DOCUMENT RESUME

ED 446 153 TM 031 934

AUTHOR McLaughlin, Don; Bandeira de Mello, Victor; Cole, Susan; Arenson, Ethan

TITLE Comparison of National Assessment of Educational Progress (NAEP) and Statewide Assessment Results: Report to Maryland on 1996 and 1998 Assessments.

INSTITUTION American Institutes for Research in the Behavioral Sciences, Palo Alto, CA.

SPONS AGENCY Department of Education, Washington, DC.; Educational Testing Service, Princeton, NJ.

PUB DATE 2000-02-00

NOTE 26p.

CONTRACT R902F9800001

PUB TYPE Reports - Evaluative (142)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Academic Achievement; Comparative Analysis; Elementary Education; Mathematics; *National Competency Tests; Reading; *State Programs; *Test Results; *Testing Programs

IDENTIFIERS *National Assessment of Educational Progress

ABSTRACT

Maryland participated in the State National Assessment of Educational Progress (NAEP) in 1996 and 1998 when samples of approximately 2,500 fourth and eighth graders in 100 schools participated in the testing sessions for mathematics (1996) and reading (1998). This report is an initial attempt to return information to Maryland about the relations between the state assessment and State NAEP data. This initial prototype report, versions of which are tailored to individual participating states, contains six sets of tables: (1) correlation of school mean state assessment and NAEP scores; (2) NAEP subscale weights in accounting for state assessment results; (3) differential sensitivity of the two assessments' results to poverty; (4) differential sensitivity of the two assessments' results to minority status; (5) differential sensitivity of the two assessments' results to urbanicity; and (6) percent of NAEP-proficient students in high, medium, and low scoring schools. At the school level, most statewide reading and mathematics assessments have been moderately or highly correlated with NAEP assessments of the same subject matter. NAEP and statewide assessments measure similar achievement gaps between rich and poor schools, and they agree in the finding that in a few states the gaps are noticeably smaller than in other states. This initial report is issued as a prototype because there are many sources of variation in the relations among test scores. Overall, however, the results are encouraging. An appendix contains a list of state assessment measures used in this report. (Contains 16 tables.) (SLD)

AMERICAN INSTITUTES FOR RESEARCH

Comparison of National Assessment of Educational Progress (NAEP) and Statewide Assessment Results:

Report to Maryland on 1996 and 1998 Assessments

Don McLaughlin
Victor Bandeira de Mello
Susan Cole
Ethan Arenson

American Institutes for Research

funded through
Cooperative Agreement R902F9800001
between the U.S. Department of Education
and the Educational Testing Service

February, 2000

JOHN C. FLANAGAN RESEARCH CENTER
1791 ARASTRADERO ROAD | PALO ALTO, CALIFORNIA 94304-1337

Comparison of National Assessment of Educational Progress (NAEP) and Statewide Assessment Results:

Report to Maryland on 1996 and 1998 Assessments

Don McLaughlin
Victor Bandeira de Mello
Susan Cole
Ethan Arenson

American Institutes for Research

funded through
Cooperative Agreement R902F9800001
between the U.S. Department of Education
and the Educational Testing Service

February, 2000

Comparison of National Assessment of Educational Progress (NAEP) and Statewide Assessment Results:

Report to Maryland on 1996 and 1998 Assessments

Maryland participated in State NAEP in 1996 and/or 1998, when samples of approximately 2,500 fourth and eighth grade students in 100 schools participated in testing sessions, completing booklets of mathematics (1996) and reading (1998) exercises. The results of the testing are available on the web (http://nces.ed.gov/nationsreportcard/). Maryland also conducts an annual statewide educational achievement testing program, and the purpose of this report is to provide information about the relations between NAEP and the statewide testing results, based on comparisons of school mean scores on NAEP and the statewide assessment in the same schools.

In cooperation with the NAEP program, many state testing programs provided school mean data in 1999. These data have been used in efforts to enhance the precision of State NAEP scores in 2000 through stratified sampling; and they are also being used to build a comprehensive picture of the state assessment context of NAEP. Statewide assessment data have now been merged with State NAEP data, and this report is an initial attempt to return information to Maryland about the relations between the two assessments. Based on review and feedback on this report, plans for additional reports will be implemented. The particular statewide achievement measure used in the analyses presented in this report are shown in the appendix.

This initial prototype report contains six sets of tables. Each set of tables includes figures for two grades (4 and 8) and two assessment areas (mathematics and reading). Your state's figures may not be included in one or more of the tables for several reasons. It may be that your state did not participate in both the 1996 and 1998 State NAEP assessments, or it may be that we were unable to carry out the extraction, merging, and analysis steps successfully in some cases. In some cases, we found the results to be anomalous, and rather than produce unwarranted controversy, we have omitted such results.

The six sets of tables are:

tables 1 – 2: Correlations of school mean state assessment and NAEP scores;

tables 3 – 6: NAEP subscale weights in accounting for state assessment results;

tables 7 – 8: Differential sensitivity of the two assessments' results to poverty;

tables 9 – 10: Differential sensitivity of the two assessments' results to minority status;

tables 11 – 12: Differential sensitivity of the two assessments' results to urbanicity; and

tables 13 – 16: Percent of NAEP-proficient students in high, medium, and low scoring schools.

In an informal survey, some states indicated an interest in information about the stability of relations over years, and some states indicated interest in identifying the distribution of NAEP equivalents of state achievement standards. These topics may be covered in future reports.

