.16
420	0.19	8.4	0.0817	0.115032	0.182590	8.138696	0.149	11.76
410	0.22	12.3	0.0946	0.133324	0.211625	11.90406	0.173	9.84
400	0.24	8.0	0.1032	0.145584	0.231086	7.784556	0.199	10.40
390	0.26	7.8	0.1118	0.157868	0.250585	7.604423	0.229	11.70
380	0.29	11.4	0.1247	0.176253	0.279767	11.08947	0.254	9.50
370	0.32	11.1	0.1376	0.194674	0.309006	10.81824	0.288	12.58
360	0.35	10.8	0.1505	0.213129	0.338300	10.54591	0.320	11.52
350	0.38	10.5	0.1634	0.231620	0.367650	10.27246	0.344	8.40
340	0.41	10.2	0.1763	0.250145	0.397056	9.997899	0.381	12.58
330	0.46	16.5	0.1978	0.280920	0.445904	16.11998	0.406	8.25
320	0.49	9.6	0.2107	0.299527	0.475440	9.451373	0.448	13.44
310	0.52	9.3	0.2236	0.318169	0.505031	9.173283	0.478	9.30
300	0.56	12	0.2408	0.342971	0.544399	11.81050	0.514	10.80
290	0.58	5.8	0.2494	0.355560	0.564380	5.794552	0.552	11.02
280	0.61	8.4	0.2623	0.374307	0.594139	8.33233	0.585	9.24
270	0.65	10.8	0.2795	0.399232	0.633702	10.68208	0.632	12.69
260	0.68	7.8	0.2924	0.418056	0.663581	7.768538	0.647	3.90
250	0.71	7.5	0.3053	0.436915	0.693516	7.483672	0.673	6.50
240	0.74	7.2	0.3182	0.455809	0.723506	7.197692	0.718	10.80
230	0.77	6.9	0.3311	0.474738	0.753552	6.910598	0.757	8.97
220	0.80	6.6	0.3440	0.493702	0.783654	6.622390	0.781	5.28
210	0.82	4.2	0.3526	0.506524	0.804006	4.274028	0.808	5.67
200	0.84	4.0	0.3612	0.519369	0.824396	4.077929	0.832	4.80
190	0.86	3.8	0.3698	0.532238	0.844823	3.881087	0.855	4.37
180	0.89	5.4	0.3827	0.551325	0.875119	5.453407	0.873	3.24
170	0.90	1.7	0.3870	0.558046	0.885788	1.813599	0.892	3.23
160	0.92	3.2	0.3956	0.570985	0.906326	3.286106	0.918	4.16
150	0.93	1.5	0.3999	0.577736	0.917041	1.607197	0.930	1.80
140	0.94	1.4	0.4042	0.584497	0.927774	1.502649	0.943	1.82
130	0.95	1.3	0.4085	0.591271	0.938526	1.397731	0.958	1.95
120	0.96	1.2	0.4128	0.598056	0.949296	1.292441	0.968	1.20
110	0.97	1.1	0.4171	0.604853	0.960085	1.186780	0.972	0.44
100	0.97	0.0	0.4171	0.605421	0.960985	0.090041	0.980	0.80

Table II-15. Distribution and Projection of 1999 Math SAT Scores for the Nation and DoDDS (cont)

	Nation						DoDDS	
s_i	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$	$t(s_i)$	$a(s_i,p,q)$	$a(s_i,p,q)/q$	$(s_i/q) \times a((s_{i-1},p,q) - a(s_i,p,q))$	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$
90	0.98	0.9	0.4214	0.612235	0.971802	0.973508	0.984	0.36
80	0.98	0.0	0.4214	0.612808	0.972712	0.072775	0.985	0.08
70	0.99	0.7	0.4257	0.619640	0.983556	0.759122	0.988	0.21
60	0.99	0.0	0.4257	0.620219	0.984475	0.055138	0.989	0.06
50	0.99	0.0	0.4257	0.620798	0.985394	0.045949	0.990	0.05
40	0.99	0.0	0.4257	0.621377	0.986313	0.036759	0.994	0.16
30	0.99	0.0	0.4257	0.621956	0.987232	0.027569	0.995	0.03
20	0.999	0.2	0.4296	0.628195	0.997134	0.198043	0.996	0.02
10	0.999	0.0	0.4296	0.628779	0.998062	0.009273	0.996	0.00
0	1	0	0.4300	0.629993	0.999989	0	1	0
Sum	—	311.68	—	—	—	306.4179	—	300.86

Table II-16. Distribution and Projection of 1999 Verbal SAT Scores for the Nation and DoDDS

	Nation						DoDDS	
s_i	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$	$t(s_i)$	$a(s_i,p,q)$	$a(s_i,p,q)/q$	$(s_i/q) \times a((s_{i-1},p,q) - a(s_i,p,q))$	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$
600	0.001	0.6	0.00043	0.000595	0.000944	0.566576	0.003	1.80
590	0.01	5.3	0.0043	0.005955	0.009452	5.019676	0.004	0.59
580	0.01	0.0	0.0043	0.005961	0.009462	0.005384	0.005	0.58
570	0.01	0.0	0.0043	0.005967	0.009471	0.005291	0.007	1.14
560	0.01	0.0	0.0043	0.005972	0.009480	0.005198	0.010	1.68
550	0.02	5.5	0.0086	0.011957	0.018979	5.224247	0.014	2.20
540	0.02	0.0	0.0086	0.011968	0.018997	0.010025	0.016	1.08
530	0.03	5.3	0.0129	0.017970	0.028524	5.049034	0.019	1.59
520	0.03	0.0	0.0129	0.017988	0.028552	0.014481	0.025	3.12
510	0.04	5.1	0.0172	0.024007	0.038106	4.872707	0.029	2.04
500	0.05	5.0	0.0215	0.030038	0.047679	4.786446	0.035	3.00
490	0.06	4.9	0.0258	0.036080	0.057270	4.699814	0.044	4.41
480	0.07	4.8	0.0301	0.042135	0.066880	4.612811	0.056	5.76
470	0.08	4.7	0.0344	0.048201	0.076509	4.525436	0.069	6.11
460	0.09	4.6	0.0387	0.054278	0.086156	4.437690	0.083	6.44
450	0.11	9.0	0.0473	0.066404	0.105404	8.661551	0.104	9.45
440	0.12	4.4	0.0516	0.072511	0.115097	4.265169	0.124	8.80
430	0.14	8.6	0.0602	0.084679	0.134410	8.304534	0.133	3.87
420	0.17	12.6	0.0731	0.102923	0.163370	12.163210	0.153	8.40
410	0.18	4.1	0.0774	0.109083	0.173147	4.008614	0.170	6.97
400	0.21	12.0	0.0903	0.127386	0.202200	11.621140	0.202	12.80
390	0.24	11.7	0.1032	0.145725	0.231309	11.352330	0.233	12.09
380	0.26	7.6	0.1118	0.158020	0.250826	7.416493	0.264	11.78
370	0.30	14.8	0.1290	0.182507	0.289693	14.380820	0.298	12.58
360	0.33	10.8	0.1419	0.200950	0.318969	10.539230	0.328	10.80
350	0.36	10.5	0.1548	0.219429	0.348300	10.265960	0.356	9.80
340	0.39	10.2	0.1677	0.237943	0.377687	9.991586	0.387	10.54
330	0.43	13.2	0.1849	0.262599	0.416823	12.914970	0.425	12.54
320	0.46	9.6	0.1978	0.281189	0.446331	9.442461	0.470	14.40
310	0.50	12.4	0.2150	0.305932	0.485607	12.175410	0.506	11.16
300	0.53	9.0	0.2279	0.324598	0.515235	8.888510	0.540	10.20

Table II-16. Distribution and Projection of 1999 Verbal SAT Scores for the Nation and DoDDS (cont)

s_i	Nation						DoDDS	
	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$	$t(s_i)$	$a(s_i,p,q)$	$a(s_i,p,q)/q$	$(s_i/q) \times a((s_{i-1},p,q) - a(s_i,p,q))$	$u(s_i)$	$s_i \times (u(s_{i-1}) - u(s_i))$
290	0.57	11.6	0.2451	0.349429	0.554650	11.430280	0.579	11.31
280	0.61	11.2	0.2623	0.374307	0.594139	11.056920	0.620	11.48
270	0.64	8.1	0.2752	0.393090	0.623953	8.049784	0.648	7.56
260	0.67	7.8	0.2881	0.411908	0.653822	7.766125	0.687	10.14
250	0.71	10.0	0.3053	0.436915	0.693516	9.923308	0.721	8.50
240	0.73	4.8	0.3139	0.449649	0.713729	4.851186	0.751	7.20
230	0.76	6.9	0.3268	0.468572	0.743766	6.908463	0.777	5.98
220	0.79	6.6	0.3397	0.487531	0.773858	6.620348	0.802	5.50
210	0.82	6.3	0.3526	0.506524	0.804006	6.331119	0.831	6.09
200	0.84	4.0	0.3612	0.519369	0.824396	4.077929	0.848	3.40
190	0.86	3.8	0.3698	0.532238	0.844823	3.881087	0.876	5.32
180	0.88	3.6	0.3784	0.545131	0.865287	3.683503	0.893	3.06
170	0.90	3.4	0.3870	0.558046	0.885788	3.485176	0.905	2.04
160	0.92	3.2	0.3956	0.570985	0.906326	3.286106	0.921	2.56
150	0.93	1.5	0.3999	0.577736	0.917041	1.607197	0.932	1.65
140	0.94	1.4	0.4042	0.584497	0.927774	1.502649	0.942	1.40
130	0.95	1.3	0.4085	0.591271	0.938526	1.397731	0.947	0.65
120	0.96	1.2	0.4128	0.598056	0.949296	1.292441	0.958	1.32
110	0.97	1.1	0.4171	0.604853	0.960085	1.186780	0.968	1.10
100	0.97	0.0	0.4171	0.605421	0.960985	0.090041	0.972	0.40
90	0.98	0.9	0.4214	0.612235	0.971802	0.973508	0.981	0.81
80	0.98	0.0	0.4214	0.612808	0.972712	0.072775	0.983	0.16
70	0.98	0.0	0.4214	0.613381	0.973621	0.063678	0.986	0.21
60	0.99	0.6	0.4257	0.620219	0.984475	0.651233	0.990	0.24
50	0.99	0.0	0.4257	0.620798	0.985394	0.045949	0.991	0.05
40	0.99	0.0	0.4257	0.621377	0.986313	0.036759	0.993	0.08
30	0.99	0.0	0.4257	0.621956	0.987232	0.027569	0.994	0.03
20	0.99	0.0	0.4257	0.622535	0.988151	0.018379	0.995	0.02
10	0.999	0.1	0.4296	0.628779	0.998062	0.099105	0.996	0.01
0	1	0	0.4300	0.629993	0.999989	0	1	0
Sum	—	305.7	—	—	—	300.6439	—	305.99

data were obtained from DoDEA on April 6, 2000. The average DoDDS math SAT score (minus 200) is the sum of the terms in the ninth column.

Table II-16 gives the corresponding data for the verbal portion of the 1999 SAT.