To add value to the statistical information about your statewide assessment, we have generated tables which show your state's statistics in the context of the distribution of similar values in other states. However, each state's assessment is unique, and the measures we have extracted from each assessment for this report are not necessarily the ones that each state would have selected, so we have not named the other states in the statistical tables. Each state's report only names the state to whom the report is addressed. Of course, if you wish to share and compare your results with other states, you are free to do so. We hope that you will find this report informative and that you will inform us about ways in which we could make it more useful to you.

The distribution of values for "other states" is based on different states in each table and in each column of multi-column tables. The other states appear in the order determined by the value of the statistic tabulated, and this is generally a different order in different columns and tables.

Correlations of school mean state assessment and NAEP scores

Tables 1 and 2 contain correlations of school mean scores on NAEP and statewide assessments, first for reading (table 1) and then for mathematics (table 2). A small number of correlations were found which were less than .3, and these are not included in the tables because we believe that they indicate a problem with our methods for extracting, merging, or analysis. In particular, we may have extracted a measure that is not the most representative for the state.

For reading, correlations range from .46 to .91 at grade 4 and from .36 to .88 at grade 8. For mathematics, correlations range from .31 to .88 at grade 4 and .55 to .92 at grade 8. The majority of the correlations in these tables are .75 or greater. Correlations greater than .70 are sufficient to warrant investigation of the potential for developing linkages for the purpose of projecting NAEP school and district scores based on statewide assessment data[1], while correlations less than .70 represent opportunities for studying differences in assessments of nominally similar domains. Lower correlations are more difficult to interpret, however, because they may also be due to differences between NAEP and the statewide assessment in exclusion and accommodation policies, motivational contexts, and time of year, and they may be due to limits on statewide assessment reliability, as well as to domain differences.

It should be pointed out that, other things being equal, the correlations are somewhat higher in states in which larger percentages of achievement variation is *between* schools as opposed to within schools. This is an artifact of use of school means for this correlation. Further analysis can separate this source of variation.

---

[1] Because no statewide assessments are completely equivalent to NAEP, the criteria for valid and reliable linkages are complex. Draft standards and methods for linkage are presented in McLaughlin (1999) *Study of the Linkages of 1996 NAEP and State Mathematics Assessments in Four States: Final Report.* (NCES)

Table 1. Correlations of School Mean State Assessment Scores with the 1998 NAEP Reading Composite School Means, by State.

State		Grade 4	State		Grade 8
	#	0.91			0.88
	#	0.90		#	0.87
Maryland	*#	0.87		#	0.86
		0.83			0.85
		0.80		#	0.84
	*	0.80	Maryland	#	0.83
		0.79			0.83
		0.78		#	0.82
		0.77			0.82
		0.77			0.79
		0.76			0.78
		0.76			0.77
	*	0.76			0.77
		0.76			0.77
	#	0.76			0.75
		0.75			0.74
		0.68			0.73
		0.66			0.72
		0.65			0.68
	*#	0.64			0.65
	#	0.57			0.60
	*	0.51			0.54
	*	0.50			0.50
		0.46			0.36

Note: * indicates that the statewide assessment was a grade 3 assessment, rather than grade 4.

# indicates that the NAEP and statewide assessment school means are based on the same sample of students in each school.

Table 2. Correlations of School Mean State Assessment Scores with the 1996 NAEP Mathematics Composite School Means, by State.

State		Grade 4	State		Grade 8
	*#	0.88	Maryland		0.92
		0.85		#	0.90
		0.85			0.89
		0.85			0.89
		0.84		#	0.87
	#	0.84		#	0.85
Maryland	*	0.83			0.85
		0.77			0.85
		0.77			0.85
		0.76			0.84
	#	0.76			0.84
		0.75			0.84
		0.74			0.83
		0.73			0.82
	*	0.72			0.79
		0.67			0.77
		0.64			0.75
	#	0.62			0.74
		0.59			0.73
		0.59			0.70
	*	0.57		#	0.70
		0.51			0.69
		0.50			0.68
		0.39			0.66
	*	0.35			0.61
		0.31			0.55

Note: * indicates that the statewide assessment was a grade 3 assessment, rather than grade 4.

# indicates that the NAEP and statewide assessment school means are based on the same sample of students in each school.

NAEP subscale weights in accounting for state assessment results

The NAEP reading and mathematics statistics cited most frequently (e.g., achievement level percentages) pertain to "composite scores," but NAEP also provides subscale scores for five mathematics subscales and three reading subscales. In mathematics, the subscales are numbers and operations, measurement, geometry, data analysis and statistics, and algebra and functions. In reading, the subscales pertain to purposes for reading: for information, for literary experience, and for performing a task (the last is not included in the grade 4 reading assessment). Examples of NAEP test items can be found on the web (http://nces.ed.gov/nationsreportcard/).

It may be of some interest to know how your state assessment is related to the NAEP subscales. Tables 3, 4, 5, and 6 provide information about the relations of statewide assessments to the NAEP subscales. If simple correlations were used for this purpose, the results would be difficult to interpret because the NAEP subscales are highly intercorrelated: both mathematics and reading abilities are cumulative, and students who are proficient in one subdomain are likely to be proficient in other areas as well. To highlight subscale differences, we have displayed standardized multiple regression coefficients[2], predicting statewide assessment school means based on NAEP subscale school means, in these tables. These coefficients can be negative, and to avoid misinterpretation (a negative regression coefficient does *not* mean a negative relation), we have omitted negative coefficients. Rows in the mathematics subscale tables are sorted by the value of the coefficient for *numbers and operations* and in the reading subscale tables by the value of the coefficient for *reading for literary experience*.

As can be seen in these tables, there is a wide range of patterns of relations across states, the exception being that for grade 4 mathematics there is a generally high "loading" on numbers and operations. A high subscale coefficient in your state may mean that your state assessment emphasizes the skills in that NAEP subdomain, but it may also mean that there is a relatively large variation on that subdomain between schools in your state. Further analysis is needed to separate these explanations.

---

[2] Standardized regression coefficients are similar to correlation coefficients in that they are dimensionless and generally vary between –1 and +1. However, when several correlated subscales are used jointly to predict a score, the standardized regression coefficients highlight the strongest relations and minimize the other relations. As a result, although correlations with all subscales may be high, only one standardized regression coefficient is usually high. As an example, if the correlations between two predictors and a criterion are 0.5 and 0.4 (i.e., both are nearly equal), and the correlation between the two predictors is 0.5, then the corresponding standardized regression coefficients are 0.4 and 0.2 (i.e., one is noticeably larger than the other).

Table 3. Standardized Beta Coefficients when Regressing School Mean State Assessment Scores with the 1996 NAEP Mathematics School Mean Grade 4 Scores of Each Content Area, by State.

State	Number and Operations	Measurement	Geometry	Data Analysis and Statistics	Algebra and Functions
	0.75	0.14	–	0.16	–
	0.75	0.11	–	–	–
	0.71	0.27	–	–	0.10
	0.68	0.20	–	–	0.06
	0.61	0.03	0.21	–	–
	0.56	0.08	0.07	0.06	0.14
	0.53	0.24	–	0.61	–
	0.52	–	–	0.05	0.36
	0.50	–	–	0.32	0.24
	0.50	–	0.06	0.03	0.30
	0.47	0.16	0.30	0.08	–
	0.46	–	0.24	0.10	0.15
	0.46	–	0.15	0.05	0.18
	0.44	–	0.05	–	0.01
	0.35	–	0.11	0.18	0.17
	0.31	0.38	–	0.18	0.02
	0.31	–	0.26	0.22	–
	0.30	0.09	0.03	0.18	0.11
	0.24	0.14	0.15	0.24	–
	0.09	0.16	0.18	0.26	–
Maryland	0.08	0.23	0.33	0.25	–
	0.05	0.24	0.20	0.17	0.15
	0.03	0.36	0.27	0.06	0.16
	0.01	0.13	–	0.13	0.21
	–	0.43	0.12	0.10	0.22

Note: " – " indicates that the standardized regression coefficient is less than zero.

Table 4. Standardized Beta Coefficients when Regressing School Mean State Assessment Scores with the 1996 NAEP Mathematics School Mean Grade 8 Scores of Each Content Area, by State.

State	Number and Operations	Measurement	Geometry	Data Analysis and Statistics	Algebra and Functions
	0.54	0.11	–	0.16	0.15
	0.48	0.70	–	0.08	–
	0.45	0.19	–	0.41	0.07
	0.36	–	0.08	–	0.57
	0.36	0.23	–	0.06	0.10
	0.32	–	0.38	0.18	–
	0.28	0.37	0.02	0.14	0.10
	0.27	0.07	0.03	0.19	0.38
	0.25	–	0.37	0.23	0.25
	0.23	0.07	0.08	–	0.62
	0.23	0.27	0.14	–	0.36
	0.17	–	0.04	0.01	0.46
	0.12	0.03	0.29	0.17	0.18
	0.11	0.04	–	0.07	0.60
	0.10	0.41	0.15	0.33	–
	0.06	0.39	0.34	0.13	–
	0.01	0.32	0.24	0.42	–
	–	0.45	0.15	0.13	0.14
	–	0.42	0.28	0.61	–
	–	0.21	0.05	0.49	–
	–	0.51	0.17	–	0.46
	–	0.46	0.26	0.30	–
	–	0.47	0.06	0.32	–
Maryland	–	0.40	0.23	0.21	0.18
	–	0.24	0.32	0.36	0.08
	–	0.31	0.28	–	0.67

Note: " – " indicates that the standardized regression coefficient is less than zero.

Table 5. Standardized Beta Coefficients when Regressing School Mean State Assessment Scores with the 1998 NAEP Reading School Mean Grade 4 Scores of Each Content Area, by State.

State	Literary Experience	Information
	0.65	0.15
	0.55	0.29
	0.55	0.24
	0.52	0.43
	0.49	0.44
	0.42	0.29
	0.34	0.46
	0.34	0.39
	0.33	0.53
	0.32	0.49
	0.32	0.49
	0.31	0.23
	0.31	0.48
	0.30	0.50
	0.28	0.55
Maryland	0.26	0.64
	0.26	0.58
	0.23	0.56
	0.22	0.29
	0.22	0.47
	0.20	0.48
	0.09	0.50
	0.09	0.71
	0.07	0.45

Table 6. Standardized Beta Coefficients when Regressing School Mean State Assessment Scores with the 1998 NAEP Reading School Mean Grade 8 Scores of Each Content Area, by State.