The "calculated" data in Table II-17 are based on the bottom lines of Tables II-15 and II-16, with the floor of 200 being added back in. These results are consistent with the hypothesis that DoDDS performance on the SAT in 1999 is about equal to the national average when participation rates are taken into consideration.

c. Using Equation (5.5) to Compare DDESS's and National 1999 SAT Scores

DDESS had lower average SAT scores than the nation in 1999, and projecting those scores from DDESS's 34% participation to the nations 43% participation can only lower them further. Thus, such projections cannot change the hypothesis that DDESS performance on the SAT in 1999 is below the national average (whether or not participation rates are taken into account). Still, equation (5.5) can be used to estimate how far the DDESS average scores would

Table II-17. Results of Projecting the Nation's 1999 SAT Scores to the DoDDS Participation Rate

	Math	Verbal	Total
Reported 1999 Average SAT Scores for the Nation at its 43% Participation Rate	511	505	1017
Calculated 1999 Average SAT Scores for the Nation at its 43% Participation Rate	511.68	505.70	1017.38
Calculated 1999 Average SAT Scores for the Nation Projected to a 63% Participation Rate	506.42	500.64	1007.06
Calculated 1999 Average SAT Scores for DoDDS at its 63% Participation Rate	500.86	505.99	1006.85
Reported 1999 Average SAT Scores for DoDDS at its 63% Participation Rate	501	506	1007

be below national average scores if participation rates are considered. In particular, equation (5.5) can be used to project scores from a 34% participation rate to a 43% participation rate by setting $p = 0.34$ and $q = 0.43$. Doing this, along with setting $m = 600$, gives

$$a(s,.34,.43) = t(s)(1.2647 - 0.00003s). \qquad (5.7)$$

Equation (5.7) was used to calculate values analogous to those presented in Tables II-15 and II-16. The results of those calculations are given in Table II-18.

Table II-18. Results of Projecting the DDESS 1999 SAT Scores to the Nation's Participation Rate

	Math	Verbal	Total
Reported 1999 Average SAT Scores for DDESS at its 34% Participation Rate	474	483	—
Calculated 1999 Average SAT Scores for DDESS at its 34% Participation Rate	473.65	482.57	956.22
Calculated 1999 Average SAT Scores for DDESS Projected to a 43% Participation Rate	472.59	481.44	954.03
Calculated 1999 Average SAT Scores for the Nation at its 43% Participation Rate	511.68	505.70	1017.38
Reported 1999 Average SAT Scores for the Nation at its 43% Participation Rate	511	505	1017

D. COMPARATIVE RESULTS FOR SEVERAL LOCAL AREAS

1. CTBS Results for Two School Districts in North Dakota

The two school districts in North Dakota with the largest percentages of children of military families are Grand Forks School District and Minot School District. Grand Forks has an enrollment of 8,559 students and Minot has an enrollment of 7,493 students. Grand Forks has 10.2 percent children from military families and Minot has 22.5 percent children from military families. Grand Forks School District educates children whose families are associated with Grand Forks Air Force Base and Minot School District educates children whose families are associated with Minot Air Force Base.

Data were obtained from the North Dakota Department of Education on 1999 CTBS percentile scores for three subjects—reading, language and mathematics—for grades 4, 6, 8 and 10.[10] 1999 CTBS percentile scores for DoDDS and DDESS were available. Table II-19 contains data on these CTBS scores.

10 Source: North Dakota Department of Education web site www.dpi.state.nd.us

Table II-19. CTBS Comparison with Grand Forks and Minot

	Grade 4	Grade 6	Grade 8	Grade 10	Average
Reading					
Grand Forks	66	66	74	74	70
Minot	69	66	67	76	70
DoDDS	68	65	65	72	68
DDESS	66	62	62	68	65
Language					
Grand Forks	66	69	72	73	70
Minot	67	68	65	73	68
DoDDS	66	69	69	72	69
DDESS	63	66	66	66	65
Mathematics					
Grand Forks	65	69	70	76	70
Minot	68	66	67	78	70
DoDDS	64	65	64	70	66
DDESS	64	64	59	64	63

The scores of Grand Forks and Minot are almost always higher than the scores of DoDDS and DDESS. The differences are captured in the averages presented in the rightmost column of Table II-19. Grand Forks has higher average scores than DoDDS and DDESS for all subjects. Minot has higher average scores than DoDDS for reading and mathematics and a lower average score than DoDDS for language. Minot has higher average scores than DDESS for all subjects.

2. CTBS Results for Two School Districts in Missouri

The two school districts in Missouri with the highest percentages of children of military families are Waynesville (58 percent) and Knob Noster (57 percent). Waynesville includes Fort Leonard Wood and Knob Noster includes Whiteman Air Force Base.

Data were obtained from the State of Missouri for these two districts on performance on the CTBS for three tests also taken by DoDEA—Mathematics, Science and Social Studies (partial data).[11] These data are for three grades, not the same in all cases, beginning with grade 3 and ending with grade 11. Since DoDEA administers these tests (as well as Reading and Language Arts) for all grades 3 through 11 there were comparable data for DoDEA.

Table II-20 presents data on the three scores for Waynesville and Knob Noster and for DoDDS and DDESS. Waynesville and Knob Noster scores are usually in the 60s, as are DoDDS

11 Source: 1999 Missouri School Directory web site www.dese.state.mo.us.

and DDESS. The scores are much higher than the national average of 50. Note that statewide Missouri averages are slightly below 60.

Table II-20. CTBS Comparison with Waynesville and Knob Noster

Mathematics	**Grade 4**	**Grade 8**	**Grade 10**	**Average**
Wynesville	66	75	70	70
Knob Noster	64	59	59	61
DoDDS	64	64	70	66
DDESS	64	59	64	62
Science	**Grade 3**	**Grade 7**	**Grade 10**	**Average**
Waynesville	74	68	65	69
Knob Noster	74	65	61	67
DoDDS	57	63	68	63
DDESS	64	59	64	62
Social Studies	**Grade 4**	**Grade 8**	**Grade 11**	**Average**
Waynesville	57	—	—	57
Knob Noster	65	69	51	62
DoDDS	67	66	74	69
DDESS	63	61	70	65

Children from Fort Leonard Wood and Whiteman Air Force Base can thus expect to receive an education which is about the same as they are receiving in the public schools if they are transferred to a base served by DoDDS or DDESS.

3. CTBS Results for Two School Districts in Nevada

The two school districts in Nevada with the largest percentages of children of military families are Churchill School District and Clark School District. Churchill has an enrollment of 4,766 students and Clark has an enrollment of 190,822 students. Churchill has 13.4 percent children from military families and Clark has 2.1 percent children from military families. Churchill is associated with Fallon Air Force Base and Clark is associated with Nellis Air Force Base.

Data were obtained from the Nevada Department of Education on 1998 CTBS percentile scores for four subjects—reading, language, mathematics and science for grades 4, 8 and 10.[12] 1998 CTBS percentile scores for DoDDS and DDESS were not available so their 1999 scores were used as an approximation. Table II-21 contains data on these CTBS scores.

12 Source: Nevada Department of Education web site www.nsn.k12.nv.us

Table II-21. CTBS Comparison with Churchill and Clark

	Grade 4	Grade 8	Grade 10	Average
Reading				
Churchill	49	56	61	55
Clark	48	52	50	50
DoDDS	68	65	72	68
DDESS	66	62	68	65
Language				
Churchill	44	51	58	51
Clark	53	49	55	52
DoDDS	66	69	72	69
DDESS	63	66	66	65
Mathematics				
Churchill	52	51	53	52
Clark	56	49	51	52
DoDDS	64	64	70	66
DDESS	64	59	64	62
Science				
Churchill	56	54	64	58
Clark	50	49	56	52
DoDDS	65	67	68	67
DDESS	67	67	64	66

Of the 12 test results in Table II-21, Churchill and Clark always have lower scores than DoDDS and DDESS except for one tie (Churchill and DDESS, science, grade 10).

4. CTBS Results for Two School Districts in Kentucky

The two school districts in Kentucky with the largest percentages of children of military families are Hardin County School District and Christian County School District. Hardin County School District has an enrollment of 12,073 students and Christian County School District has an enrollment of 8,007 students, according to the National Military Impacted Schools Association. Hardin County School District has 8.7 percent children from military families and Christian County School District has 8.5 percent children from military families. Hardin County School District educates children whose families are associated with Fort Knox and Christian County School District educates children whose families are associated with Fort Campbell.

Data were obtained from the State of Kentucky on 1999 CTBS percentile scores for three subjects—reading, language and mathematics—for grades 3, 6 and 9.[13] 1999 CTBS percentile scores for DoDDS and DDESS were available. Table II-22 contains data on these CTBS scores.

Table II-22. CTBS Comparison with Hardin and Christian

	Grade 3	Grade 6	Grade 9	Average
Reading				
Hardin County	48	50	50	49
Christian County	44	46	43	44
DoDDS	58	65	71	65
DDESS	62	62	67	64
Language				
Hardin County	47	48	45	47
Christian County	42	44	43	43
DoDDS	61	69	71	67
DDESS	62	66	67	65
Mathematics				
Hardin County	50	50	42	47
Christian County	42	42	35	40
DoDDS	58	65	66	63
DDESS	63	64	59	62

The scores of Hardin County School District and Christian County School District are always much lower than the scores of DoDDS and DDESS. The differences are captured in the averages presented in the rightmost column of Table II-22.

5. SAT Results for the Washington Metropolitan Area

Following are 1999 SAT scores and participation rates for selected school districts in the Washington metropolitan area[14] as well as for DoDDS and DDESS.

---

13 Source: Spring 1999 Kentucky School and District Results; CTBS/5 Survey Edition

14 Source: FY 2000 Metropolitan Area Boards of Education Guide, Produced by Fairfax County Public Schools, October 1999.

School District	SAT Verbal	SAT Math	SAT Total	% Participation
Montgomery County, MD	540	556	1096	79
Prince George's County, MD	499	440	939	53
Fairfax County, VA	541	553	1094	84
Prince William County, VA	519	505	1024	61
Washington, DC	494	478	972	77
DoDDS	506	501	1007	63
DDESS	483	474	957	34

Montgomery County, Maryland and Fairfax County, Virginia score significantly higher and have higher participation rates than DoDDS and DDESS. Prince George's County, Maryland scores lower than DoDDS and DDESS; its participation rate is lower than DoDDS and higher than DDESS. Prince William County, Virginia scores higher than DoDDS and DDESS; its participation rate is lower than DoDDS and higher than DDESS. Washington, DC scores lower than DoDDS and higher than DDESS. Its participation rate is higher than both.

Montgomery County, Maryland and Fairfax County, Virginia are two examples of affluent counties that have SAT scores much higher than those of DoDDS and DDESS. Recall that the CTBS scores of Montgomery County are also higher than DoDDS and DDESS. Fairfax County does not administer the CTBS.

E. INCONSISTENCY AMONG NAEP, CTBS, AND SAT RESULTS

NAEP results in DoDDS and DDESS are well above average in reading and writing, about average in mathematics and above average in science. CTBS results are in the middle to high 60^{th} percentiles for DoDDS and in the low to middle 60^{th} percentiles for DDESS. SAT results, however, are average for DoDDS and far below average for DDESS.

What is the source of this inconsistency?

1. Aptitude Versus Achievement

It is possible that the SAT scores measure intellectual aptitude for college, that DoDDS students taking the test are of average aptitude among college-bound students, and that DDESS students taking the test are of below-average aptitude among college-bound students.

Coupled with this it is also possible that DoDDS and DDESS teach students very well in the sense of enabling mastery of the material taught at various grade levels as measured by the NAEP and CTBS achievement tests. The NAEP results are particularly persuasive in this regard. Additionally, the CTBS scores for DoDDS and DDESS are far higher than in many states and are comparable with the CTBS scores attained by students in good school districts in selected states.

2. The Effect of the Antilles Schools

The low DDESS score might be due to the fact that, in the Antilles, the first language of many children is Spanish. The effect of the Antilles scores is illustrated by the following. There are eight high schools in DDESS. Data in the DDESS 1997-1998 Accountability Report give that the average SAT score for Camp Lejeune High School, Guam High School and Quantico Middle/High School was 1028. There were no data for Fort Knox and Fort Campbell (the ACT was reported). Adding data for the three high schools in the Antilles to the data for the three previous high schools reduces the average DDESS SAT score to 948.

3. A General Explanation

In the NAEP, on a national basis, DoDDS and DDESS perform better than most states in grade 4 reading and better than almost all states in Grade 8 reading and writing. DoDDS and DDESS perform about average in Grade 4 and Grade 8 mathematics and above average in Grade 8 science.

On the CTBS/Terra Nova DoDDS consistently scores in the middle to high 60th percentiles and DDESS consistently scores in the low to middle 60th percentiles. In DoDDS and DDESS CTBS/Terra Nova performance tends to be higher in secondary school.

In the performance of minority students and in the performance of students in the lower percentiles, DoDDS and DDESS tend to do much better than the national average in the NAEP.

These three results, taken together, indicate that DoDEA may not be dealing with a proportional number of students who are representative of the bottom levels of the nation's students with respect to economic and cultural background. Instead, the DoDEA test scores might be comparable with those attained by the portion of nation's students whose economic and cultural backgrounds are equivalent to those of DoDEA students, which may be much higher, on average, than for the nation's students as a whole.

Also confirmatory of this explanation is the observation that NAEP test scores tend to rise with grade level, as do the Terra/Nova percentiles in secondary school. This may correlate with the relative economic and cultural levels of the parents of these children. Higher-ranking officer and enlisted personnel have been competitively selected from the larger pool of candidates and so should have stronger economic and cultural characteristics.

Finally, the seeming anomaly of the SAT test results may have a simple explanation. Setting aside the DDESS results due to the effects of the Antilles, the DoDDS participation rate of 63 percent may be equivalent to the nationwide participation rate of 43 percent if one-quarter of the nation's seniors in high school have significantly lower levels of economic and cultural advantages than the seniors in DoDDS. A 63 percent participation rate of a top 75 percent cohort is 47 percent, quite similar to the national participation rate of 43 percent.

These concepts are discussed further in Section F, which follows.

F. A THEORETICAL MODEL OF DODEA STUDENT SCORES

Throughout this section, let "DoDEA students" mean the students attending DoDEA schools, and let "DoDEA system" mean all of DoDEA except for these students.

1. The Underlying Hypothesis

The basic hypothesis here is that DoDEA students are significantly above average in a particular way, where "above average" means above the national average in intelligence and in academic test-taking ability. (The way that they are above average is described below.) Being above average, they should be expected to achieve above average scores on tests taken by all, or by a representative sample, of DoDEA students. In particular, the above average scores that they achieved on NAEP and CTBS tests may be: 1) partially due to their above average intelligence and academic test-taking ability, and partially due to the DoDEA system which, accordingly, would be better than the average national public education system, or 2) entirely due to their above average intelligence and academic test-taking ability, and so the DoDEA system may be about equal to the average national public education system, or 3) less than they could have achieved, given their above average intelligence and academic test-taking ability, had not a below average DoDEA system been holding them back. This hypothesis implies that one of these three cases must hold, but it does not indicate which one holds.

The hypothesized way that the DoDEA students are above average is as follows. On an absolute basis, the bottom end (e.g., the number of students who score below a given value on a given test) of DoDEA students is, on average, a lower percentage of the total number of DoDEA students than is the corresponding bottom-end percentage for the nation's public school students. Conversely, on a conditional basis, the scores of the students who are in the top end have about the same distribution for DoDEA as they have for the nation's public school students.

If the number of bottom-end DoDEA students is a lower percentage of the total number of DoDEA students than is the corresponding percentage for the nation's public school students, then the number of top-end DoDEA students is a higher percentage of the total number of DoDEA students than is the corresponding percentage for the nation's public school students. Thus, this hypothesis (at least partially) explains why the DoDEA SAT participation rate is higher than the national SAT participation rate. However, given that a DoDEA student is in the top end (and so is likely to take the SAT), this hypothesis says that the DoDEA student is, on average, about the same as the national average top-end student (who is also likely to take the SAT). Thus, this hypothesis also (at least partially) explains why DoDDS students' scores on the SAT are roughly equivalent to the national average even though their NAEP and CTBS scores are much higher. (DDESS SAT scores are lower than the national average due to the Antilles schools.)

2. Theoretical Rationale for this Hypothesis

A theoretical rationale for this hypothesis is as follows.

First, this rationale argues that a child's intelligence (and academic test-taking ability) is correlated with his or her parents' (or guardian's) income and intelligence at the low end of parents' income and intelligence levels. Of course, it may be correlated with parents' income and intelligence throughout the spectrum of income and intelligence levels; however, while this broader correlation is sufficient, it is not necessary here. The argument needed here is only that this correlation must hold at the low end. Thus, for example, a statistical test that covered the whole spectrum of parents' income and intelligence levels and found little or no correlation with their children's intelligence would not necessarily invalidate this argument, while one that covered this spectrum and found a high correlation except at the low end might do so, depending on whether anything was able to be concluded about the low end.

Second, this rationale argues that, with the exception of a statistically insignificant number of cases, the lowest compensation level of the sponsors of DoDEA fourth grade students is some pay-grade level, say that of that of an E-5 with 6 years of service. That is, the DoDEA students' NAEP and CTBS scores are essentially achieved by children who have a parent or guardian who is at least an E-5 over 6. [This "E-5 over 6" level is a judgmental estimate. We have no data to support it, but it seems like a reasonably conservative estimate.] If this is the case, then, in terms of intelligence, this parent or guardian must have passed the tests necessary to enlist in a military service, and must have performed well enough to have been promoted four times and have been accepted for re-enlistment. In terms of income, the average annual regular military compensation of an E-5 over 6 in 1999 was over $31,000. (Regular military compensation includes basic pay, average basic allowance for housing, basic allowance for subsistence, and an estimate of the federal tax advantage of these allowances. It does not include state or local tax advantages, nor does it include any benefits, such as medical benefits, retirement, or annual leave. It does not include an estimate of any other income, such as spousal income, that a household might have.)

Third, this rationale argues that there are a statistically significant number of students in the nation's public schools whose parents or guardian do not have the intelligence level necessary to pass the tests to enlist in a military service and then to perform well enough to be promoted four times and to be accepted for re-enlistment. It is likewise argued that a statistically significant number of students in the nation's public schools come from households whose annual income is significantly less than $31,000.

This third argument means that the students whose parents' intelligence levels or income are significantly below that of an E-5 over 6 form a statistically significant set of students who are in the nation's public schools and so who are considered in national average test scores. By the first argument above, these students, on average, lower the nation's average test scores. By the second argument, the corresponding set of students in DoDEA schools is not statistically significant, and so such students do not substantially affect DoDEA test scores.

Fourth, this rationale argues that this is the only statistically significant difference in intelligence and test-taking ability between DoDEA students and the nation's public school students. In particular, while there are some civilian parents whose incomes or intelligence levels are greater than those of any military parent, the number of children of such parents who are in public schools is arguably an insignificant percentage of the total number of public school students.

Finally, this rationale argues that, conditional on their intelligence being over some threshold, the statistical characteristics of the intelligence of students are essentially independent of their parents' intelligence levels or incomes. (This argument is only needed to explain why the DoDDS SAT score is about the same as the national average. In particular, the average total

DoDDS SAT score over 1997, 1998, and 1999 is 1014, while the national average over these three years is 1017.)

3. The Model

For any set of students, S, and any academic test, t, let:

$s_t(x,S)$ = the probability that a randomly (i.e., uniformly and independently) chosen student from S would score at least x on test t if that student were to take that test,

$a_t(S)$ = the (expected) average score that would be achieved by the students in S on test t if all of those students were to take that test.

For any given school grade, let

M = the set of DoDEA (military) students in that grade,

N = the set of the nation's public school students in that grade,

L = the subset of the students in N whose parents' or guardian's intelligence levels or household incomes are below that of an E-5 over 6,

K = the subset of the students in N who are not in L,

O = a subset of K that consists on a statistically insignificant number of students,

H = the subset of the students in K who are not in O.

Note that $N = L \cup H \cup O$, and, because O is insignificantly small,

$$s_t(x,N) \cong s_t(x, L \cup H) \text{ for all relevant } x\text{, and}$$

$$a_t(N) \cong a_t(L \cup H).$$

The non-SAT portion of the model can be stated as follows.

Hypothesis: There exists a set O as described above such that

$$s_t(x,M) > s_t(x,L)$$

and

$$s_t(x,M) = s_t(x,H)$$

for all relevant x. Therefore:

$$a_t(M) > a_t(L)$$

and

$$a_t(M) = a_t(H),$$

so

$$a_t(M) > a_t(L \cup H) \cong a_t(N).$$

To consider the SAT portion of the model, for any set of students, S, let

$p(x,S)$ = the probability that a student in S, who would (on average) score x on the SAT if that student were to take the SAT, does take the SAT. It is assumed that

4. Some Practical Considerations

First, other than giving the average annual regular military compensation of an E-5 over 6, no data are presented here to support the hypothesis above.

Second, even if this hypothesis is accepted as being valid (perhaps because educational researchers have already established the first part of the theoretical rationale above, and the rest is "obviously true"), data are still needed to determine which of the three cases described in the first paragraph of Section 2 holds. One path towards obtaining such data is to ignore household income, and, instead, to concentrate on estimating the distribution of the intelligence (somehow defined) of the parents of the nation's public school students, of DoDEA's student's parents, and on determining the correlation of such intelligence to student's scores. This path has many risks.

A potentially less risky path is to ignore parent's intelligence (except indirectly through income) and, instead, to concentrate on estimating the distribution of the household income of the nation's public school students, of DoDEA's students, and on determining the correlation of household income to students' scores.

III. RESOURCES

Three primary measures of resources are considered in this chapter. They are expenditures per pupil, class size, and teacher education.

Data on expenditures per pupil for states are from the National Center for Educational Statistics. Data for the District of Columbia are estimated. The DoDEA provided DoDDS and DDESS data. Data on class size and on teacher education are from the NAEP; this is the only national source that includes consistent data on these items for states, the District of Columbia, DoDDS and DDESS.

The purpose of presenting these data is to span the range of the most crucial inputs.[1] Expenditures are to some extent fungible, and in the aggregate should be the best measure of resources. Comparative class size from the NAEP should be correct with respect to comparisons across states, and it is one of the most widely-accepted measures of inputs to school quality. Teacher qualifications are another of the most widely accepted measures of inputs to school quality. They reflect the commitment of resources as well as the level of training of the teachers. One caveat, however, is that research has shown that student performance is correlated with teacher intelligence as measured by test scores.[2] Advanced degrees and certificates in teaching methodology are not necessarily related to student achievement.

A. EXPENDITURES PER STUDENT

Table III-1 presents average expenditures per student for each of the fifty states, the District of Columbia, DoDDS and DDESS, ranked from highest to lowest.[3]

---

1 For a discussion of the most important inputs see Eric A. Hanushek and Others, *Making Schools Work: Improving Performance and Controlling Costs*, The Brookings Institution, 1994 and Gary Burtless (Editor), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, The* Brookings Institution, 1996.

2 See Marci Kanstoroom and Chester E. Finn, Jr. (Editors), *Better Teachers, Better Schools*, Thomas B. Fordham Foundation, July 1999.

3 U.S. Department of Education, National Center for Educational Statistics, Common Core of Data. Preliminary student membership and number of teachers, and estimates of revenues, expenditures, and pupil/teacher ratio for public elementary and secondary schools, by state, for grades prekindergarten through 12. School year 1998-1999 data published in Education Week, Volume 18, January 13, 2000.