State	Literary Experience	Information	Task
	0.40	0.30	0.18
	0.38	0.47	0.06
	0.38	0.44	0.16
	0.33	0.24	0.28
	0.33	0.49	–
	0.30	0.42	0.19
	0.30	0.23	0.34
	0.26	0.19	0.14
	0.25	0.39	0.19
	0.25	0.45	0.15
	0.22	0.11	0.47
	0.18	0.53	0.22
	0.18	0.56	0.15
	0.18	0.43	0.34
	0.17	0.34	0.23
	0.15	0.43	0.28
	0.15	0.29	0.20
	0.13	0.28	0.50
	0.12	0.23	0.46
	0.07	0.40	0.24
	–	0.47	0.41
Maryland	–	0.25	0.65
	–	0.40	0.52
	–	0.37	0.29
	–	0.31	0.18

Note: " – " indicates that the standardized regression coefficient is less than zero.

Differential sensitivity of NAEP and statewide assessment results to poverty

An important issue confronting American education is the gap in achievement outcomes between students from rich and poor families. Educational policymakers strive for schools that promote both excellence and equity. Statewide assessments and NAEP both provide information about the gap in achievement, but do they provide the same answers to the question? We can address this question using the percent of students eligible for the Federal free school lunch program as a measure of poverty.

The results in tables 7 and 8 show how large the gap is between schools in the NAEP sample serving rich and poor students, as measured by your state assessment and as measured by NAEP. Specifically, these tables show what fraction of a standard deviation corresponds to a difference of 20 percentage points in the count of free lunch eligible students in enrollment.[3] The entries are ordered from largest to smallest gap, as measured by NAEP.

For example, the first entry in table 7 (i.e., –1.36) indicates that in one state, a school with a particular percent free lunch eligible enrollment (e.g., 45%) would on average, have a 1.36 s.d. lower statewide assessment grade 4 mathematics mean than a school with 20 percent lower free lunch eligible enrollment (e.g., 25%). The figure to the right of the –1.36 indicates the corresponding gap as measured by NAEP, (i.e., 1.46 s.d.). That these two figures are very similar indicates that in that state NAEP and the statewide assessment are measuring virtually the same size gap. Examination of tables 7 and 8 shows that this is not true in all states.

The results in tables 7 and 8 show that, as measured by NAEP, the 20 percent difference in poverty has a similar association with achievement in many states, but a few states stand out with noticeably smaller gaps. Further analyses are needed to identify the source of the smaller gap, whether it might be due to a difference in the operationalization of the poverty measure or to a real breakthrough in educational equity.

[3] The "standard deviation" used in these analyses is the standard deviation of school means in the state, which for NAEP is typically 7 or 8 points.

Table 7. Difference in Standardized Mathematics Assessment School Means Associated with a 20 Percent Increase in Free-Lunch Eligible Students, by State.

State	4th Grade State Assessment	4th Grade 1996 NAEP	State	8th Grade State Assessment	8th Grade 1996 NAEP
	-1.36	-1.46		-0.86	-1.39
	-0.77	-0.80		-1.44	-0.89
	-0.77	-0.79		-1.00	-0.87
	-0.78	-0.78		-0.88	-0.86
	-0.72	-0.74		-0.89	-0.83
	-0.42	-0.70	Maryland	-0.80	-0.77
	-0.63	-0.69		-0.79	-0.76
	-0.47	-0.69		-0.72	-0.75
Maryland	-0.59	-0.68		-0.88	-0.74
	-0.50	-0.66		-0.78	-0.74
	-0.56	-0.64		-0.84	-0.74
	-0.60	-0.64		-0.79	-0.73
	-0.31	-0.61		-0.43	-0.68
	-0.48	-0.56		-0.72	-0.67
	-0.50	-0.54		-0.52	-0.61
	-0.31	-0.52		-0.60	-0.56
	-0.38	-0.52		-0.59	-0.51
	-0.38	-0.38		-0.31	-0.33

Note: Entries are fractions of the standard deviation of school means in a state. NAEP and statewide assessment results in this table are for the same sample of schools.

Table 8. Difference in Standardized Reading Assessment School Means Associated with a 20 Percent Increase in Free-Lunch Eligible Students, by State.

State	4th Grade State Assessment	4th Grade 1996 NAEP	State	8th Grade State Assessment	8th Grade 1996 NAEP
	-0.87	-0.79		-1.75	-1.34
	-0.72	-0.76		-0.82	-0.88
	-0.75	-0.76		-0.80	-0.73
	-0.97	-0.72		-0.77	-0.73
	-0.72	-0.68	Maryland	-0.71	-0.72
	-0.74	-0.65		-0.83	-0.70
	-0.68	-0.65		-0.81	-0.68
	-0.75	-0.63		-0.70	-0.67
	-0.62	-0.63		-0.78	-0.66
Maryland	-0.60	-0.62		-0.75	-0.66
	-0.62	-0.61		-0.66	-0.64
	-0.70	-0.60		-0.88	-0.62
	-0.49	-0.58		-0.67	-0.58
	-0.37	-0.57		-0.54	-0.57
	-0.68	-0.57		-0.59	-0.57
	-0.54	-0.55		-0.65	-0.51
	-0.40	-0.55		-0.89	-0.47
	-0.52	-0.44		-0.58	-0.45
	-0.54	-0.44		-0.69	-0.18

Note: Entries are fractions of the standard deviation of school means in a state. NAEP and statewide assessment results in this table are for the same sample of schools.

Differential sensitivity of NAEP and statewide assessment results to minority concentrations

Another important issue confronting American education is the gap in achievement outcomes between students from white and minority families.[4] Measurement of progress towards equal opportunity in education was one of the founding goals of NAEP, and the interpretation of test scores has been confounded by race/ethnic differences throughout the 20th century. Statewide assessments and NAEP both provide information about the gap in achievement, but do they provide the same answers to the question?

The results in tables 9 and 10 show how large the gap is between NAEP schools serving white and minority students is in your state, as measured by your state assessment and as measured by NAEP. Specifically, these tables show what fraction of an achievement standard deviation corresponds to a difference of 20 percent in the count of minority students in enrollment. The entries are ordered from largest to smallest gap, as measured by NAEP.

For example, the first entry in table 9 (i.e., –0.47) indicates that in one state, a school with a specified percent minority enrollment (e.g., 60%) would, on average, have a 0.47 s.d. lower statewide assessment grade 4 mathematics mean than a school with a 20 percent lower percent minority enrollment (e.g., 40%). The "standard deviation" used in these analyses is the standard deviation of school means. The figure to the right of the –0.47 indicates the corresponding gap as measured by NAEP, (i.e., –0.91 s.d.). In this case, the gap as measured by NAEP is noticeably larger than the gap as measured by the statewide assessment. Examination of tables 9 and 10 shows that this is not true in all states.

The results in tables 9 and 10 show that, as measured by NAEP, the 20 percent difference in minority has a similar association with achievement in nearly all of the states, but a few states stand out with much smaller gaps. Further analyses are needed to identify the source of the smaller gap, whether it might be due to a difference in patterns of minority enrollment or to real breakthroughs in equal educational achievement.

[4] The NAEP race/ethnic question, used for this analysis, has the following categories: White, Black, Hispanic, Native American, and Asian. For these analyses, all responses other than "White" are included in the "Minority" category.

Table 9. Difference in Standardized Mathematics Assessment School Means Associated with a 20 Percent Increase in Minority Enrollment, by State.

State	4th Grade State Assessment	4th Grade 1996 NAEP	State	8th Grade State Assessment	8th Grade 1996 NAEP
	-0.47	-0.91		-0.72	-0.79
	-0.66	-0.65		-0.77	-0.73
	-0.53	-0.64		-0.73	-0.73
	-0.26	-0.62		-1.00	-0.64
	-0.60	-0.62		-0.66	-0.62
	-0.48	-0.58		-0.66	-0.62
	-0.58	-0.57		-0.67	-0.59
	-0.50	-0.52		-0.66	-0.59
	-0.48	-0.52		-0.48	-0.56
	-0.48	-0.51		-0.57	-0.55
	-0.47	-0.51		-0.57	-0.54
	-0.47	-0.49		-0.55	-0.53
	-0.43	-0.49		-0.50	-0.52
	-0.52	-0.48		-0.52	-0.52
	-0.10	-0.47		-0.65	-0.51
	-0.38	-0.47		-0.38	-0.48
	-0.31	-0.46		-0.45	-0.47
	-0.20	-0.45		-0.09	-0.46
Maryland	-0.40	-0.44	Maryland	-0.49	-0.45
	-0.10	-0.39		-0.37	-0.45
	-0.27	-0.39		-0.49	-0.45
	-0.19	-0.36		-0.48	-0.43
	-0.21	-0.27		-0.39	-0.40
	-0.51	-0.23		-0.05	-0.32
				-0.26	-0.06

Note: Entries are fractions of the standard deviation of school means in a state. NAEP and statewide assessment results in this table are for the same sample of schools.

Table 10. Difference in Standardized Reading Assessment School Means Associated with a 20 Percent Increase in Minority Enrollment, by State.

State	4th Grade State Assessment	4th Grade 1998 NAEP	State	8th Grade State Assessment	8th Grade 1998 NAEP
	-0.92	-0.69		-0.83	-0.76
	-0.84	-0.68		-1.18	-0.74
	-0.57	-0.65		-0.63	-0.70
	-0.64	-0.63		-0.84	-0.67
	-0.43	-0.55		-0.80	-0.60
	-0.55	-0.52		-1.10	-0.57
	-0.54	-0.52		-0.57	-0.56
	-0.57	-0.52		-0.57	-0.54
	-0.59	-0.51		-0.62	-0.54
	-0.50	-0.50		-0.57	-0.53
	-0.52	-0.46		-0.52	-0.53
	-0.44	-0.46		-0.44	-0.51
	-0.58	-0.46		-0.58	-0.51
	-0.44	-0.46		-0.68	-0.50
	-0.19	-0.45		-0.66	-0.50
	-0.55	-0.45		-0.53	-0.48
	-0.64	-0.44		-0.45	-0.47
	-0.40	-0.44		-0.47	-0.46
	-0.55	-0.42		-0.47	-0.45
	-0.44	-0.42	Maryland	-0.39	-0.37
	-0.28	-0.42		-0.28	-0.30
Maryland	-0.37	-0.38		-0.47	-0.29
	-0.36	-0.27		-0.33	-0.25
				-0.03	-0.23

Note: Entries are fractions of the standard deviation of school means in a state. NAEP and statewide assessment results in this table are for the same sample of schools.

Differential sensitivity of NAEP and statewide assessment results to rural-urban differences

In many states, there are concerns about providing equal educational resources to urban and rural schools. The problems, constraints, and strengths of schools in rural and urban settings[5] are different, and state agencies can allocate resources to address gaps in achievement if they know those gaps to exist. An important question is whether different achievement tests would yield the same information about these gaps.