Table III-1. Expenditures per Pupil (School Year 1998-1999)

State	1999 Enrollment (thousands)	1999 Expenditures (millions)	1999 Expenditures per Student
DoDDS	**76**	**780**	**10,215**
New Jersey	1,300	13,100	10,077
DDESS	**36**	**348**	**9,734**
Alaska	134	1,200	9,500
Connecticut	545	5,100	9,358
New York	2,900	26,200	9,034
Colorado	699	348	8,500
District of Columbia	80	656	8,200
Massachusetts	964	7,800	8,091
Rhode Island	154	1,200	7,792
Pennsylvania	1,800	14,000	7,778
Wisconsin	888	6,900	7,770
Delaware	113	841	7,442
New Hampshire	195	1,400	7,179
Illinois	2,000	14,300	7,150
Wyoming	94	670	7,128
Michigan	1,700	12,100	7,118
West Virginia	296	2,100	7,095
Maryland	837	5,800	6,930
Minnesota	858	5,900	6,876
Maine	220	1,500	6,818
Vermont	105	704	6,705
Indiana	989	6,500	6,572
Oregon	543	3,500	6,446
Nation	**46,100**	**296,642**	**6,435**
Ohio	1,800	11,400	6,333
Washington	1,000	6,300	6,300
Nebraska	291	1,800	6,186
Virginia	1,100	6,800	6,182
Iowa	503	3,100	6,163
Kentucky	646	3,800	5,882
California	5,800	34,100	5,879
Montana	161	945	5,870
New Mexico	329	1,900	5,775
North Carolina	1,200	6,900	5,750
Kansas	470	2,700	5,745
Louisiana	754	4,300	5,703
Florida	2,300	13,000	5,652
South Carolina	644	3,600	5,590
Georgia	1,400	7,800	5,571
Missouri	921	5,100	5,537
North Dakota	114	627	5,500
Texas	3,900	21,400	5,487
Nevada	311	1,700	5,466
Hawaii	187	1,000	5,348
Oklahoma	627	3,200	5,104
South Dakota	142	699	4,923
Alabama	759	3,600	4,743
Tennessee	909	4,300	4,730
Arkansas	456	2,100	4,605
Arizona	829	3,800	4,584
Mississippi	502	2,300	4,582
Idaho	245	1,100	4,490
Utah	447	1,800	4,027

DoDDS has the highest expenditures and DDESS is third. Their expenditures are about the same as those of the other highest-spending states, which are mostly in the Northeast. Their expenditures per student are more than twice as large as those of the bottom few states; these states are mostly in the South and in the West. There is no adjustment for cost of living in the data, so it is conceivable that the costs of education inputs are somewhat less in the South and in the West, but assumably not enough less to compensate for the differences. Also, the aggregation of school districts into statewide averages may hide some factors related to the relative costs of education inputs.

There is some argument that DoDDS expenditures, which include allowances for cost of living overseas and permanent change of station, should be adjusted downward to be comparable with expenditures in the United States. If those allowances are removed for DoDDS, the cost per pupil is calculated at $8,706, which is still at the high end of the scale. DDESS expenditures include cost of living overseas and change of station allowances for Guam and the Antilles. If these costs are removed, the cost per pupil for DDESS is $9,646. Since DoD has military dependents living overseas and their children are being educated in DoD schools overseas, it is quite reasonable that the cost of living allowances and change of station costs should be included in cost per pupil analyses since they are part of the operating expenditures of the schools.

B. CLASS SIZE

Data are given in Table III-2 for class size in Grade 4 Reading.[4] The average percentage of classes with 25 or fewer students is presented for 36 states, the District of Columbia, DoDDS and DDESS.

Counting the District of Columbia as a state, there are seven states with a higher percentage of classes with 25 or fewer students than DoDDS and 19 states with a higher such percentage than DDESS out of the 37 states listed. There are 29 states with a lower such percentage than DoDDS and 17 states with a lower such percentage than DDESS out of these states. Thus, DDESS is near the top and DoDDS is about at the level of the average state. However, the DoDDS percentage is 75 percent as compared to the national average of 64 percent. Several very large states, and California in particular, have a small percentage of classes with 25 or fewer students.

---

4 NCES Web Site at http: //nces.ed.gov/nationsreportcard/TABLES/REA1998/XS/Gr04/TCH/XSR13061.HTM

Table III-2. Grade 4 Reading Class Size Data

State	1999 Enrollment (thousands)	% of 1998 Classes With Less Than 26 Students
Maine	220	98
Texas	3,900	97
Kansas	470	94
New Mexico	329	94
Oklahoma	627	93
Connecticut	545	92
Wyoming	94	90
West Virginia	296	89
DDESS	**36**	**89**
Rhode Island	154	84
Massachusetts	964	82
Virginia	1,100	82
New Hampshire	195	81
South Carolina	644	81
Tennessee	909	79
Wisconsin	888	79
Arkansas	456	78
District of Columbia	80	78
Missouri	921	78
Montana	161	77
Alabama	759	75
DoDDS	**76**	**75**
Mississippi	502	74
Kentucky	646	73
Louisiana	754	69
Georgia	1,400	67
Minnesota	858	66
Nation	**46,100**	**64**
North Carolina	1,200	63
New York	2,900	62
Colorado	699	60
Delaware	113	56
Maryland	837	52
Oregon	543	52
Washington	1,000	51
Arizona	829	49
Florida	2,300	46
Hawaii	187	43
Nevada	311	35
California	5,800	21

Analogous data are given in Table III-3 for class size in Grade 8 Reading for 35 states, the District of Columbia, DoDDS and DDESS.[5] This time, DDESS is not bettered by any state and DoDDS is bettered by three states. Both are far above the national average.

A caveat is that class size data are not reported for a fairly large number of states.

C. TEACHER QUALIFICATIONS

Data are given in Table III-4 for percent of teachers with advanced degrees (final degrees or certificates higher than the bachelor's degree) as recorded in the NAEP Grade 4 Reading test.[6] These data are given for 39 states, the District of Columbia, DoDDS and DDESS.

DoDDS and DDESS have among the highest percentages of teachers with advanced degrees.

Data are given in Table III-5 for percent of teachers with advanced degrees (final degrees or certificates higher than the bachelor's degree) as recorded in the NAEP Grade 8 Reading Test.[7] These data are given for 36 states, the District of Columbia, DoDDS and DDESS.

DDESS has a much higher percent of teachers with advanced degrees than does any state in the nation, and the percent of teachers with advanced degrees in DoDDS is exceeded by only two states.

Teacher education in DoDDS and DDESS is far greater than in the nation as a whole.

A caveat is that teacher education data are not reported for a fairly large number of states.

---

5 NCES Web Site at http: //nces.ed.gov/nationsreportcard/TABLES/REA1098/XS/GR08/TCH/XSR23144.HTM

6 NCES Web Site at http: //nces.ed.gov/nationsreportcard/TABLES/REA1098/XS/GR08/TCH/XSR113014.HTM

7 NCES Web Site at http: //nces.ed.gov/nationsreportcard/TABLES/REA1098/XS/GR08/TCH/XSR23012.HTM

Table III-3. Grade 8 Reading Class Size Data

State	1999 Enrollment (thousands)	% of 1998 Classes With Less Than 26 Students
DDESS	**36**	**95**
Maine	220	92
Wyoming	94	84
Arkansas	456	84
DoDDS	**76**	**82**
Kansas	470	80
Connecticut	545	80
New Mexico	329	77
West Virginia	296	76
Massachusetts	964	76
Texas	3,900	75
Oklahoma	627	72
Rhode Island	154	72
Missouri	921	72
Mississippi	502	71
Montana	161	69
South Carolina	644	67
Virginia	1,100	66
Alabama	759	66
Wisconsin	888	65
Louisiana	754	63
District of Columbia	80	60
North Carolina	1,200	60
New York	2,900	58
Nation	**46,100**	**57**
Georgia	1,400	56
Tennessee	909	54
Hawaii	187	54
Kentucky	646	49
Colorado	699	46
Delaware	113	46
Minnesota	858	45
Oregon	543	44
Arizona	829	44
Maryland	837	36
Washington	1,000	35
Florida	2,300	33
Nevada	311	33
California	5,800	19

Table III-4. Grade 4 Reading Teacher Data

State	1999 Enrollment (thousands)	% of 1998 Teachers With More Than a Bachelor's Degree
Connecticut	545	80
New York	2,900	75
Kentucky	646	73
DoDDS	**80**	**65**
Rhode Island	154	64
DDESS	**30**	**63**
South Carolina	644	60
Alabama	759	59
Massachusetts	964	57
Colorado	699	54
West Virginia	296	53
Michigan	1,700	52
Washington	1,000	51
Maryland	837	50
Tennessee	909	49
Georgia	1,400	48
Wisconsin	888	48
Arizona	829	47
Delaware	113	46
District of Columbia	80	46
Minnesota	858	46
Missouri	921	44
Nation	**46,100**	**43**
Hawaii	187	42
New Mexico	329	41
Nevada	311	39
Kansas	470	38
New Hampshire	195	36
Oklahoma	627	36
Oregon	543	36
Florida	2,300	33
Mississippi	502	33
Virginia	1,100	33
Maine	220	32
California	5,800	31
Montana	161	31
Arkansas	456	28
Louisiana	754	28
North Carolina	1,200	27
Iowa	503	22
Wyoming	94	22
Texas	3,900	21
Utah	447	20

Table III-5. Grade 8 Reading Teacher Data

State	1999 Enrollment (thousands)	% of 1998 Teachers With More Than a Bachelor's Degree
DDESS	**36**	**95**
Connecticut	545	84
New York	2,900	80
DoDDS	**76**	**80**
Kentucky	646	71
District of Columbia	80	69
Massachusetts	964	66
Alabama	759	63
West Virginia	296	63
South Carolina	644	55
Tennessee	909	55
Washington	1,000	55
Colorado	699	53
Georgia	1,400	53
Oregon	543	53
Rhode Island	154	53
Arizona	829	49
Delaware	113	48
Maryland	837	48
Missouri	921	48
Nation	**46,100**	**47**
Wisconsin	888	46
Virginia	1,100	44
New Mexico	329	41
California	5,800	40
Kansas	470	40
Hawaii	187	39
Minnesota	858	39
Louisiana	754	38
Montana	161	38
Nevada	311	38
Maine	220	37
Florida	2,300	36
Mississippi	502	36
North Carolina	1,200	34
Oklahoma	627	33
Utah	447	32
Arkansas	456	30
Wyoming	94	30
Texas	3,900	27

D. WASHINGTON METROPOLITAN AREA COMPARISONS

The Metropolitan Area Boards of Education Guide provides data with common definitions for the Washington area school districts in Maryland and Virginia. The data are for 1999.[8]

Table III-6 presents a wide variety of comparative data for Montgomery County, Maryland, Fairfax County, Virginia, Prince George's County, Maryland, Prince William County, Virginia, DoDEA, DoDDS and DDESS. Data for the four counties were available from the guide. Data for DoDEA, DoDDS and DDESS were furnished by DoDEA.

DoDEA resources seem to be generally consistent with these other school systems.

Cost per pupil for DoDEA is much higher. This is partially attributable to the overseas living allowances of DoDDS teachers.

Teacher staffing ratios are within the range of those of the other school systems.

Teacher salaries for experienced teachers are higher in DoDDS and lower in DDESS. Montgomery County and Fairfax County pay higher salaries than Prince George's County and Prince William County. DoDDS salaries are comparable with the former two school systems while DDESS salaries are comparable with the latter two school systems.

Teacher salaries for new teachers are lower in DoDDS and DDESS than in the other school systems.