The results in tables 11 and 12 show how large the gap between students in rural and urban schools is in your state, as measured by your state assessment and as measured by NAEP. Specifically, these tables show what fraction of a standard deviation in achievement increment corresponds to a rural locale. The entries are ordered from smallest to largest rural/small town advantage, as measured by NAEP.

For example, the first entry in table 11 indicates that in one state, rural schools would, on average, have a 0.91 s.d. lower statewide assessment grade 4 mathematics mean than other schools. The "standard deviation" used in these analyses is the standard deviation of school means. The figure to the right of the –0.91 indicates the corresponding gap as measured by NAEP, (i.e., –1.10 s.d.). That these two figures are similar indicates that in that state NAEP and the statewide assessment are finding a similar gap between rural and urban schools. Examination of tables 11 and 12 shows that in many states the size of this gap depends on the assessment.

The results in tables 11 and 12 show that, as measured by NAEP, the rural-urban difference in achievement varies widely across states, from substantially favoring rural and small town schools in some states and substantially favoring schools in metropolitan areas and large towns in other states. These differences probably reflect differences in the nature of "rural" and "urban" communities between states, but they may also reflect differences in educational policies. Further analyses are needed to clarify the sources of urban-rural differences.

[5] "Rural" refers to locales with fewer than 2,500 residents and small towns with fewer than 10,000 residents. "Urban" refers to five other categories of locale, including large towns outside metropolitan areas, and large and midsize cities and their suburbs inside metropolitan areas.

Table 11. Difference in Standardized Mathematics Assessment School Means Associated with School Locale, by State.

State	4th Grade State Assessment	4th Grade 1996 NAEP	State	8th Grade State Assessment	8th Grade 1996 NAEP
	-0.91	-1.10		-0.73	-0.82
	-0.30	-0.52		-0.82	-0.59
	0.06	-0.52		-0.26	-0.59
	0.03	-0.49		-0.38	-0.58
	-0.64	-0.42		-0.59	-0.57
	-0.49	-0.42		-0.50	-0.56
	0.01	-0.36		-0.65	-0.56
	-0.04	-0.35		-0.14	-0.35
	-0.36	-0.30		-0.27	-0.32
	-0.21	-0.26		0.01	-0.21
	-0.50	-0.25		-0.47	-0.20
	-0.45	-0.20		-0.05	-0.16
	0.42	-0.15		-0.10	-0.11
	0.09	-0.10		-0.37	-0.06
	0.11	0.02		-0.20	-0.01
Maryland	0.33	0.08		0.35	0.06
	-0.06	0.11		0.15	0.06
	0.17	0.12		0.37	0.08
	-0.33	0.15		0.32	0.09
	-0.05	0.44		0.08	0.12
	0.12	0.44	Maryland	0.39	0.20
	0.30	0.46		0.30	0.44
	0.62	0.82		0.64	0.51
				0.49	0.68

Note: Entries are fractions of the standard deviation of school means in a state. NAEP and statewide assessment results in this table are for the same sample of schools.

Table 12. Difference in Standardized Reading Assessment School Means Associated with School Locale, by State.

State	4th Grade State Assessment	4th Grade 1998 NAEP	State	8th Grade State Assessment	8th Grade 1998 NAEP
	-0.97	-0.72		-0.94	-0.87
	-0.70	-0.64		-0.86	-0.71
	-0.23	-0.58		-0.54	-0.48
	-0.45	-0.50		-0.14	-0.46
	-0.52	-0.47		-0.15	-0.38
	-0.61	-0.46		-0.94	-0.36
	-0.46	-0.43		-0.38	-0.23
	-0.47	-0.38		-0.23	-0.19
	-0.40	-0.24		-0.21	-0.21
	-0.25	-0.12		-0.13	-0.09
	-0.24	-0.11		-0.12	-0.07
	0.01	-0.10		0.12	-0.05
	-0.18	-0.10		-0.15	0.07
	0.26	0.02	Maryland	0.07	0.39
Maryland	0.10	0.23		0.32	0.45
	0.10	0.22		0.23	0.46
	0.09	0.23		0.26	0.54
	0.59	0.32		0.51	0.55
	0.35	0.42		0.32	0.71
	0.00	0.55		0.06	0.85
	0.45	0.55		0.87	0.89
	0.69	0.74		0.77	0.93

Note: Entries are fractions of the standard deviation of school means in a state. NAEP and statewide assessment results in this table are for the same sample of schools.

Statewide school means associated with NAEP "proficient" achievement level percentages

NAEP has set three criterion "cutpoints" on its scales, minimum scores for "advanced" categorization, "proficient" categorization, and "basic" categorization of achievement. Most states have also set categorizations of performance on their statewide assessments, and there is some interest in the relations between these different categorizations of achievement. As a first step in linking the "achievement levels," we have matched mean statewide assessment scores and NAEP percentages of proficient achievement in the NAEP schools.

Tables 13, 14, 15, and 16 display the ranges of estimated percentages of proficient students for NAEP schools that are in (a) the lowest quartile, (b) the middle half, and (c) the highest quartile in terms of mean statewide assessment scores. Tables 13 and 14 are for mathematics and tables 15 and 16 are for reading. In each table, the states are sorted by the maximum NAEP-proficient percentage among schools in the middle half of the sample.

There are a great many numbers in these tables, and only one row in each table is immediately pertinent to your state. For example, the first row in table 13 has 6 entries: 2 each for the lowest quartile, middle half, and highest quartile of schools in one state. The first two entries in table 13, 0.0% and 37.5%, give the range of NAEP grade 4 mathematics proficiency percentages in the schools in the lowest quartile of the NAEP school sample in one state. The "quartiles" are defined based on statewide assessment means. That is, of approximately 100 public elementary schools participating in State NAEP in that state, the roughly 25 schools in the lowest quartile were estimated to have percentages of grade 4 students NAEP-proficient in mathematics ranging from 0.0% to 37.5%.

One expects to find higher percentages of NAEP proficient students in schools in the upper quartile in terms of statewide assessment results. However, as can be seen by examining the first row of table 13, there is a great deal of overlap in percentages of proficient students in schools whose statewide assessment means are in different quartiles. At least one school in the top quartile in this state had a percent proficient of only 1.5%, while at least one school in the bottom quartile had 37.5% proficient students. The major reasons for the overlap are (a) that NAEP statistics are based on the random sample of students participating in NAEP in the school and (b) that NAEP and the statewide assessment are not testing exactly the same skills.

Overall, the top quartile of schools generally had greater variation in NAEP proficiency than the bottom quartile did. This pattern could mean that NAEP is more sensitive to variations between schools in the upper part of the achievement distribution, while statewide assessments are more sensitive to variations in the lower part of the achievement distribution, which is consistent with state policies of orienting statewide assessments more to identify schools that need help than to identify outstanding performing schools. On the other hand, this pattern may just be an artifact of the fact that NAEP is a "low-stakes" test, on which students in some high achieving schools (according to the statewide assessment results) do not perform up to their potential.

Table 13. Grade 4 Mathematics Ranges of NAEP Percent Proficient, for Upper and Lower Quartiles and Middle Half of Schools, as Defined by Statewide Assessment Means.

	Lower Quartile		Middle Half of Schools		Upper Quartile	
	Percent Proficient		Percent Proficient		Percent Proficient	
	min	max	min	max	min	max
Maryland	0.0%	37.5%	0.0%	91.3%	1.5%	76.2%
	0.0%	34.8%	0.9%	71.8%	9.1%	58.6%
	0.0%	22.0%	0.0%	64.0%	0.0%	76.3%
	0.0%	21.0%	0.0%	60.1%	0.0%	70.0%
	0.0%	24.6%	0.7%	58.3%	16.9%	82.5%
	0.0%	16.1%	0.0%	56.1%	12.3%	82.8%
	4.9%	41.0%	6.0%	55.4%	12.1%	69.1%
	0.0%	32.6%	1.7%	55.4%	4.0%	50.1%
	0.0%	42.2%	1.0%	55.3%	0.0%	80.8%
	0.0%	36.1%	6.8%	53.9%	13.1%	66.7%
	2.0%	41.0%	7.1%	53.2%	14.2%	47.4%
	0.0%	35.6%	1.3%	50.9%	15.2%	64.4%
	0.0%	25.3%	0.0%	49.5%	2.8%	57.3%
	0.0%	35.1%	2.5%	46.4%	14.5%	62.0%
	0.0%	29.0%	0.0%	46.3%	15.5%	75.1%
	0.0%	35.8%	0.0%	44.6%	0.0%	49.8%
	0.0%	12.7%	0.0%	41.6%	11.8%	80.0%
	0.0%	26.1%	3.4%	41.0%	8.3%	76.5%
	0.0%	24.6%	1.0%	35.7%	12.6%	53.4%
	0.7%	24.1%	0.0%	34.0%	5.7%	70.0%
	0.0%	11.7%	0.0%	32.9%	3.3%	51.9%
	0.0%	21.5%	1.0%	32.8%	8.7%	64.6%
	0.0%	9.0%	0.0%	30.5%	1.5%	62.7%
	0.0%	31.6%	0.8%	28.8%	6.7%	66.1%
	0.0%	8.9%	0.0%	23.1%	0.0%	65.2%

Note: School quartiles are defined by the school mean on the statewide assessment.

Table 15. Grade 4 Reading Ranges of NAEP Percent Proficient, for Upper and Lower Quartiles and Middle Half of Schools, as Defined by Statewide Assessment Means.

	Lower Quartile		Middle Half of Schools		Upper Quartile	
	Percent Proficient		Percent Proficient		Percent Proficient	
	min	max	min	max	min	max
	0.0%	29.8%	15.7%	73.0%	30.4%	87.0%
	0.0%	46.0%	13.9%	72.4%	20.8%	65.5%
	0.0%	50.8%	20.8%	68.6%	9.2%	63.1%
	0.0%	44.1%	8.4%	67.7%	20.0%	87.2%
	4.3%	40.0%	10.0%	67.6%	19.6%	70.3%
Maryland	0.0%	26.8%	1.9%	66.6%	20.0%	69.3%
	10.1%	41.9%	13.0%	63.3%	23.1%	80.0%
	0.0%	60.8%	6.3%	61.7%	18.8%	72.8%
	0.0%	38.0%	9.0%	59.3%	22.7%	75.3%
	0.8%	25.8%	6.9%	58.2%	15.7%	63.0%
	1.0%	19.5%	1.6%	58.1%	1.4%	72.8%
	0.0%	38.9%	0.8%	57.9%	0.0%	61.5%
	0.0%	23.6%	3.7%	57.7%	10.1%	51.8%
	4.3%	36.3%	12.8%	55.3%	28.6%	63.2%
	0.0%	34.5%	6.3%	55.2%	27.0%	66.5%
	0.0%	16.6%	0.0%	55.1%	12.8%	61.5%
	2.4%	36.3%	6.1%	55.0%	18.3%	62.3%
	0.0%	21.1%	1.0%	52.0%	11.0%	53.6%
	0.0%	27.5%	1.9%	51.3%	18.8%	53.8%
	0.0%	16.3%	0.8%	51.1%	5.0%	51.8%
	0.0%	37.1%	5.2%	49.8%	16.3%	67.1%
	1.4%	30.0%	7.7%	49.6%	14.0%	63.1%
	0.0%	24.8%	2.1%	40.8%	9.3%	57.5%
	0.0%	25.7%	0.0%	37.1%	16.3%	76.5%