Salary and benefits assuming a $45,000 salary are high in DoDDS due to allowances. DDESS adjusted salaries are lower than the others. Both DoDDS and DDESS have much lower costs for medical insurance than do the other school systems.

---

8 FY 2000 Metropolitan Area Boards of Education Guide, Fairfax County Public Schools, October 1999

Table III-6. Comparative Data for Washington Area School Districts

Category	Mont-gomery	Fairfax	Prince George's	Prince William	DoDEA	DoDDS	DDESS
Enrollment 1988	125,035	148,036	128,369	50,166	110,900		
Enrollment 1999	127,918	151,979	130,972	52,948	112,145[1]	76,366	35,779
Number of Schools	185	205	185	68	226		
Operating Fund, millions	$1,076.5	$1,180.4	$844.9	$329.75	$1,224,352	$780,060	$348,29 0
Construction Fund, millions	$110.5	$139.5	$73.65	$77.6	$79.8[2]		
Cost per Pupil	$8,256	$7,731	$6,166	$6,339	$10,900[3]	$10,215	$9,734
Teachers Staffing Ratio (Elementary)	24.1:1	24.5:1 to 26.5:1	25.0:1	24.6:1	23:1 grades 1-3[4] 25:1 gr. 4-6		
Teachers Staffing Ratio (Intermed/Middle)	25.0-1	27.0:1	30.0:1	19.8:1		23:1	
Teachers Staffing Ratio (Secondary)	30.1:1	28.0:1	21.4:1	20.0:1		21-25:1	
Teacher Salary, 9th Step w/M.A.	$46,667	$45,953	$40.176	$41,828		$48,937 (FY 99)[5]	$40,688 (FY 99)[6]
Teacher Salary w/B.A., Beginning	$31,669	$30,761	$30,577	$30,328		$27,945 (FY 99)[5]	$28,423 (FY 99)[6]
Teacher Salary, Average	$51,267	$48,497	$42,469	$41,389		$48,937 (FY 99)[5]	
Salary & Benefits for a $45,000 Salary	$62,649	$60,153	$58,393	$60,267		$62,173	$52,109
Average Class Size (Elementary)	14.60	15.20	N/A	16.70	17[7]	23 in gr. 1-3 25 in gr. 4-6	15[7]

1. Does not include 265 students at a junior college that closed in 1999.
2. Includes $78,408,000 for military construction (costs greater that $500,000) and $1,428,000 for procurement (purchases greater than $100,000).
3. Includes headquarters/consolidated costs, and overseas housing allowance, repair and maintenance costs, military base support services, dorm expenses, and change of station costs.
4. A reduced teacher-staffing ratio of 18.1 for grades 1-3 is to be completed by FY 05.
5. Salary tables are under negotiations.
6. A 3% increase is expected in 2000 for state side schools.
7. When all teachers scale positions are included.

IV. COLLEGE ATTENDANCE

The purpose of this chapter is to examine the quality of the colleges attended by DoDEA graduates.

The first analysis addresses college attendance at the high end. Are DoDEA graduates over-represented, proportionally represented, or under-represented at the top universities and colleges?

The second analysis addresses college attendance across-the-board, focusing on the question of whether the colleges attended by DoDEA graduates are better than, equal to, or worse than would be expected given the SAT scores of the students.

The former analysis deals with a relatively small proportion of the students. The latter analysis considers all of the college-bound students.

A. ATTENDANCE AT TOP-TIER UNIVERSITIES AND COLLEGES

1. Data

Post-secondary plans of 1999 DoDEA high school graduates were as follows, according to information received from DoDEA:

4-year college / university	1697
2-year college / university	499
Subtotal	2196
Other	795
Total	2991

Of the 2196 students planning to attend 4-year or 2-year colleges, 1,654 reported the colleges they planned to attend. Of those 1,654 students, 142 planned to attend schools that are among the U.S. News top 50 national universities, and 9 planned to attend schools that are among the U.S. News top 40 national liberal arts colleges.

2. First Comparison—Percentage of Those Attending College Who Attend Top Colleges

Of the 1,654 students reporting their intentions, 151 planned to attend top universities and colleges. Thus, at least 8.9 percent of the college-bound students (151 of the 1697 college-bound students) planned to attend top universities and colleges.

In 1996 there were 8.8 million students in 4-year colleges (Source: Digest of Education Statistics 1998). Let us assume that this number has increased to 9.0 million in 1999. In 1999, there were 663,535 undergraduate students in the top 50 national universities and top 40 national liberal arts colleges. Thus, the attendance at top universities and colleges is about 7.4 percent of all 4-year colleges.

Some of the 663,535 students attending the top universities and colleges are foreign students. So 7.4 percent is an overestimate of the attendance of American students at these institutions.

3. Second Comparison—Percentage of Those Graduating from High School Who Attend Top Colleges

Of all DoDEA students graduating from high school in 1999, at least 151 of 2991, or at least 5.0 percent, planned to attend top universities and colleges.

In 1998, there were 2.8 million students in the high school graduating class (Source: Digest of Education Statistics 1998). Assume that this number has increased to 2.9 million in 1999 (about the same increase as in the previous year). Also assume that the number of freshman attending the top universities and colleges is one-fourth of the total number of undergraduates, 663,535, which is 165,884. Then the attendance at top universities and colleges would be 165,884 of 2.9 million, or 5.7 percent.

Some of the 663,535 students attending the top universities and colleges are foreign students. So 5.7 percent is an overestimate of the attendance of American students at these institutions.

4. Conclusions

Of the students planning to attend college, DoDEA students are overrepresented at top universities and colleges. Of the students who graduate from high school, DoDEA students may be underrepresented at top universities and colleges.

Concerning the first conclusion, the number of DoDEA graduates planning to attend top universities and colleges may be underestimated due to the sample of 1654 out of 2196.

Therefore, the percentage of DoDEA graduates planning to attend top universities and colleges is surely greater than the percent of the nation's graduates attending these schools.

Concerning the second conclusion, if the number of DoDEA graduates attending top universities and colleges is about as reported in the 1,654 observations, then the percentage of DoDEA graduates attending these colleges is smaller than the corresponding percentage of the nation's graduates. If the number of DoDEA graduates attending these colleges is more than as reported in those 1,654 observations, then the percentage of DoDEA graduates attending these colleges may be smaller or larger than the corresponding percentage of the nation's graduates.

B. ATTENDANCE AT ALL UNIVERSITIES AND COLLEGES

1. Introduction

The methodology of this analysis is to estimate the quality of a college based on the average SAT score of the students attending that college, and then to compare the quality of the colleges attended by DoDEA graduates with the average SAT score of the DoDEA graduates.

The quality of the colleges attended by last year's DoDEA graduates must be compared with an estimate of the SAT scores achieved by those graduates because the actual scores achieved by those graduates is not known.

One approach for doing this is to compare an estimate of the average SAT score of the colleges to be attended with the average SAT score of all DoDEA graduates who took the SAT last year.

A second approach is to first estimate the number of DoDEA graduates who are in their first year at colleges that report the average SAT score of their incoming students. Call this number *N*. Next, assume that the top *N* DoDEA graduates last year, in terms of their SAT scores, are the DoDEA students who are now attending those colleges. Then compare the estimate of the average SAT score of those colleges to the average SAT score of the top *N* DoDEA graduates.

Thus, the major difference between the two approaches is as follows. The first approach compares a college-based score to the average SAT score of all DoDEA graduates who took the SAT last year, while the second approach compares a college-based score to the average SAT score of the top *N* DoDEA graduates who took the SAT last year for some data-dependent value of *N*. The second approach is taken here.

2. Estimating *N*

As stated above, 2,196 of last year's DoDEA graduates planned to attend a 2-year or 4-year college. A detailed list, provided by DoDEA, gives (sometimes imprecise) names for the colleges and the number of DoDEA graduates planning to attend each of those colleges. However, that list includes only 1,654 such attendees, a discrepancy of 542 students. It will be seen that the result of the comparison described above depends on the data for the "missing" 542 students.

First, concerning the 1,654 students for whom colleges were given, note the following.

An initial sample examined data for 40 randomly selected students from that list. This sample yielded 23 students attending SAT-reporting 4-year colleges, 2 attending ACT-reporting 4-year colleges, 2 attending "non-SAT, non-ACT" 4-year colleges, 7 attending 2-year colleges, 3 attending overseas divisions of the University of Maryland, 2 attending unspecified divisions/campuses of that university, and 1 apparently incomplete entry (which is presumably not a SAT-reporting college). It was estimated that those 23 SAT-reporting 4-year colleges had an average SAT score of 1112 for the students they admitted last year.

A second sample examined data for 50 additional randomly selected students from that list. These additional samples yielded 31 students attending SAT-reporting 4-year colleges, 4 attending ACT-reporting 4-year colleges, 3 attending "non-SAT, non-ACT" 4-year colleges, 4 attending 2-year colleges, 3 attending overseas divisions of the University of Maryland, 3 attending unspecified divisions/campuses of that university, 1 attending a multi-campus system (SUNY), and 1 attending an acting school. It was estimated that these 31 SAT-reporting 4-year colleges had the same average SAT score of 1112 for the students they admitted last year.

The fact that both samples yielded the same average SAT score of 1112 is remarkable.

It is assumed that none of the 5 students in these samples who are attending unspecified divisions/campuses of the University of Maryland are attending any of that school's SAT-reporting campuses. The same assumption is made for the SUNY student. Combining these two sets of samples with this assumption gives that 54 of the 90 students sampled are attending SAT-reporting 4-year colleges whose average SAT score is 1112.

Accordingly, it is assumed here that

$$N = S \times (54 / 90) = 0.6S,$$

where S is the number of students on the detailed list (i.e., 1654) plus the number of the 542 missing students who are attending the same types of colleges in the same proportions as the students on the detailed list. In other words,

$$S = 1654 + X$$

where X is the number of DoDEA graduates who are not on the detailed list, but who are attending the same types of colleges in the same proportions as the graduates who are on that list.

Three alternative assumptions concerning X are considered here. Assumption 1, X is about equal to 0. Assumption 2, X is about equal to 542. Assumption 3, X is about half way between, i.e., about 271.

Before going on, it should be noted that the six students attending ACT-reporting colleges could be included by using equivalent SAT scores for those schools. Also, some of the other students could be included by estimating SAT scores for their schools. This inclusion would be appropriate if the DoDEA students attending those schools took the SAT even though they were not required to do so, and it is not appropriate otherwise. The latter case is assumed here.

Assumption 1: Assume that the detailed list is essentially complete as far as DoDEA graduates attending SAT-reporting 4-year colleges are concerned, so X is about 0. Then S is about 1654, so N is about 0.6×1654, which is about 992.

Assumption 2: Assume that the 542 missing students are attending about the same types of colleges in about the same proportions as the students represented on the detailed list are attending, so X is about 542. Then S is about 2196, so N is about 0.6×2196, which is about 1318.

Assumption 3: Assume that half of the 542 missing students are attending about the same types of colleges in about the same proportions as the students represented on the detailed list are attending, and half are not, so X is about 271. Then S is about 1925, so N is about 0.6×1925, which is about 1155.

3. Estimating the Average SAT Score of the Top *N* DoDEA Graduates

In its 1999 Graduating Senior DoDEA Students' SAT I Total Score Statistics, DoDEA lists total SAT scores for every 5th percentile from the 0th to the 100th percentile. By assuming linearity within these 5-percentile intervals, the average total SAT score of the top N DoDEA students who took the SAT last year can be estimated for any feasible value of N. For $N = 992$, this estimate is about 1144. For $N = 1318$, this estimate is about 1086. For $N = 1155$, this estimate is about 1115.

4. Conclusions

The assumptions and data described here lead to the following observations.