Note: School quartiles are defined by the school mean on the statewide assessment.

Conclusions

At the school level, most statewide reading and mathematics assessments are moderately or highly correlated with NAEP assessments of the same subject matter. In mathematics, at grade 4 most statewide assessments are more closely aligned with the "numbers and operations" subscale of NAEP, but at grade 8, the alignment is more likely to be closer to the "measurement" or "algebra and functions" subscale. NAEP and statewide assessments measure similar achievement gaps between rich and poor schools, and they agree in the finding that in a few states the gaps are noticeably smaller than in the majority of states. Statewide assessment school means are more closely related to NAEP school percentages of proficient students in some states than others; and in many states, NAEP appears to be more sensitive to performance variations at high performance levels, while statewide assessments are more focused on performance variations in low performing schools.

This initial report on the results of matching statewide assessment and NAEP school means is issued as a prototype, and with trepidation, because there are many sources of variation in the relations among test scores. States were asked for data in whatever format was easiest to provide, and the specific nature of the data varies from state to state. In some states we have used scale scores, in others average percentiles or the percent of students in the school who have satisfied a statewide performance criterion. In some cases, we were able to use matched student scores for the correlations, while in others published school means were used; and in a few cases, we found it necessary to use scores from an adjacent age cohort. These sources of variation may result in misleading conclusions, and we hope that you will explore with us the factors that might affect the relations between the results in your state and the results in other states.

Overall, we find these results encouraging, and we hope that you do also. In the future, we plan to explore the relations between statewide assessments and NAEP in greater depth and breadth, focusing on the topics that you find important. In particular, we have not addressed issues about the measurement of trends over time by NAEP and statewide assessments and about effects of exclusions and accommodations on comparisons, either over time or between states. Ultimately, it is our hope that NAEP will facilitate the use and interpretation of statewide assessments in the service of your state's educational policy development.

Appendix

List of state assessment measures used in this report

State	Math Grade 4 96	Math Grade 8 96	Reading Grade 4 98	Reading Grade 8 98
	SAT9 scale	SAT9 scale	SAT9 NCE	SAT9 NCE
	SAT9 %ile	SAT9 %ile	SAT9 %ile	SAT9 %ile
	STAR 1998 scale	STAR 1998 scale	STAR 1998 scale	STAR 1998 scale
	CMT (1)	CMT (1)	CMT scale	CMT scale
		DSTP scale		DSTP scale
		FCAT scale	FCAT scale	FCAT scale
	ITBS-g3 scale	ITBS scale	ITBS-g3 scale	ITBS scale
	SAT8-g3 raw	SAT8 scale	SAT8-g3 raw	SAT8 scale
		ISAT scale		ISAT 1997 scale
	ISTEP-g3 %pass	ISTEP %pass		
		KIRIS %ile		CTBS index
	MEAP scale	MEAP scale	MEAP scale	MEAP scale
Maryland	MSPAP-g3 %excellent	MSPAP %excellent	MSPAP-g3 %exc'nt	MSPAP %excellent
	MEA scale	MEA scale	MEA scale	MEA scale
	MEAP scale		MEAP scale	
		MBST 1998 avg score		MBST 1998 avg score
	MAP-g5 scale	MAP scale		
	ITBS NCE	ITBS NCE	ITBS NCE	ITBS NCE
	Various NCE	various NCE	various NCE	various NCE
	NCEOGT scale	NCEOGT scale	NCEOGT scale	NCEOGT scale
		CTBS NCE		
		CTBS scale	CTBS scale	CTBS scale
	CTBS %ile		CTBS 1997 %ile	CTBS 1997 %ile
			OCCT g5 %satisftry	OCCT-g8 %satisfactory
	OSA-g3 1998 scale	OSA scale	OSA-g3 scale	OSA scale
			MAT7 scale	MAT7 scale
	BSAP scale	BSAP scale	BSAP scale	BSAP scale
	TCAP scale	TCAP scale	TCAP scale	TCAP scale
	TAAS NCE	TAAS NCE		
	VSAP-g3 1997 %ile	VSAP 1997 %ile	VSAP-g3 1997 %ile	SOL scale
	WASL 1977 %pass		WASL %pass	
	WKCE 1997 NCE	WKCE 1997 NCE	WKCE 1997 NCE	WKCE 1997 NCE
	SAT %ile		SAT 1997 %ile	

(1) At grade 4, total objectives point score; at grade 8, number of objectives mastered.

Note: These are not all of the measures used by individual states in their statewide assessments. They are measures selected from data provided, for use in this report. For additional information on state assessments, please consult the *Annual Survey of State Student Assessment Programs*, published by the Council of Chief State School Officers, Washington, DC.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").