The number of last year's DoDEA graduates who entered SAT-reporting 4-year colleges is about 992 under Assumption 1, about 1318 under Assumption 2, and about 1155 under Assumption 3. The average SAT score of the top 992 DoDEA graduates last year is about 1144, for the top 1318 graduates it is about 1086, and for the top 1155 graduates it is about 1115.

The average SAT score of all of the incoming students at those SAT-reporting 4-year colleges last year was about 1112.

Accordingly, in terms of SAT scores, DoDEA graduates in 1999 would be attending lower-scoring colleges than their scores predict given Assumption 1 (if the students on the list were the top 1654 students). They would be attending higher-scoring colleges than their scores predict given Assumption 2 (if the 1654 students on the list were fully-representative of all of the students). And they would be attending colleges whose scores are equivalent to their scores given Assumption 3 (if the 1654 students on the list were semi-representative of all of the students).

Comparisons concerning the quality of the colleges attended by DoDEA graduates thus depend on the characteristics of the missing observations on where students attended college.

V. TEACHER QUALITY

This chapter examines the quality of the teachers in DoDEA compared to the quality of the teachers in public and private schools in the United States. It focuses on the intellectual quality of the teachers, on the assumption that brighter individuals make better teachers.

To motivate this approach, we observe that the Carnegie Forum on Education and the Economy reported the following:

"Teachers should have a good grasp of the ways in which all kinds of physical and social systems work; a feeling for what data are and the uses to which they should be put, an ability to help students see patterns of meaning where others see only confusion.... They must be able to learn all the time, as the knowledge required to do their work twists and turns with new challenges and the progress of science and technology.... We are describing people of substantial intellectual accomplishment."[1]

The basic intellectual capability of the teachers is, therefore, a critical input to student learning.

A. BACKGROUND

The research literature lends support to the proposition that instructors with stronger intellectual capabilities and academic backgrounds are, other things being equal, more effective teachers. From Ballou and Podgursky, the following research findings are extracted.[2]

Hanushek in 1971 investigated the relationship between the achievement of California third graders and the characteristics of their second and third grade teachers, including experience, hours of graduate education, and scores on a test of verbal ability. Of all teacher characteristics, scores on the test of verbal ability were the most important determinant of student learning.[3]

---

1 Carnegie Forum on Education and the Economy: Task Force on Teaching as a Profession, *A Nation Prepared: Teachers for the 21st Century*, New York: Carnegie Corporation, 1986, page 25.

2 Dale Ballou and Michael Podgursky, *Teacher Pay and Teacher Quality*, Kalamazoo, Michigan: W.E. Upjohn Institute, 1997, page 9.

3 Eric Hanushek, "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data," *American Economic Review*, Volume 61, Number 2, 1971, pages 280-288.

Winkler in 1975 found a positive association between test score gains and the "prestigiousness" of the teacher's undergraduate college. Prestigious institutions included Stanford and the University of California system; nonprestigious schools were represented primarily by the California state college system.[4]

Summers and Wolfe in 1977 found that student test score gains between third and sixth grade varied positively with the quality of their teacher's undergraduate college.[5]

Webster in 1988 found a significant positive correlation between teachers' scores on the Wesman Personnel Classification test (a test of verbal and quantitative ability) and the scores of middle school students on the Iowa Test of Basic Skills and of secondary school students on the Iowa Tests of Educational Development.[6]

Ehrenberg and Brewer in 1994 found a positive association between student test score gains from tenth to twelfth grades and the selectivity of the colleges attended by teachers at their school.[7]

Monk in 1994 found a strong positive association between the subject matter preparation (college course work) of high school mathematics and science teachers and their students' achievement test scores.[8]

The fact that researchers employing a variety of data sets and test instruments have found a positive association between teacher's tested ability and student learning attests to the robustness of this relationship.[9]

---

4 Donald R. Winkler, "Educational Achievement and School Peer Group Composition," *Journal of Human Resources*, Volume 10, Spring 1975, pages 189-204.

5 Anita A. Summers and Barbara L. Wolfe, "Do Schools Make a Difference?," *American Economic Review*, Volume 67, Number 4, 1977, pages 639-652.

6 William J. Webster, "Selecting Effective Teachers", *Journal of Educational Research*, Volume 91, Number 4, 1988, pages 245-253.

7 Ronald G. Ehrenberg and Dominic J. Brewer, "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," *Economics of Education Review*, Volume 13, Number 1, 1993, pages 1-17.

8 David H. Monk, "Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement," *Economics of Education Review*, Volume 13, Number 2, 1994, pages 125-145.

9 Ballou and Podgursky, op. cit., page 11.

B. SELECTIVITY OF THE UNDERGRADUATE SCHOOLS OF TEACHERS

1. Public School Teachers and Private School Teachers in the United States

The quality of the undergraduate schools of public school teachers and private school teachers in the United States was estimated by Ballou and Podgursky[10] from data on institutions awarding the undergraduate degrees of teachers taken from the national schools and staffing survey of 1987-1988.[11] The schools were classified according to the Barron's profiles of American colleges for 1991.[12]

Barron's rankings are based on the scores of matriculating students on the SAT or ACT, their high school standing and the percentage of applicants admitted. The category of Ballou and Podgursky's designated as "selective" corresponds to Barron's highest two ratings—"most competitive" and "highly competitive." The category "above average" corresponds to Barron's "very competitive." The category "average" corresponds to Barron's "competitive." The category "below average" corresponds to the rest of the Barron's categories.

Ballou and Podgursky obtained the following results for public school teachers and private school teachers in the United States.

	Percent of Public School Teachers	Percent of Private School Teachers
Selective	6.5	10.7
Above average	14.8	13.1
Average	49.0	44.5
Below Average	25.3	18.1
Unrated College	3.7	12.7
No Bachelor's Degree	0.7	2.9

2. DoDEA Teachers

A sample was taken from a data base of 8565 teachers in DoDEA. Teacher's records were sorted by social security number. Every 70th record was extracted, yielding a sample of 122

10 Ballou and Podgursky, op. cit., page 131.

11 Susan P. Choy et al., *Schools and Staffing the United States: A Statistical Profile*, 1987-1988, National Center for Educational Statistics, 1992.

12 *Barron's Profiles of American Colleges 1991*, Barron's Educational Services.

teacher's records. The undergraduate colleges for these teachers were obtained from the computerized data base or from the personnel files of the teachers.

The schools were classified according to the Barron's profiles of American colleges for 2000.[13]

The following results were obtained for DoDEA teachers.

	Percent of DoDEA Teachers
Selective	9.8
Above Average	16.4
Average	47.5
Below Average	19.7
Unrated College	6.6

3. Comparison of Public School Teachers, Private School Teachers, and DoDEA Teachers

Following are the results for public, private and DoDEA teachers:

	Public	Private	DoDEA
Selective	6.5	10.7	9.8
Above Average	14.8	13.1	16.4
Average	49.0	44.5	47.5
Below Average	25.3	18.1	19.7
Unrated College	3.7	12.7	6.6
No Bachelor's Degree	0.7	2.9	0.0

This yields the following cumulative comparison.

13 *Barron's Profiles of American Colleges 2000*, Barron's Educational Services, 1998.

	Public	Private	DoDEA
Selective or Better	6.5	10.7	9.8
Above average or Better	21.3	23.8	26.2
Average or Better	70.3	68.3	73.7
Below Average or Better	95.6	86.4	93.4
Unrated College or Better	99.3	99.1	100.0
No bachelor's Degree or Better	100.0	100.0	100.0

DoDEA has half again as large a percent of teachers from selective schools as do the public schools and almost as large a percentage as the private schools. DoDEA has 26.2 percent teachers from above average or better schools while the public schools have 21.3 percent and the private schools have 23.8 percent. DoDEA has 73.7 percent teachers from average or better schools while the public schools have 70.3 percent and the private schools have 68.3 percent. (One caveat is that the 12.7 percent of private school teachers from unrated colleges might affect these findings.)

On average, DoDEA teachers attended higher-quality undergraduate schools and are thus presumably better-qualified intellectually compared to public school teachers. Their qualifications are also better than private school teachers in two of the measures—above average or better and average or better—though again these findings might be affected by the unrated colleges.

C. SAT AND ACT SCORES OF UNDERGRADUATE SCHOOLS OF DODEA TEACHERS

1. Data

Table V-1 presents SAT and ACT data for the undergraduate schools for the 122 observations on DoDEA teachers. These data were obtained from the College Board Handbook for 2000.[14] ACT data were converted to SAT data using a table from the College Board.[15] Table V-1 also presents the competitiveness ratings used in the previous section.

---

[14] *College Handbook 2000*, The College Board, 1999.

[15] *Admission Staff Handbook for the SAT Program 1999-2000*, The College Board, 1999.

Table V-1. SAT and ACT Scores and Competitiveness Ratings

Number		25 V	75 V	25 M	75 M	25 ACT	75 ACT	25 SAT	75 SAT	Avg. SAT	Avg. SAT	Competi-tiveness
1		490	590	490	600					1085		C
2		500	540	470	520					1015		LC
3		580	670	590	680					1260		HC
4		470	550	490	590					1050		C
5		460	540	440	530					985		LC
6		480	570	480	580					1055		C
7		410	500	400	500					905		C+
8		560	640	570	640					1205		HC
9		430	520	430	520					950		C
10		500	590	500	590					1090		VC
11		480	530	480	580					1035		C
12						19	24	910	1110		1010	C
13		520	620	540	640					1160		VC
14	no test											C
15		420	530	410	530					945		C+
16		540	640	560	670					1205		HC
17		500	590	500	600					1095		C
18		490	590	480	590					1075		C
19		570	670	600	710					1275		HC
20		500	600	500	600					1100		VC
21		530	610	520	590					1125		VC
22		540	640	560	670					1205		VC
23		430	520	410	510					935		C
24		410	550	440	570					985		C
25		460	550	440	540					995		C
26		460	560	460	550					1015		LC
27		450	580	430	550					1005		C
28		540	630	550	630					1175		VC
29		460	570	460	570					1030		C
30	no test											LC
31		500	600	500	480					1040		C
32		460	550	460	550					1010		C
33		560	670	570	670					1235		HC
34		460	570	450	560					1020		C
35		470	570	470	560					1035		C
36		530	630	540	650					1175		VC
37		480	610	480	600					1085		C
38		480	610	480	600					1085		C
39		550	640	550	640					1190		HC
40		450	540	440	530					980		C
41		400	550	400	550					950		C
42		490	550	470	540					1025		C
43		500	580	500	580					1080		VC
44		560	670	580	680					1245		HC
45		490	600	500	610					1100		C
46	no test											C
47		520	620	520	620					1140		VC
48		510	610	510	610					1120		C+
49	no test											LC
50		500	610	510	640					1130		VC
51		540	630	560	650					1190		HC
52		490	580	480	570					1060		LC
53						19	24	910	1110		1010	LC
54		440	550	430	550					985		LC
55						20	24	950	1110		1030	C
56						19	24	910	1110		1010	C
57						19	24	910	1110		1010	C
58		490	610	490	610					1100		VC
59		480	610	470	590					1075		C
60		470	590	460	590					1055		C
61						19	24	910	1110		1010	LC
62						19	25	910	1140		1025	C
63						19	24	910	1110		1010	C
64						19	24	910	1110		1010	C
65						18	24	870	1110		990	NC
66						18	24	870	1110		990	NC

Table V-1. SAT and ACT Scores and Competitiveness Ratings (cont)

Number		25 V	75 V	25 M	75 M	25 ACT	75 ACT	25 SAT	75 SAT	Avg. SAT	Avg. SAT	Competi-tiveness
67	no test											NC
68	no test											C+
69		460	550	460	550					1010		C
70		460	580	460	570					1035		C
71						19	25	910	1140		1025	LC
72						19	24	910	1110		1010	C
73		480	600	480	600					1080		C
74						18	24	870	1110		990	C
75		423	662	438	624					1074		VC
76	no test											LC
77	unk school											
78						21	26	990	1180		1085	C
79	no test											C
80	no test											C
81	no test											
82		530	640	560	670					1200		VC
83		530	640	560	670					1200		VC
84		480	580	480	590					1065		VC
85						19	24	910	1110		1010	C
86		420	540	410	530					950		LC
87		470	560	470	560					1030		C
88	no school											
89		520	610	520	640					1145		VC
90		430	600	430	560					1010		C
91		490	600	490	610					1095		VC
92		530	650	540	660					1190		VC
93						19	25	910	1140		1025	NC
94		470	590	470	580					1055		LC
95	no test											LC
96						18	24	870	1110		990	NC
97						18	24	870	1110		990	NC
98		430	550	420	540					970		LC
99						19	24	910	1110		1010	C
100		490	590	470	580					1065		C+
101		450	560	460	570					1020		C
102	no test											LC
103		480	610	490	600					1090		NC
104		500	600	500	590					1095		C
105		440	560	430	560					995		LC
106		450	570	450	580					1025		C
107		580	680	570	660					1245		VC
108	no school											
109		520	620	540	650					1165		HC
110		520	620	540	650					1165		HC
111		570	670	600	710					1275		HC
112	unk school											
113		420	530	440	560					975		C
114		490	600	500	620					1105		VC
115		520	620	540	650					1165		HC
116		480	600	480	600					1080		C
117		530	650	500	620					1150		C+
118		580	670	540	630					1210		
119	unk school											
120						18	24	870	1110		990	C
121						17	20	830	950		890	C+
122		580	670	540	630					1210		
Avg.										1087	1005	

For the SAT, data were extracted for the 25th and 75th percentile of the verbal and mathematics scores of the matriculating freshmen. They were converted to an average total score. For the ACT, data were extracted for the 25th and 75th percentile for the composite score. They were converted to an average SAT score.

For the SAT alone there were 85 observations for which the average score was 1087.

For the ACT alone there were 22 observations for which the average score was 1005.

For the SAT and ACT there were 107 observations for which the average score was 1070.

If it is assumed that the missing seven observations have the very low score of 900, the combined score of the 122 observations is 1046.

2. Interpretations

The average score of students who took the SAT in 1999 was 1017. The participation rate was 43 percent. The average score of all 4-year institutions is not known. Since many do not require the SAT, it is not unreasonable to assume that it would be well below 1000. Thus the average score of institutions from which DoDEA teachers graduated, which is presumably between 1046 and 1070, is far higher than the average score of all 4-year institutions.

However, the average score of institutions from which all public school teachers have graduated is not known.

One surrogate might be the scores of the historically teachers colleges in Pennsylvania. Table V-2 contains data[16] on the SAT scores of the colleges presently designated by titles including the suffix "University of Pennsylvania," such as Indiana University of Pennsylvania. (its previous name was Indiana State Teacher's College). Not reducing the average score for the unknown score of Cheyney, the average score is 999.

Graduates of teaching programs in public colleges tend to be certified and are thus easier to hire. Ballou and Podgursky state that "public school officials do not appear to give special weight to a strong academic background when recruiting teachers. This difference in priorities is likely to be one of the reasons private schools obtain a comparatively high percentage of teachers from good colleges with academic majors." For instance, data from 1987-1888 show that 95.9 percent of public school teachers and 55.9 percent of non-religious private school teachers are certified.[17]

---

16 Data are from *College Handbook 2000*, The College Board, 1999.

17 Ballou and Podgursky, op. cit., pages 140-141.

Table V-2. Selected Colleges in Pennsylvania

Name	25 V	75 V	25 M	75 M	Average (V+M)
Bloomsburg	460	550	460	560	1015
California	430	520	410	510	935
Cheyney					
Clarion	420	530	410	530	945
East Stroudsburg	430	520	430	520	950
Edinboro	490	540	470	520	1010
Indiana	490	580	480	570	1060
Kutztown	450	540	440	530	980
Lock Haven	430	530	430	530	960
Mansfield	430	520	420	530	950
Millersville	480	580	480	580	1060
Shippensburg	480	570	470	570	1045
Slippery Rock	480	570	470	570	1045
West Chester	480	560	470	550	1030
Average (not counting Cheyney)					999

It is also worth noting that SAT scores of students by intended college major are relatively low for those intending to be teachers. For instance, the fall 1996 SAT scores of high school seniors nationwide by intended college major were as follows:[18]

Education -- 964

Business -- 982

Biological science -- 1091

Language and literature -- 1150

Physical science -- 1170

Mathematics -- 1178

The DoDEA SAT scores are well above the SAT scores of all 4-year institutions, of all students taking the SAT, of the historically teachers colleges in Pennsylvania, and of the intended education majors nationwide.

---

[18] Robert P. Strauss, "Who Gets Hired to Teach? The Case of Pennsylvania," in Marci Kanstoroom and Chester E. Finn, Jr. (Editors), *Better Teachers, Better Schools*, Thomas Fordham Foundation, 1999, page 112.

As discussed in Section A of this chapter, several studies have tied achievement of students to the colleges of their teachers. It seems very likely that the colleges of the DoDEA teachers are better than the colleges of public school teachers, based on the selectivity data on colleges presented in Section B of this chapter as well as the SAT data on colleges presented in Section C of this chapter.

VI. MEASURING THE DODEA CONTRIBUTION TO STUDENT ACHIEVEMENT

The question remains: How much of the reason for the better-than-average scores achieved by DoDEA students is due to the DoDEA students being better than average, and how much is due to the DoDEA educational system being better than average?

This chapter proposes a simple method that DoDEA could use to address this question. Since this method requires data not currently available, it is not applied here. However, the data required should not be difficult for DoDEA to collect in the future, if it decides to do so. Thus, availability of data need not be a problem if DoDEA decides to implement this method. Further, this method quantitatively considers all nine combinations of the three possible cases that: DoDEA students are better than, are about equal to, or are worse than the national average student, with the three possible cases that the DoDEA system is better than, is about equal to, or is worse than the national average educational system. Thus, this method can also address situations in which DoDEA students do not outscore the national average, should such situations occur in the future.

A. OVERVIEW

An overview of this method is as follows.

Suppose that a test is given to all, or to a representative number, of the nation's students, including DoDEA students, who are in a certain grade, say grade g, in a certain year. For the DoDEA students, calculate the "discounted" fractional portion of their education that they received attending a DoDEA school. Find a continuous monotonic curve that approximates the scores achieved on that test by DoDEA students as a function of this fraction of education received in DoDEA. Let $s(x)$ be this function, so that $s(1)$ is the estimated average score of those students who have received all of their education from DoDEA, and $s(0)$ is the estimated average score of those students who are just entering DoDEA for the first time. Let $\bar{s}$ be the overall average score achieved by DoDEA students on that test, and let $\bar{n}$ be the overall average score achieved by the nation's students on that test. Then, concerning that test,

$s(1) - s(0)$ is a measure of the potential contribution (if $s(1) \geq s(0)$), or harm (if $s(1) < s(0)$), to DoDEA student scores that would be due to the DoDEA educational system if all DoDEA students in grade g were to have attended DoDEA schools for all of grades 1 through g,

$\bar{s} - s(0)$ is a measure of the actual contribution (if $\bar{s} \geq s(0)$), or harm (if $\bar{s} < s(0)$), to DoDEA student scores due to the DoDEA educational system,

$s(0) - \bar{n}$ is a measure of the amount by which DoDEA students are better than (if $\bar{s} \geq \bar{n}$), or worse than (if $\bar{s} < \bar{n}$), the national average student, and

$\bar{s} - \bar{n}$ is the average amount by which DoDEA students outscored (if $\bar{s} \geq \bar{n}$), or were outscored by (if $\bar{s} < \bar{n}$), the nation's students.

Thus, for example, if $\bar{s} - \bar{n}$, $\bar{s} - s(0)$, and $\bar{s} - \bar{n}$ are all non-negative, then $s(0) - \bar{n}$ is the portion of $\bar{s} - \bar{n}$ that is due to the DoDEA students being better than average, and $\bar{s} - s(0)$ is the portion of $\bar{s} - \bar{n}$ that is due to the DoDEA system being better than average.

Average results over several tests given to several grades can be computed by taking weighted averages of percentile-normalized scores.

A more detailed description of this method is given next.

B. DETAILS

Suppose that a test is given to all, or to a representative number, of the nation's students, including DoDEA students, either at the beginning of grade $g + 1$ or at the end of grade g. (Given some experience with this method as described here, it may be possible to extend it to cover tests given in mid-year—provided that reasonable values for d, as defined below, are far enough above 0.) Let $\boldsymbol{P}$ be the set of DoDEA students who take this test. Suppose that, if the test is given at the end of grade g, then DoDEA can obtain the scores of those students who are just entering (or re-entering) the DoDEA system in grade $g + 1$ and who took the test, but not at a DoDEA school. Include these students in $\boldsymbol{P}$. (Assume that students who were in grade g and took the test, and then were held back, if any, are considered in some appropriate manner.)

Let d be an "educational yearly discount factor" (i.e., a weighting factor) in the following sense. For a test taken at the end of grade g (or, equivalently, the start of grade $g + 1$), grade g counts as 1 year, and, for $0 \leq d \leq 1$, grade $g - 1$ counts as d of a year, grade $g - 2$ counts as d^2 of a year, and so on back to first grade, which counts as d^{g-1} of a year. Thus, relative to a test taken at the end of grade g, a student would have a total of

VI. MEASURING THE DODEA CONTRIBUTION TO STUDENT ACHIEVEMENT

The question remains: How much of the reason for the better-than-average scores achieved by DoDEA students is due to the DoDEA students being better than average, and how much is due to the DoDEA educational system being better than average?

This chapter proposes a simple method that DoDEA could use to address this question. Since this method requires data not currently available, it is not applied here. However, the data required should not be difficult for DoDEA to collect in the future, if it decides to do so. Thus, availability of data need not be a problem if DoDEA decides to implement this method. Further, this method quantitatively considers all nine combinations of the three possible cases that: DoDEA students are better than, are about equal to, or are worse than the national average student, with the three possible cases that the DoDEA system is better than, is about equal to, or is worse than the national average educational system. Thus, this method can also address situations in which DoDEA students do not outscore the national average, should such situations occur in the future.

A. OVERVIEW

An overview of this method is as follows.

Suppose that a test is given to all, or to a representative number, of the nation's students, including DoDEA students, who are in a certain grade, say grade g, in a certain year. For the DoDEA students, calculate the "discounted" fractional portion of their education that they received attending a DoDEA school. Find a continuous monotonic curve that approximates the scores achieved on that test by DoDEA students as a function of this fraction of education received in DoDEA. Let $s(x)$ be this function, so that $s(1)$ is the estimated average score of those students who have received all of their education from DoDEA, and $s(0)$ is the estimated average score of those students who are just entering DoDEA for the first time. Let $\bar{s}$ be the overall average score achieved by DoDEA students on that test, and let $\bar{n}$ be the overall average score achieved by the nation's students on that test. Then, concerning that test,

$s(1) - s(0)$ is a measure of the potential contribution (if $s(1) \geq s(0)$), or harm (if $s(1) < s(0)$), to DoDEA student scores that would be due to the DoDEA educational system if all DoDEA students in grade g were to have attended DoDEA schools for all of grades 1 through g,

$\bar{s} - s(0)$ is a measure of the actual contribution (if $\bar{s} \geq s(0)$), or harm (if $\bar{s} < s(0)$), to DoDEA student scores due to the DoDEA educational system,

$s(0) - \bar{n}$ is a measure of the amount by which DoDEA students are better than (if $\bar{s} \geq \bar{n}$), or worse than (if $\bar{s} < \bar{n}$), the national average student, and

$\bar{s} - \bar{n}$ is the average amount by which DoDEA students outscored (if $\bar{s} \geq \bar{n}$), or were outscored by (if $\bar{s} < \bar{n}$), the nation's students.

Thus, for example, if $\bar{s} - \bar{n}$, $\bar{s} - s(0)$, and $\bar{s} - \bar{n}$ are all non-negative, then $s(0) - \bar{n}$ is the portion of $\bar{s} - \bar{n}$ that is due to the DoDEA students being better than average, and $\bar{s} - s(0)$ is the portion of $\bar{s} - \bar{n}$ that is due to the DoDEA system being better than average.

Average results over several tests given to several grades can be computed by taking weighted averages of percentile-normalized scores.

A more detailed description of this method is given next.

B. DETAILS

Suppose that a test is given to all, or to a representative number, of the nation's students, including DoDEA students, either at the beginning of grade $g + 1$ or at the end of grade g. (Given some experience with this method as described here, it may be possible to extend it to cover tests given in mid-year—provided that reasonable values for d, as defined below, are far enough above 0.) Let $\boldsymbol{P}$ be the set of DoDEA students who take this test. Suppose that, if the test is given at the end of grade g, then DoDEA can obtain the scores of those students who are just entering (or re-entering) the DoDEA system in grade $g + 1$ and who took the test, but not at a DoDEA school. Include these students in $\boldsymbol{P}$. (Assume that students who were in grade g and took the test, and then were held back, if any, are considered in some appropriate manner.)

Let d be an "educational yearly discount factor" (i.e., a weighting factor) in the following sense. For a test taken at the end of grade g (or, equivalently, the start of grade $g + 1$), grade g counts as 1 year, and, for $0 \leq d \leq 1$, grade $g - 1$ counts as d of a year, grade $g - 2$ counts as d^2 of a year, and so on back to first grade, which counts as d^{g-1} of a year. Thus, relative to a test taken at the end of grade g, a student would have a total of

$$m(d,g) = 1 + d + d^2 + \cdots + d^{g-1} = \begin{cases} 1 & \text{if } d = 0 \\ \dfrac{1-d^g}{1-d} & \text{if } 0 < d < 1 \\ g & \text{if } d = 1 \end{cases}$$

discounted years of education. Since no confusion will result, $m(d,g)$ is written as m below.

It is assumed that DoDEA can estimate a reasonable value (or range of values) for d.

For each DoDEA pupil p in $\boldsymbol{P}$ and for each grade g' from 1 through g, let

$H(p,g')$ = the fraction of pupil p's education in grade g' that pupil p received attending a DoDEA school.

For example, $H(p,g') = 1$ if pupil p attended a DoDEA school for all of grade g', and $H(p,g') = 0$ if pupil p did not attend a DoDEA school for any significant portion of grade g'.

For each DoDEA pupil p in $\boldsymbol{P}$, let

$$F(p) = \frac{1}{m} \sum_{g'=1}^{g} d^{g-g'} H(p,g').$$

Note that $0 \le F(p) \le 1$, and that $F(p)$ is the "discounted" fractional portion of pupil p's education that was received attending a DoDEA school.

For each DoDEA pupil p in $\boldsymbol{P}$, let $S(p)$ be the score that pupil p achieved on the test in question. Let

$$\bar{s} = \frac{1}{|\boldsymbol{P}|} \sum_{p \in \boldsymbol{P}} S(p),$$

so that, as in Section A, $\bar{s}$ is the average score achieved by the DoDEA students on that test. Suppose that the minimum possible score on the test in question is 0, and that its maximum score is u. Then $0 \le S(p) \le u$ for p in $\boldsymbol{P}$.

Consider the set $\{(F(p),S(p)) \mid p \in \boldsymbol{P}\}$ as giving $|\boldsymbol{P}|$ points on the x,y plane, where the $F(p)$ are $|\boldsymbol{P}|$ values of an independent variable, say x, and the $S(p)$ are the corresponding $|\boldsymbol{P}|$ values of the dependent variable, say y. Suppose that a function $s(x)$ can be found such that

1) s maps $[0,1]$ into $[0,u]$,
2) s is continuous and monotonic, and
3) the points $\{(F(p),s(F(p)) \mid p \in \boldsymbol{P}\}$ provide a reasonable fit to the points $\{(F(p),S(p)) \mid p \in \boldsymbol{P}\}$.

Then, as in Section A above, the function $s(x)$ approximates the scores achieved on the test by the DoDEA students as a function of the fraction of their education that each of those students received in DoDEA. The method proposed here assumes that such a functional approximation can be found.

Given $s(x)$, and given $\bar{n}$ as defined above, the quantities $s(1) - s(0)$, $\bar{s} - s(0)$, $\bar{s} - \bar{n}$, and $\bar{s} - \bar{n}$ have the interpretations given above.

Since results from individual tests could, at times, have a high degree of variability, it might be desirable compute results based on scores achieved over several tests. The method proposed here can be applied to scores taken from multiple tests, once these scores have been normalized by percentiles.

C. IMPLEMENTATION ISSUES

In addition to collecting the data and calculating the results, the following issues must be addressed in order to implement the method described here.

1. Some suitable functional form for $s(x)$ must be selected. A bounded linear form could be used; but the data might suggest alternatives. Accordingly, this issue may best be addressed after data have been obtained.

2. Once a functional form for $s(x)$ has been selected, some method (such as least squares) for determining its parameters must be selected and, if the method is new, implemented.

3. Some method for estimating d must be selected and implemented. This method could be entirely judgmental, mostly analytic, or any of various combinations of these approaches.

D. EXTENSIONS

1. As mentioned above, it could be useful to be able to handle mid-year tests.

2. For $1 \leq g' < g$, let $c(g',g)$ be the weighting factor to be applied to grade g' when estimating its impact, relative to grade g, on student scores achieved at the end of grade g. Then Section B proposes setting $c(g',g) = d^{g-g'}$ for some d such that $0 \leq d \leq 1$. Other specifications for $c(g',g)$ could be used if data or judgment warranted.

VII. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Conclusions are presented for the quantitative analysis of educational quality based on an assessment of:

- Student Achievement
- Resources
- College Attendance
- Teacher Quality
- DoDEA Contribution to Student Achievement

1. Student Achievement

Three measures of student achievement are considered in the analysis. They are the National Assessment of Educational Progress (NAEP), CTBS/Terra Nova, and the Scholastic Aptitude Test (SAT). We conclude that the DODEA students are doing well on the NAEP. Specifically, in reading and writing, DoDDS and DDESS score significantly higher than the national average. In Grade 4 mathematics, DoDDS and DDESS are slightly higher. In Grade 8 mathematics, DoDDS is higher and DDESS is lower. In Grade 8 science, DoDDS and DDESS are higher. Achievement of minority students in DoDDS and DDESS is significantly better than that of minority students in the nation. The lower-scoring students in DoDDS and DDESS significantly outperform the lower-scoring students in the nation. The higher-scoring students in DoDDS and DDESS slightly outperform the higher-scoring students in the nation.

We conclude that the DoDEA students are doing very well on the CTBS/Terra Nova, but caution the reader that the high percentiles may be optimistic due to the "Lake Wobegon Effect."[1] Comparisons with Montgomery County, Maryland, show that Montgomery County students have higher scores than DoDDS and DDESS. Though Montgomery County is relatively wealthy, it has a similar ethnicity in student enrollment. Comparisons with other school districts across the nation show DoDDS and DDESS performance to usually, but not always, be better.

---

1 Studies of nationally-normed tests have found that all states were above average; norms developed by contractors have insured that the results for all customers for their tests would be above average.

We conclude that the DoDDS students are doing about average on the SAT. We conclude that the DDESS students are doing below average on the SAT; but if the Antilles students are removed from the population, the DDESS students are doing about average.

2. Resources

The primary measures of resources considered in the analysis are expenditures per student, class size, and teacher education. We conclude that DoDEA resources are well above average. Expenditures per student are among the highest in the nation. Expenditures are comparable to those of the highest-spending states, which are mostly in the Northeast. Even if costs of living overseas and costs of permanent change of station are removed (to increase comparability with expenditures within the United States), the cost per student is still at the high end of the scale. Class size is very good when compared with the national average.

DoDDS and DDESS have among the highest percentages of teachers with advanced degrees in the nation. DDESS has a higher percent of teachers with advanced degrees than does any state in the nation, and the percent of teachers with advanced degrees in DoDDS is exceeded by only two states.

3. College Attendance

We conclude that the quality of planned college attendance is very good for those DoDEA students planning to attend college, and is a little below the national average for the average DoDEA student. However, missing observations on where students plan to attend college could have an impact on these comparisons.

4. Teacher Quality

We compared the quality of teachers in DoDEA to the quality of teachers in public and private schools in the United States in terms of estimated intellectual capabilities of the teachers. We conclude that DoDEA teachers have excellent qualifications based on their attendance at higher-quality undergraduate schools.

5. DoDEA Contribution to Student Achievement

The study team addressed the issue of whether the quality of the DoDEA student performance is better than average because the students themselves are better than average or because the DoDEA educational system is better than average. Data were not available to resolve this question. However, we offer a methodology for estimating the contribution of DoDEA to student achievement.

B. RECOMMMENDATIONS

The following recommendations are made:

- Continue to focus on and deliver good student achievement as evidenced in Chapters II and IV.
- When recruiting new teachers and managing existing teachers, concentrate on maintaining DoDEA's high level of intellectual capability of its teachers.
- Introduce a system to measure the contribution of DoDEA to student achievement. Collect and maintain a database of student time spent in DoDEA schools and associated measures of student achievement and college attendance.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2000	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Review of Department of Defense Education Activity (DoDEA) Schools Volume II: Quantitative Analysis of Educational Quality	DASW 01-98-C-0067 BE-6-1786
6. AUTHOR(s) Lowell Bruce Anderson, Jerome Bracken, Marilyn C. Bracken	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses 1801 N. Beauregard Street Alexandria, VA 22311	8. PERFORMING ORGANIZATION REPORT NUMBER IDA Paper P-3544
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) OUSD (P&R) The Pentagon Washington, DC 20301	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited.	

13. ABSTRACT (Maximum 200 words)

This report provides an independent review of DoDEA schools focused on the quality of education of the overall system. It reports the findings from extensive interviews with military leaders, parents, administrators, teachers, students, and other individuals in Europe, the Pacific, and CONUS and presents detailed analyses and comparisons of data and additional research on numerous issues identified. The review focused on student achievement, curriculum, teachers and administrators, school environment, and DoDEA management and relationships. The DoDEA school system provides students with an above average to excellent education. Interview comments and detailed analysis did not result in identification of any overall dissatisfaction or serious problems with the DoDEA school system. Analysis and comparisons with other systems indicate that DoDEA performs well and compares favorably in many areas. Improvements can be made that should further enhance the quality of education. Renewed cooperation and collaboration among stakeholder groups indicate a healthy, positive relationship that will help students. There are 53 recommendations that should have an impact on student learning and improve the overall quality of education. All recommendations can be implemented, but some will take an investment of additional resources.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Quality of education, DoD Education Activity, DoDEA, DoDDS, DDESS, student achievement, NAEP, CTBS/Terra Nova, SAT, curriculum, AP courses/tests, content standards, vocational education, distance learning, professional development, SAC, IAC, School Board, staffing standards, PTR	94
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL