Mapping 2005 State Proficiency Standards Onto the NAEP Scales

Research and Development Report

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
Institute of Education Sciences
NCES 2007–482

THIS PAGE INTENTIONALLY LEFT BLANK.

U.S. Department of Education
Institute of Education Sciences
NCES 2007–482

Mapping 2005 State Proficiency Standards Onto the NAEP Scales

Research and Development Report

June 2007

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

National Center for Education Statistics
Peggy G. Carr
Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

June 2007
The NCES World Wide Web Home Page address is http://nces.ed.gov.
The NCES World Wide Web Electronic Catalog is http://nces.ed.gov/pubsearch.

Suggested Citation
National Center for Education Statistics (2007). *Mapping 2005 State Proficiency Standards Onto the NAEP Scales* (NCES 2007-482). U.S. Department of Education. Washington, DC: Author.

For ordering information on this report, write to
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398
or call toll free 1-877-4ED-Pubs or online at http://www.edpubs.org.

Content Contact:
Taslima Rahman
202-502-7316
taslima.Rahman@ed.gov

Executive Summary

Under the No Child Left Behind Act (NCLB), states are required to report the percentages of students achieving proficiency in reading and mathematics for grades 3 through 8. For each subject and grade combination, the percentages vary widely across states. For grades 4 and 8, these percentages can be compared to the estimated percentages of students achieving proficiency with respect to the standard established by the National Assessment of Educational Progress (NAEP). Again, large discrepancies are observed. This variation could derive from differences in both content standards and student academic achievement from state to state, as well as from differences in the stringency of the standards adopted by the states. Unfortunately, there is no way to directly compare state proficiency standards because states are free to select the tests they employ and to establish their own performance standards.

This report presents the results of applying a methodology for mapping state proficiency standards in reading and mathematics onto the appropriate NAEP scale, employing data from the 2004–05 academic year. The mapping exercise was carried out for both grades 4 and 8. For each of the four subject and grade combinations, the NAEP score equivalents to the states' proficiency standards vary widely, spanning a range of 60 to 80 NAEP score points. Although there is an essential ambiguity in any attempt to place state standards on a common scale, the ranking of the NAEP score equivalents to the states' proficiency standards offers an indicator of the relative stringency of those standards.

There is a strong negative correlation between the proportions of students meeting the states' proficiency standards and the NAEP score equivalents to those standards, suggesting that the observed heterogeneity in states' reported percents proficient can be largely attributed to differences in the stringency of their standards. There is, at best, a weak relationship between the NAEP score equivalents for the state proficiency standard and the states' average scores on NAEP. Finally, most of the NAEP score equivalents fall below the cut-point corresponding to the NAEP *Proficient* standard, and many fall below the cut-point corresponding to the NAEP *Basic* standard.

These results should be employed cautiously, as differences among states in apparent stringency can be due, in part, to reasonable differences in the assessment frameworks, the types of item formats employed, and the psychometric characteristics of the tests. Moreover, there is some variation among states in the proportion of NAEP sample schools that could be employed in the analysis.

THIS PAGE INTENTIONALLY LEFT BLANK.

Foreword

The Research and Development (R&D) series of reports at NCES has been initiated to

- share studies and research that are developmental in nature. The results of such studies may be revised as the work continues and additional data become available;

- share the results of studies that are, to some extent, on the "cutting edge" of methodological developments. Emerging analytical approaches and new computer software development often permit new and sometimes controversial analyses to be done. By participating in "frontier research," we hope to contribute to the resolution of issues and improved analysis; and

- participate in discussions of emerging issues of interest to educational researchers, statisticians, and the Federal statistical community in general.

The common theme in all three goals is that these reports present results or discussions that do not reach definitive conclusions at this point in time, either because the data are tentative, the methodology is new and developing, or the topic is one on which there are divergent views. Therefore, the techniques and inferences made from the data are tentative and subject to revision. To facilitate the process of closure on the issues, we invite comment, criticism, and alternatives to what we have done. Such responses should be directed to

Marilyn Seastrom
Chief Statistician
Statistical Standards Program
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006-5651

THIS PAGE INTENTIONALLY LEFT BLANK.

Contents

List of Tables

Table

List of Figures

Figure

THIS PAGE INTENTIONALLY LEFT BLANK.

Introduction

Under the No Child Left Behind Act (NCLB), each state can select the tests and set the proficiency standards for reading and mathematics by which it determines its standing with respect to the requirements of adequate yearly progress (AYP). An apparent consequence is that the percentages of students deemed proficient vary widely across states for a given subject and grade. One explanation that has been offered is that this heterogeneity is largely due to differences in the stringency of the standards adopted by the states. Unfortunately, there is no way to directly compare state standards. However, for grades 4 and 8, percentages of students in states reaching proficiency can be compared to the estimated percentages of students achieving proficiency as defined by the National Assessment of Educational Progress (NAEP). When these comparisons are carried out, it is evident that there are substantial differences in the two sets of percentages.

In this report, the results of applying a methodology for mapping state proficiency standards in reading and mathematics onto the appropriate NAEP scale are presented, employing data from the 2004–05 academic year. The mapping exercise was carried out for both grades 4 and 8. Although there is an essential ambiguity in any attempt to place state standards on a common scale, the relative ranking of the NAEP score equivalents to the states' proficiency standards offers (a) a credible indicator of the relative stringency of the standards, and (b) a more useful basis for policy discussion than the differences in percentages referred to above.

McLaughlin and Bandeira de Mello (2002, 2003) created a methodology for mapping the state standards onto the NAEP scale, based on combining data from the National Longitudinal School-Level State Assessment Score Database[1] with data from NAEP. Braun and Qian (in press) proposed a modified version of this methodology and applied it to grades 4 and 8 data from 2000 for mathematics and 2002 for reading. For each state standard, the methodology yields a point on the NAEP scale that is an estimate of the NAEP equivalent to that standard. It also produces an estimate of the variance associated with the estimated NAEP equivalent.

This report presents the results of applying the methodology of Braun and Qian to state test data from the 2004–05 academic year and NAEP 2005 results. The NAEP score equivalents derived from this methodology facilitate inferences about how the states' standards for these four subject and grade combinations compare to each other, as well as to NAEP performance standards.

[1] The National Longitudinal School-Level State Assessment Score Database (NLSLSASD; www.schooldata.org) is constructed and maintained by the American Institutes for Research (AIR) for NCES. Its purpose is to collect and validate data from state testing programs across the country. It contains assessment data for approximately 80,000 public schools in the United States and is updated annually.

The difficulty in placing all states' standards on a common scale is due to a number of factors. Among them are differences among states (and with NAEP) in subject frameworks, assessment format, psychometric characteristics of the tests, and so on. Nonetheless, both McLaughlin and Bandeira de Mello and Braun and Qian conclude that most of the heterogeneity across states in the NAEP score equivalents can be attributed to differences in the stringency of the proficiency standards set by the states.

The report is organized as follows: Section 1 provides a brief description of the estimation method, including variance estimation. Section 2 describes the data resources employed, and Section 3 presents the results of the analysis. Section 4 provides discussion and conclusions. The text is augmented with four appendices: Appendix A describes the use of sample weights in NAEP; Appendix B treats the estimation of the variance of the estimated NAEP scale score equivalents; Appendix C provides supplementary plots; and Appendix D presents results based on an analysis of data from the 2003 assessments.

Section 1: Methodology

Outline of the Methodology

The procedure is carried out separately for each state. In the description that follows, mathematics data are used for illustrative purposes. An identical procedure was used for the reading data.

1) *Based on the proportions of students in NAEP-sampled schools who meet each state's performance standard on that state's assessment, estimate the proportion, P, of students in the state as a whole who meet the state's standard for proficiency*. The schools in each state's NAEP sample are identified and matched with their records in the National Longitudinal School-Level State Assessment Score Database (NLSLSASD). This database provides an estimate of the proportion of students meeting the state standard in each school. Using the school weights from the NAEP school sample design, a ratio estimate of P, $\overline{p}_w$, is derived.[2]
2) *Based on the NAEP sample of schools and students within schools, estimate the distribution of scores on the NAEP assessment for the state as a whole*. This is the procedure that is carried out to generate the results contained in the NCES report that follows each NAEP assessment. Let $\hat{F}$ denote the estimated distribution.[3]
3) *Find the point on the NAEP score scale at which the estimated proportion of students in the state scoring above that point equals the estimated proportion of students in the state meeting the state's own performance standard*. Using the results of 1 and 2, the NAEP equivalent to the state performance standard is obtained by finding the point y_{WAM} on the NAEP scale that is the $\left(1-\overline{p}_w\right)^{\text{th}}$ quantile:[4]

$$y_{\text{WAM}} = \hat{F}^{-1}\left(1-\overline{p}_w\right).$$

4) *Compute an estimate of the variance of the estimated NAEP score equivalent*. This computation uses standard NAEP methods to obtain variance estimates given NAEP's complex sample design and latent ability measurement procedures (Allen, Carlson, Johnson, & Mislevy, 2001).

Figure 1 illustrates the mapping procedure. The dashed curve on the left side represents an estimate of the state distribution of scores on the state test, based on all students in the

---

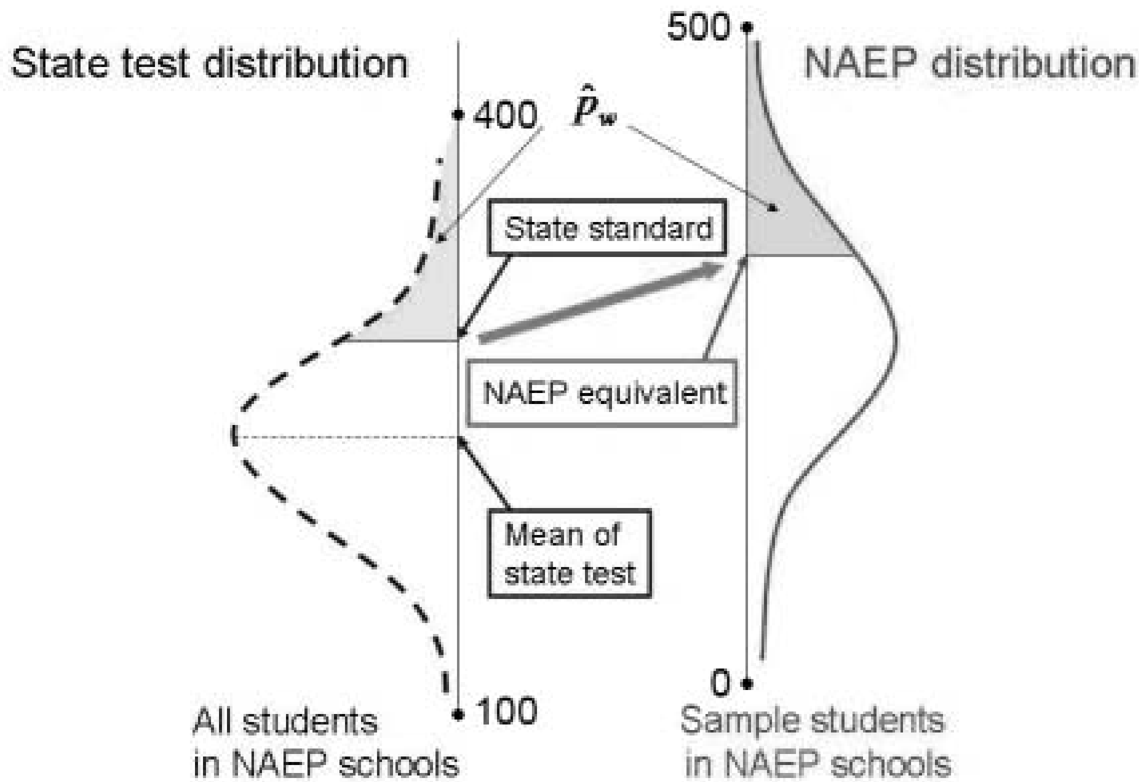
[2] Student weights are not required for this calculation because essentially all students in each NAEP sample school contribute to the estimation of *P*.

[3] For this calculation, both school and student weights are employed.

[4] WAM is an acronym for "weighted aggregate mapping," which is the term employed in Braun and Qian (in press) to distinguish this approach from the approach presented in McLaughlin and Bandeira de Mello (2003).

schools selected for the state's NAEP sample.[5] The area in the upper tail of this distribution above the state standard is an estimate of the proportion of students in the state meeting or exceeding that standard and is denoted by $\hat{p}_w$. In practice, only $\hat{p}_w$ need be obtained from the data. The curve on the right side represents the estimated distribution of NAEP scores for the state, based on the students in the state's NAEP sample schools who took the NAEP assessment. The estimated NAEP equivalent to the state standard, y_{WAM}, is the point on the NAEP scale, such that the corresponding upper tail area of the NAEP distribution also equals $\hat{p}_w$. Or, in other words, the estimated NAEP equivalent to the state standard is the point at the upper end of the state's NAEP score range that yields a proportion of the distribution above that point equal to the proportion of students with proficient scores on the state test.

Figure 1. The schematic of the mapping procedure



The above description is accurate in the ideal situation in which all the NAEP sample schools are available for estimating $\hat{p}_w$. In actuality, for some of the NAEP sample schools, information on the proportion of students meeting the proficiency standard is unavailable. Accordingly, for the estimation procedures described in 1 and 2 above, the subset of NAEP sample schools with information on the proportion of students meeting the proficiency

[5] The scale endpoints (100 and 400) are for illustrative purposes only and are intended to highlight the fact that the state test score scale and the NAEP score scale are different.

standard on the state test was employed. (In the remainder of this report, this subset of schools is referred to as having complete data.)

That not all NAEP sample schools can be employed in the analysis accounts for the need to estimate both $\hat{p}_w$ and the NAEP score distribution from precisely those schools that contribute data to the analysis, as well as the need to use the school-level design weights for estimation. With this strategy, the results should be relatively unaffected by any defects in the school sample. Braun and Qian (in press) document that for most states $\hat{p}_w$ is very close to the proportion proficient reported by the state.

NAEP Design Weights

State NAEP samples are obtained through a two-stage probability sampling design. The first stage constitutes a probability sample of schools containing the relevant grade. The second stage involves the selection of a random sample of students within each school. To account for the unequal probabilities of selection and to allow for adjustments for nonresponse, each school and each student was assigned separate sampling weights.[6] If these weights are not employed in the computation of the statistics of interest, the resulting estimates can be biased. With this caution in mind, appropriate school weights were applied in the estimation of the proportion of students above the standard on the state assessment. In general, the school weight equals the inverse of the approximate school selection probability. Appendix A provides a more detailed description of the sampling weights.

Because school weights are not retained in the NAEP database, for this study the school weights were computed in two steps. First, the sum of the student design weights for each school was calculated, and then this sum was divided by the number of grade-eligible students.[7] Details of the creation of school design weights for NAEP can be found in *NAEP 1998 Technical Report* (Qian, Kaplan, Johnson, Krenzke, & Rust, 2001, chapter 11). The student weights and the estimated school weights are used to estimate the distribution of NAEP scores in a state, employing information from the NAEP sample schools with complete data.

The Ratio Estimator for the Target Proportion. Let P_k be the proportion of students achieving the standard at school *k,* and w_k be the corresponding school weight. Further, let M_k be the number of students who were grade-eligible at school *k* (including all students with disabilities and English language learners). The total number of students meeting the standard is $\sum_{l=1}^{N} P_l \cdot M_l$, where *N* is the total number of public schools in the

[6] Students with disabilities and English language learners who cannot be assessed, even with the accommodations that NAEP provides, are not considered nonrespondents, but are excluded from the population of inference. Their performance is not included in estimates of the NAEP score distributions.
[7] Note that this calculation was carried out only for the subset of NAEP sample schools with complete data. School and student weights were not adjusted for schools lost from the NAEP school sample due to the unavailability of state test performance data.

state containing the relevant grade. The statewide target proportion of students meeting the standard is approximately

$$P = \frac{\sum_{l=1}^{N} P_l \cdot M_l}{\sum_{l=1}^{N} M_l}.$$

Using Horvitz–Thompson estimators (Cochran, 1977), the numerator and denominator of P are estimated separately from the state's NAEP school sample. For example, $\sum_{l=1}^{n} w_l M_l$ estimates the total number of eligible students in the state, and $\sum_{l=1}^{n} w_l \left(P_l \cdot M_l\right)$ estimates the total number of students meeting the standard. The target proportion, P, of students meeting the standard can be estimated by a ratio estimator:

$$\overline{p}_w = \frac{\sum_{l=1}^{n} w_l \left(P_l \cdot M_l\right)}{\sum_{l=1}^{n} w_l M_l}.$$

When survey variables are observed without error from every respondent to a stratified and clustered sample such as NAEP, the usual complex-sample variance estimators quantify the uncertainty associated with sample statistics (Skinner, Holt, & Smith, 1989). Since a specific NAEP score is not assigned to individual students participating in the NAEP assessments (even to those who responded to the cognitive items), additional statistical analyses are required to properly quantify the uncertainty associated with inferences about score distributions (Allen et al., 2001). For details of variance estimation for NAEP, see Appendix B.

Section 2: Data Resources

Data from the 2005 NAEP assessment[8] were used to implement the mapping of state proficiency standards onto the NAEP scale. The data were extracted from the 2005 NAEP restricted-use data set.[9]

Information on the proportions of students meeting state test standards for 2004–05 was retrieved from the NLSLSASD. Typically, the NLSLSASD presents for each school the percent of students meeting or exceeding each achievement standard established by the

---

[8] Note that in 2005 there were no separate state and national samples. Since 2002, national results have been reported using the aggregate of the states.

[9] Licenses and more information about the restricted-use data sets may be obtained from NCES by visiting the following website: http://nces.ed.gov/nationsreportcard/researchcenter/license.asp.

state.[10] For this project, the achievement standard that best matched the percent proficient reported by the state to the U.S. Department of Education was chosen as the appropriate standard to be mapped to the NAEP scale.[11]

Some states were not in the database and, consequently, the number of jurisdictions (the 50 states and the District of Columbia) represented in the results displayed in the next section ranges from 32 to 36, depending on the subject and grade combination. Details are provided in the Results section.

Section 3: Results

Tables 1–4 contain the grade 4 and grade 8 results for reading and mathematics, respectively. For each state included in the analysis, each table displays the number of schools in the NAEP sample and the number of schools employed in the mapping. This last quantity is simply the number of schools in the NAEP sample that could be matched to the schools with usable state test performance data. Each table also displays an estimate of the statewide percent proficient, the estimate of the NAEP score equivalent to the state's standard, and the estimated standard error of the NAEP score equivalent.

[10] For almost all states, some schools in the NAEP school sample were either missing from the NLSLSASD, or the required datum was not listed. In those cases, the number of schools available for estimation is smaller than the number of schools in the NAEP school sample. For each subject and grade combination, there were six to eight jurisdictions in which the proportion of NAEP sample schools employed in the estimation was less than 0.9.

[11] The U.S. Department of Education compiles a database that contains all the information provided by the states in compliance with the regulations of No Child Left Behind. This information was not employed in the estimation procedure but only used for validation purposes. Three states (Maine, New York, and Ohio) apparently did not report the statewide percent proficient, so it was not possible to carry out the check. The data available to the Department of Education do not contain school-level data required to carry out the methodology.

Table 1. Results of mapping state standards to the grade 4 NAEP reading scale: 2005

State	State name	Number of schools in the NAEP sample	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska[1]	157	97	0.79	182	2.6
AR	Arkansas	151	144	0.53	217	1.2
CA	California	445	421	0.48	210	0.9
CO	Colorado	147	135	0.86	186	1.6
CT	Connecticut	132	132	0.66	212	1.0
FL	Florida	169	159	0.71	202	1.0
GA	Georgia[1]	176	156	0.87	175	2.2
HI	Hawaii	132	131	0.56	205	1.1
IA	Iowa	130	125	0.77	197	1.2
ID	Idaho	157	148	0.87	185	2.9
IN	Indiana	138	138	0.72	199	1.1
KY	Kentucky	149	148	0.67	206	1.6
LA	Louisiana	136	134	0.65	198	2.0
MA	Massachusetts	202	199	0.48	234	0.8
MD	Maryland	125	123	0.82	187	1.4
MS	Mississippi	127	116	0.88	161	2.0
MT	Montana[1]	241	194	0.81	197	1.5
NC	North Carolina	175	168	0.82	183	1.6
ND	North Dakota[1]	261	194	0.76	204	0.8
NJ	New Jersey	135	134	0.81	191	1.6
NM	New Mexico[1]	161	135	0.50	208	1.2
NV	Nevada	120	113	0.48	212	1.4
NY	New York	190	186	0.71	207	1.5
OH	Ohio	201	198	0.77	199	1.9
OK	Oklahoma	176	175	0.82	182	1.8
SC	South Carolina	119	118	0.35	228	1.3
TN	Tennessee	139	137	0.88	170	2.3
TX	Texas	383	376	0.81	190	1.0
WA	Washington	136	133	0.80	197	1.6
WI	Wisconsin	169	169	0.83	189	1.8
WV	West Virginia	195	190	0.80	186	1.3
WY	Wyoming[1]	170	146	0.47	228	0.7

[1] The proportion of NAEP sample schools employed in the estimation was less than 0.9.

NOTE: NAEP reading cut scores at grade 4 are 208 for *Basic* and 238 for *Proficient.* The following states' grade 4 reading test data were not used in the analysis or received special treatment: ME and MI—results deleted due to discrepancies between state assessment data and the state document; CA and LA—reading data not available for the state assessment, so English Language Arts (ELA) data used; MA—reading data not available for state assessment, so "Language" data variable used; AZ, DC, DE, IL, KS, MN, MO, OR, PA, and VA—neither reading nor ELA data available in the state data file; AL, NH, RI, SD, UT, and VT—state assessment data not available; NE—state results are based on assessments developed by each local education agency.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Table 2. Results of mapping state standards to the grade 8 NAEP reading scale: 2005

State	State name	Number of schools in the NAEP sample	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska[1]	102	54	0.82	230	1.2
AR	Arkansas[1]	125	112	0.57	254	1.2
AZ	Arizona	132	125	0.63	244	1.3
CA	California	374	356	0.39	262	0.8
CO	Colorado	120	108	0.86	229	2.1
CT	Connecticut	106	102	0.77	242	1.7
DC	District of Columbia[1]	42	28	0.44	244	0.9
DE	Delaware[1]	43	37	0.80	242	0.9
FL	Florida	161	155	0.44	265	1.5
GA	Georgia	124	116	0.83	224	2.2
HI	Hawaii	67	64	0.37	262	1.4
IA	Iowa	111	109	0.72	250	1.0
ID	Idaho	101	93	0.82	235	2.5
IL	Illinois	190	187	0.72	245	1.2
IN	Indiana	107	105	0.66	249	1.5
KS	Kansas	117	114	0.78	242	1.4
LA	Louisiana	112	110	0.54	251	1.4
MD	Maryland	107	105	0.68	245	1.7
MS	Mississippi	115	104	0.58	247	1.4
NC	North Carolina	139	132	0.88	217	1.5
ND	North Dakota[1]	182	134	0.72	255	0.9
NJ	New Jersey	111	110	0.74	250	1.3
NM	New Mexico[1]	106	86	0.52	251	1.2
NY	New York	182	173	0.49	268	1.1
OH	Ohio	142	135	0.80	241	1.5
OK	Oklahoma	147	142	0.71	244	1.9
OR	Oregon	119	116	0.64	254	1.3
PA	Pennsylvania	110	104	0.64	258	1.7
SC	South Carolina	108	104	0.30	276	1.3
TN	Tennessee	112	111	0.87	222	1.5
TX	Texas	278	270	0.83	225	1.0
WI	Wisconsin	118	117	0.86	229	2.1
WV	West Virginia	110	107	0.80	228	1.7
WY	Wyoming	78	77	0.39	278	1.2

[1] The proportion of NAEP sample schools employed in the estimation was less than 0.9.
NOTE: NAEP reading cut scores at grade 8 are 243 for *Basic* and 281 for *Proficient.* The following states' grade 8 reading test data were not used in the analysis or received special treatment: ME, MT, and VA—discrepancies exist between the state assessment data and the state document; CA and LA—reading data not available for state assessment, so ELA data used; KY, MA, MI, MN, MO, NV, and WA—neither reading nor ELA data available in the state data file; AL, NH, RI, SD, UT, and VT—state assessment data not available; NE—state results are based on assessments developed by each local education agency.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Table 3. Results of mapping state standards to the grade 4 NAEP mathematics scale: 2005

State	State name	Number of schools in the NAEP sample	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska[1]	153	108	0.71	222	1.4
AR	Arkansas	151	144	0.53	236	1.0
CA	California	446	421	0.51	231	0.7
CO	Colorado	146	135	0.90	201	1.7
CT	Connecticut	132	132	0.78	221	1.0
FL	Florida	169	159	0.63	230	0.8
GA	Georgia[1]	176	156	0.75	215	1.4
HI	Hawaii	132	131	0.30	247	1.2
IA	Iowa	130	124	0.80	219	1.1
ID	Idaho	158	148	0.91	207	1.9
IN	Indiana	138	138	0.72	225	1.1
KS	Kansas	139	134	0.85	218	1.4
LA	Louisiana	136	134	0.63	223	1.0
MA	Massachusetts	202	200	0.39	255	1.0
MD	Maryland	125	124	0.78	215	1.1
MI	Michigan	141	131	0.73	222	1.7
MO	Missouri	159	158	0.41	242	1.2
MS	Mississippi	127	117	0.79	206	1.3
NC	North Carolina	175	168	0.91	203	1.2
ND	North Dakota[1]	261	194	0.80	224	0.8
NJ	New Jersey	135	134	0.81	221	1.3
NM	New Mexico[1]	162	135	0.39	233	1.3
NV	Nevada	118	112	0.52	230	0.9
NY	New York	190	186	0.87	207	1.5
OH	Ohio	201	199	0.65	233	1.3
OK	Oklahoma	177	175	0.74	218	0.9
SC	South Carolina	119	118	0.39	246	1.2
TN	Tennessee	139	137	0.87	200	1.6
TX	Texas	382	376	0.82	219	1.0
WA	Washington	136	133	0.60	236	1.1
WI	Wisconsin	169	169	0.74	225	1.4
WV	West Virginia	195	190	0.75	215	1.1
WY	Wyoming[1]	164	146	0.39	251	0.7

[1] The proportion of NAEP sample schools employed in the estimation was less than 0.9.

NOTE: NAEP mathematics cut scores at grade 4 are 214 for *Basic* and 249 for *Proficient.* The following states' grade 4 mathematics test data were not used in the analysis or received special treatment: ME and MT—discrepancies exist between the state assessment data and the state document; AZ, DC, DE, IL, KY, MN, OR, PA, and VA—data not available in the file; AL, NH, RI, SD, UT, and VT—state assessment data not available; NE—state results are based on assessments developed by each local education agency.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Table 4. Results of mapping state standards to the grade 8 NAEP mathematics scale: 2005

State	State name	Number of schools in the NAEP sample	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska[1]	101	59	0.65	268	0.9
AR	Arkansas[1]	125	112	0.34	288	1.0
AZ	Arizona	131	125	0.61	265	1.1
CO	Colorado[1]	121	108	0.74	258	1.6
CT	Connecticut	106	102	0.76	257	2.3
DC	District of Columbia[1]	42	28	0.40	252	1.4
DE	Delaware[1]	43	37	0.56	275	1.0
FL	Florida	162	155	0.58	269	1.3
GA	Georgia	124	116	0.69	255	1.2
HI	Hawaii	67	64	0.20	296	1.2
IA	Iowa	111	109	0.76	262	1.1
ID	Idaho	103	93	0.70	266	1.7
IL	Illinois	190	187	0.54	276	1.5
IN	Indiana	107	105	0.70	266	1.5
KY	Kentucky	117	115	0.37	285	1.4
LA	Louisiana	112	110	0.56	264	1.6
MA	Massachusetts	131	128	0.42	301	1.3
MD	Maryland	107	105	0.53	276	1.7
MI	Michigan	116	111	0.61	269	1.9
MO	Missouri	131	129	0.15	311	1.4
MS	Mississippi	115	104	0.53	262	1.5
NC	North Carolina	140	133	0.84	247	1.2
ND	North Dakota[1]	184	135	0.65	277	1.1
NJ	New Jersey	111	110	0.64	273	1.4
NM	New Mexico[1]	106	86	0.24	287	1.8
NY	New York	182	173	0.56	275	0.9
OH	Ohio	143	136	0.63	274	1.1
OK	Oklahoma	148	142	0.67	258	1.0
OR	Oregon	119	116	0.65	269	1.4
PA	Pennsylvania	110	104	0.62	272	1.1
SC	South Carolina	107	104	0.24	305	1.1
TN	Tennessee	112	111	0.88	230	1.6
TX	Texas	278	270	0.61	273	0.8
WI	Wisconsin	118	117	0.75	263	1.4
WV	West Virginia	110	107	0.71	253	1.1
WY	Wyoming	80	77	0.37	293	0.9

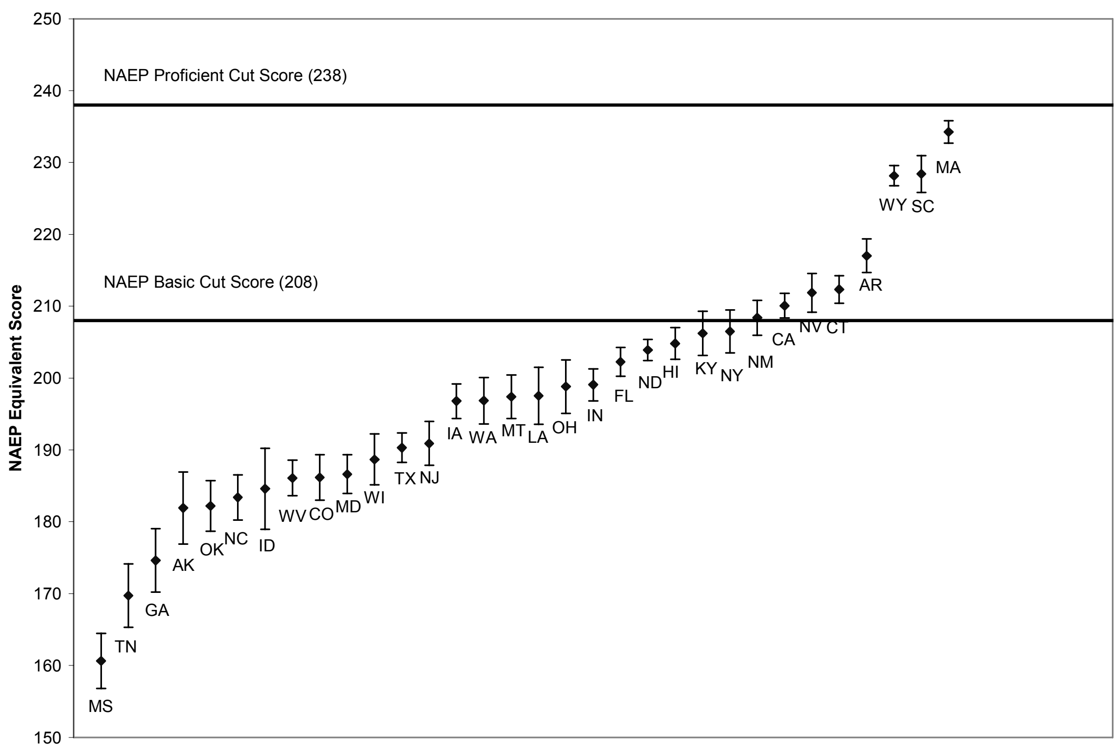
[1] The proportion of NAEP sample schools employed in the estimation was less than 0.9.
NOTE: NAEP mathematics cut scores at grade 8 are 262 for *Basic* and 299 for *Proficient.* The following states' grade 8 mathematics test data were not used in the analysis or received special treatment: ME, MT and VA—discrepancies exist between the state assessment data and the state document; CA, KS, MN, NV, and WA—data not available in the state assessment file; AL, NH, RI, SD, UT, and VT—state assessment data not available; NE—state results are based on assessments developed by each local education agency.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figures 2–5 display the ordered estimated NAEP score equivalents together with their estimated standard errors for the four subject and grade combinations (reading, grades 4 and 8; mathematics, grades 4 and 8). The estimated standard errors are relatively small compared to the range of the estimated NAEP score equivalents. The error bands in the figures extend plus or minus 1.96 standard errors on either side of the estimated NAEP score equivalent for the state.

Reading—Grade 4

There were 32 states in the grade 4 reading analysis. As shown in figure 2, the estimated NAEP score equivalents range from 161 (Mississippi) to 234 (Massachusetts), and the median estimated standard error is 1.5. As can be seen from figure 2, the margin of error for all but 10 of the estimated NAEP score equivalents falls below the cut-point of the NAEP *Basic* achievement level. There is also a negative correlation[12] of -0.88 (with a standard error of 0.094) between the estimated NAEP score equivalents and the statewide percents proficient; that is, the larger the NAEP score equivalent, the lower the percent of students in a state deemed proficient (see figure C-1 in Appendix C).

Figure 2. NAEP score equivalents of states' proficiency standards for reading, grade 4: 2005



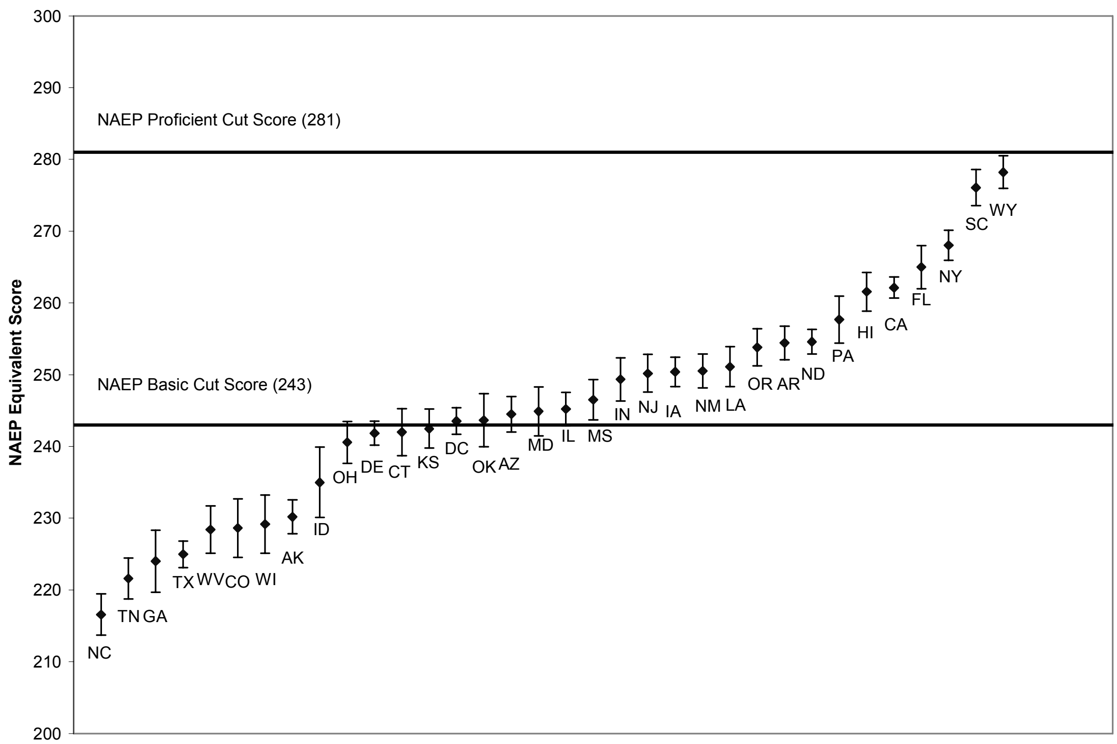
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

[12] Correlations calculated are standard Pearson correlations.

Reading—Grade 8

There were 34 states in the grade 8 reading analysis. As shown in figure 3, the estimated NAEP score equivalents range from 217 (North Carolina) to 278 (Wyoming), and the median estimated standard error is 1.4. As can be seen from figure 3, the margin of error for 9 of the estimated NAEP score equivalents falls below the cut-point of the NAEP *Basic* achievement level. There is also a negative correlation of -0.85 (with a standard error of 0.101) between the estimated NAEP score equivalents and the statewide percents proficient (see figure C-2 in Appendix C).

Figure 3. NAEP score equivalents of states' proficiency standards for reading, grade 8: 2005

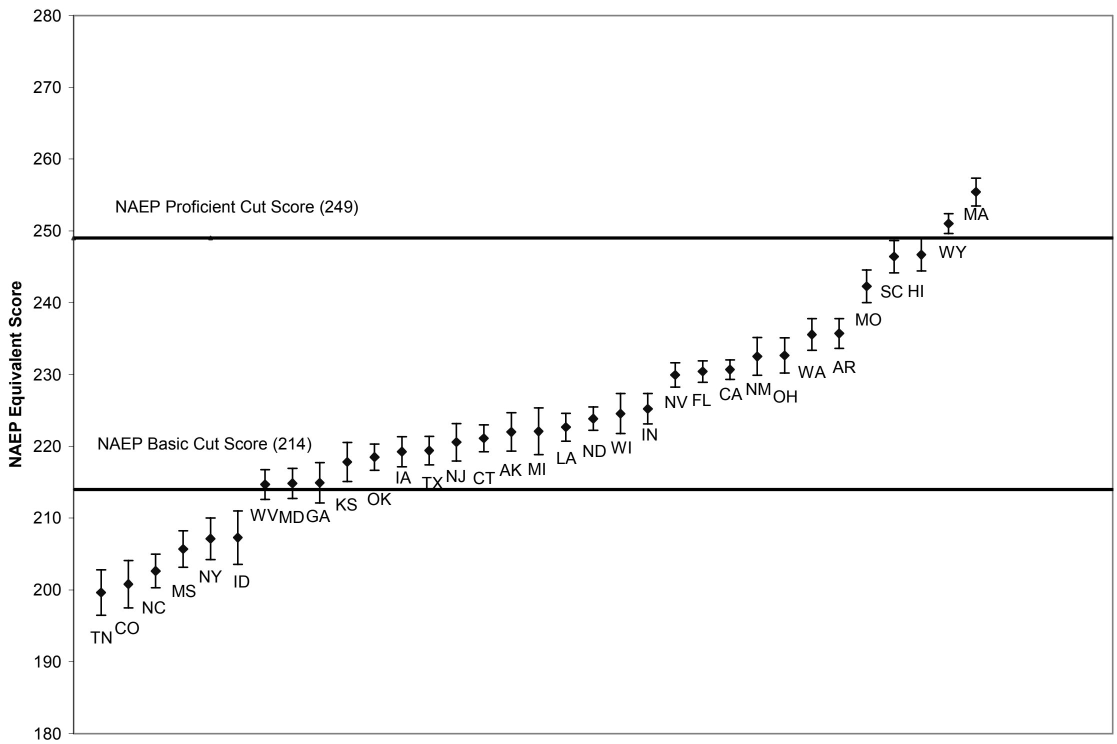


SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Mathematics—Grade 4

There were 33 states in the grade 4 mathematics analysis. As shown in figure 4, the estimated NAEP score equivalents range from 200 (Tennessee) to 255 (Massachusetts), and the median estimated standard error is 1.2. As can be seen from figure 4, the margin of error for 6 of the estimated NAEP score equivalents falls below the cut-point of the NAEP *Basic* achievement level, and 2 lie above the cut-point of the NAEP *Proficient* achievement level. There is also a negative correlation of -0.91 (with a standard error of 0.081) between the estimated NAEP score equivalents and the statewide percents proficient (see figure C-3 in Appendix C).

Figure 4. NAEP score equivalents of states' proficiency standards for mathematics, grade 4: 2005

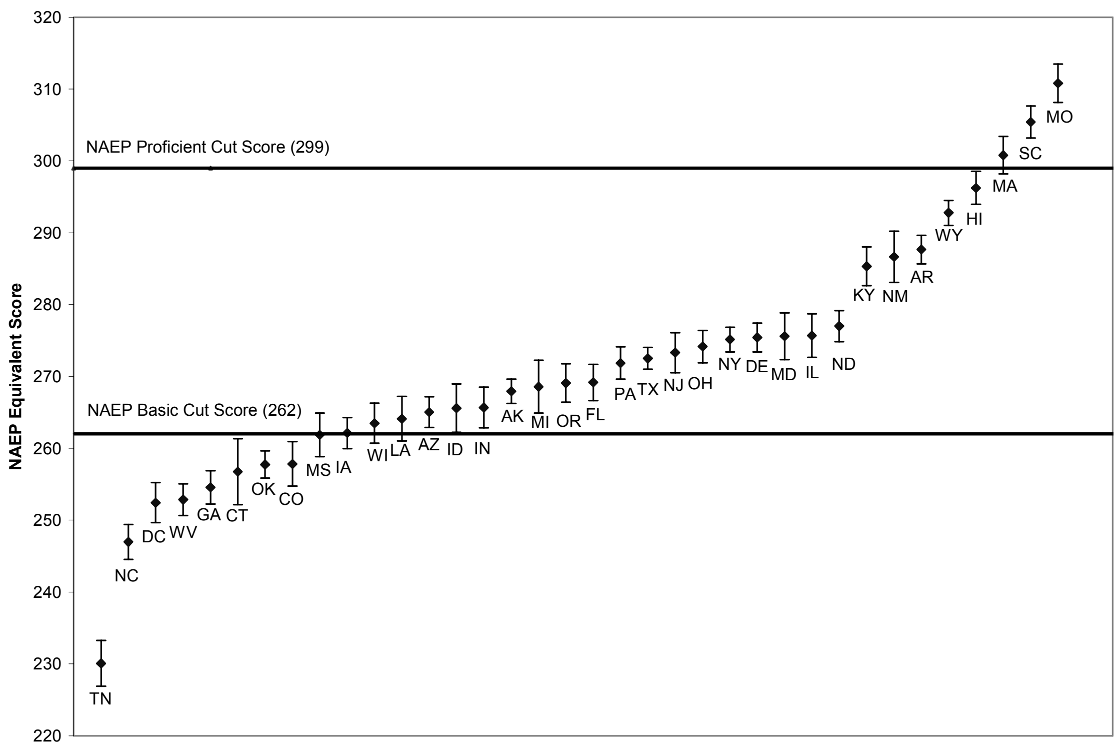


SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Mathematics—Grade 8

There were 36 states in the grade 8 mathematics analysis. As shown in figure 5, the estimated NAEP score equivalents range from 230 (Tennessee) to 310 (Missouri), and the median standard error is 1.3. As can be seen from figure 5, the margin of error for 8 of the estimated NAEP score equivalents falls below the cut-point of the NAEP *Basic* achievement level, and the margin of error for 2 lies above the cut-point of the NAEP *Proficient* achievement level. There is also a negative correlation of -0.83 (with a standard error of 0.101) between the estimated NAEP score equivalents and the statewide percents proficient (see figure C-4 in Appendix C).

Figure 5. NAEP score equivalents of states' proficiency standards for mathematics, grade 8: 2005



SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Appendix D contains the results of applying the mapping methodology described in Section 1 to state test data from the 2002–03 academic year and the NAEP assessments administered in the spring of 2003. As with the data from 2005, the analysis was carried out for reading and mathematics for both grades 4 and 8. Between 2002–03 and 2004–05, some states could have changed their tests, revised their standards, or improved the performance of their students. Thus, the results of the two analyses may not be directly comparable for some states. Nonetheless, for each subject and grade combination, the results of the two analyses are very similar; that is, for most states included in both analyses, the estimated NAEP score equivalents to the states' proficiency standards are numerically close.

Section 4: Discussion and Conclusions

The usefulness of these results depends on two arguments: first, that the estimated NAEP score equivalents are both well estimated and stable; second, that one can interpret the results as indicating that an important factor in explaining why two states have substantially different proportions of students meeting each state's proficiency standard is where they have set their standards, in addition to differences in the tests used and differences in the distributions of the relevant skills in their student populations.

With respect to the first argument, the estimated standard errors of the NAEP score equivalents are generally small in comparison to the range of the NAEP score equivalents. Stability would be best addressed by carrying out the mapping procedure using data only from students with particular characteristics (e.g., female students, African American students). Because of current data limitations, that is possible only for a few states. An alternative is to examine, for each state, the correlation between performance on the state test and on NAEP. This can be done at the school level. For example, using the NLSLSASD files, for each state one can compute the rank correlation across schools between the percent proficient on the state test and the estimated NAEP mean.[13] For grade 4 mathematics, the median correlation across the states in the analysis is 0.73, with an interquartile range of (0.63, 0.78). For grade 8 mathematics, the median correlation is 0.81, with an interquartile range of (0.71, 0.86). For grade 4 reading, the median correlation is 0.73, with an interquartile range of (0.63, 0.79). For grade 8 reading, the median correlation is 0.76, with an interquartile range of (0.62, 0.82). These figures are consistent with at least a modest degree of alignment for more than three-quarters of the states in each analysis. Ideally, the quantitative analysis should be supplemented by an intensive examination of the degree of alignment between the state test frameworks and the NAEP frameworks. This has not been done.

With respect to the second argument, the essential problem is that one must reason from the observed results (e.g., figures C-1–C-4) back to the situation with respect to the relative stringency of states' standards. The plausibility of the second argument is supported by the observation that there is a weak relationship between states' percent proficient and states' performance on NAEP. There is also a weak relationship between states' NAEP means and their NAEP score equivalents. Figures C-5–C-8 present scatter-plots of the state NAEP means and the state estimated NAEP score equivalents for reading (grades 4 and 8) and mathematics (grades 4 and 8). The correlations between where states set their proficiency standards and how they perform on NAEP in the four charts are 0.01 (with a standard error of 0.177) for grade 8 reading, 0.11 (with a standard error of 0.179) for grade 4 mathematics, 0.23 (with a standard error of 0.167) for grade 8 mathematics, and 0.27 (with a standard error of 0.176) for grade 4 reading. Note also that

[13] For the results that follow, schools with less than 5 students in the NAEP sample were excluded from the calculations.

the heterogeneity among the NAEP equivalents is much greater than among NAEP means.

For each subject, the stringency of the states' standards for proficiency appears to be articulated in the sense that the Spearman rank correlations between the states' NAEP score equivalents for grades 4 and 8 are 0.78 in reading (with a standard error of 0.090) and 0.77 in mathematics (with a standard error of 0.105). That is, for both reading and mathematics, states with apparently less stringent standards in grade 4 tend to have apparently less stringent standards in grade 8.

A recent report by Kingsbury, Olson, Cronin, Hauser, and Houser (2003) presents findings similar to those presented here. It describes an effort to map the proficiency standards for 12 states onto a common scale, which is used to report test scores for the Northwest Evaluation Association (NWEA) assessment battery. This exercise was carried out in both reading and mathematics for grades 3–10, employing data collected between 1999 and 2003.[14] The authors found substantial heterogeneity among the NWEA score equivalents of the state proficiency standards, as well as a strong negative correlation between the percent proficient and the NWEA score equivalent to the state's proficiency standard.

In view of the limitations of the data available, inferences concerning the NAEP score equivalents should be made with due caution. As indicated at the outset, in some states a number of schools in the NAEP sample could not be included in the analysis because the required state test data were not available at the individual school level. The loss of these schools could introduce some bias. In other states, the relevant state assessment was labeled "English/ Language Arts" rather than "Reading," so the degree of alignment between the two assessments could be lower than for other states. In any case, for each subject and grade combination, state assessment frameworks, as well as the test structures and item formats employed, will differ from those of the corresponding NAEP assessment. These differences can add noise to the comparisons with NAEP. Finally, states differ in the numbers and proportions of students with disabilities or English language learners that are excluded from either the state assessment or NAEP (or both). Such differences can also contribute to differences in the estimated NAEP score equivalents. Consequently, the estimated variance associated with each NAEP equivalent provides only a lower limit to the uncertainty to be associated with that value. At the same time, it is highly unlikely that the sources of bias discussed above could yield the broad range of NAEP score equivalents obtained.

[14] In contrast to the present situation, NWEA has available individual student scores on both the state test and the (common) NWEA scale.

References

Allen, N., Carlson, J.E., Johnson, E.G., & Mislevy, R.J. (2001). Scaling procedures, chapter 12. In N. Allen, J. Donoghue, & T. Schoeps (Eds.), *The NAEP 1998 technical report* (NCES 2001–509). Washington DC: National Center for Education Statistics.

Braun, H.I., & Qian, J. (in press). *Mapping state performance standards onto the NAEP scale*. Princeton, NJ: Educational Testing Service.

Cochran, W.G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley & Sons.

Kingsbury, G.G., Olson, A., Cronin, J., Hauser, C., & Houser, R. (2003). *The state of state standards: Research investigating proficiency levels in fourteen states.* Portland OR: Northwest Evaluation Association.

McLaughlin, D., & Bandeira de Mello, V. (2002, April). *Comparison of state elementary school mathematics achievement standards, using NAEP 2000*. Paper presented at the American Educational Research Association Annual Meeting, New Orleans.

McLaughlin, D., & Bandeira de Mello, V. (2003, June). *Comparing state reading and math performance standards using NAEP*. Paper presented at the National Conference on Large-Scale Assessment, San Antonio.

Qian, J., Kaplan, B., Johnson, E., Krenzke, T., & Rust, K. (2001). State weighting procedures and variance estimation, chapter 11. In N. Allen, J. Donoghue, & T. Schoeps (Eds.), *The NAEP 1998 technical report* (NCES 2001–509). Washington, DC: National Center for Education Statistics.

Rubin, D.B., & Schenker, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association*, *81*, 366-374.

Skinner, C.J., Holt, D., & Smith, T.M.F. (1989). *Analysis of complex surveys*. New York: John Wiley.

Wolter, K. (1985). *Introduction to variance estimation*. New York: Springer-Verlag.

Appendix A

Sampling Weights in NAEP

Formally, let *N* be the total number of schools in a state and M_k be the number of students who were grade-eligible at school *k*. Therefore, the total number of eligible students in the state is $\sum_{l=1}^{N} M_l$. Let *n* be the number of schools in the state NAEP sample. Let π_k be the school selection probability, which is proportional to its size M_k, and let $\pi_{i|k}$ be the conditional probability of selection for student *i* in school *k*. Suppose that *b* students are randomly selected from school *k*. Then the unconditional selection probability of student *i* in school *k* is

$$\pi_{ki} = \pi_k \cdot \pi_{i|k} = \frac{a \cdot M_k}{\sum_{l=1}^{N} M_l} \cdot \frac{b}{M_k},$$

where *a* is a constant of normalization. Then the weight of student *i* in school *k* is

$$w_{ki} = w_k \cdot w_{i|k} = \frac{1}{a \cdot M_k / \sum_{l=1}^{N} M_l} \cdot \frac{1}{b / M_k}.$$

This formula is only an approximation because students are selected without replacement, and the vicissitudes of fieldwork necessitate modifications to the ideal weights. For example, nonresponse adjustments to the weights are employed in NAEP to account for effects of schools and students who were selected but did not participate. Weights are not used to account for the effects of students with disabilities and English language learners who were unable to participate in NAEP, even with accommodations. Such students are not part of NAEP's population of inference. In any case, the weight of school *k* in a state NAEP sample is approximately

$$w_k = \frac{1}{a \cdot M_k / \sum_{l=1}^{N} M_l},$$

which equals the inverse of the approximate school selection probability. Because the weights in NAEP samples reflect the effects of oversampling, nonresponse adjustments, and trimming, the actual school weight will differ somewhat from $\sum_{l=1}^{N} M_l / (a \cdot M_k)$.

THIS PAGE INTENTIONALLY LEFT BLANK.

Appendix B

Estimating the Variance of the Estimated NAEP Score Equivalents

a. Two Components of Total Variance

The approach to variance estimation is based on the procedures developed by NAEP for the estimation of the variances of reporting statistics (Allen et al., 2001). The total variance of the estimate of the NAEP score equivalent of a state standard consists of two components: (a) the error due to sampling schools and students and (b) the uncertainty in the estimate of the distribution of the state's performance on NAEP based on data obtained from assessed students in the NAEP sample. The sampling error is estimated by the jackknife replicate resampling procedure (JRR) applied both to schools (for the state data) and to students (for the NAEP data). The uncertainty associated with the estimation of the NAEP distribution is estimated by utilizing the variability among the sets of plausible values generated for the sample.

b. The NAEP Jackknife Replicate Resampling (JRR) Approach

The JRR procedure for NAEP involves the formation of a large number of strata, typically consisting of pairs of schools. In NAEP, there are usually 62 strata. For the *j*th replicate, one school in the *j*th stratum is randomly deleted, and an appropriate set of weights is computed. The calculation of the 62 jackknife replicate weights for NAEP state samples can be found in the *NAEP 1998 Technical Report* (Allen et al., 2001). See also Wolter (1985).

To implement the JRR in this setting, both the jackknife replicate weights for students and the jackknife replicate weights for schools are required. These are formed by the same procedure described in Section 12.4.1 of Allen et al. (2001). For the *j*th replicate, the *j*th jackknife replicate weights for schools are applied to estimate the corresponding proportion of students meeting the standard, $\overline{p}_{w,(j)}$. The corresponding NAEP score equivalent is $y_{\mathrm{WAM},(j)}$, the $\left(1-\overline{p}_{w,(j)}\right)$th quantile of the distribution of NAEP scores based on that same replicate and employing the same replicate weights for students. Finally, the estimated variance of y_{WAM} that is due to sampling is

$$v_J\left(y_{\mathrm{WAM}}\right)=\sum_{j=1}^{62}\left(y_{\mathrm{WAM},(j)}-y_{\mathrm{WAM}}\right)^2.$$

c. Estimation of Measurement Uncertainty and Total Variance

The uncertainty due to measurement is estimated by carrying out the estimation procedure outlined in the text for each of the *M* = 5 sets of plausible values. Let the

NAEP score equivalent of a state standard estimated by the *m*th set of plausible values be $y_{\text{WAM},m}$, $m = 1,\ldots, M$, and denote the mean of $y_{\text{WAM},m}$ by $\bar{y}_{\text{WAM},\cdot}$. Finally, let

$$B = \sum_{m=1}^{M} \frac{\left(y_{\text{WAM},m} - \bar{y}_{\text{WAM},\cdot} \right)^2}{M - 1}.$$

Then the total variance is estimated by

$$v_T\left(y_{\text{WAM}} \right) = v_J\left(y_{\text{WAM}} \right) + \left(1 + M^{-1} \right) B,$$

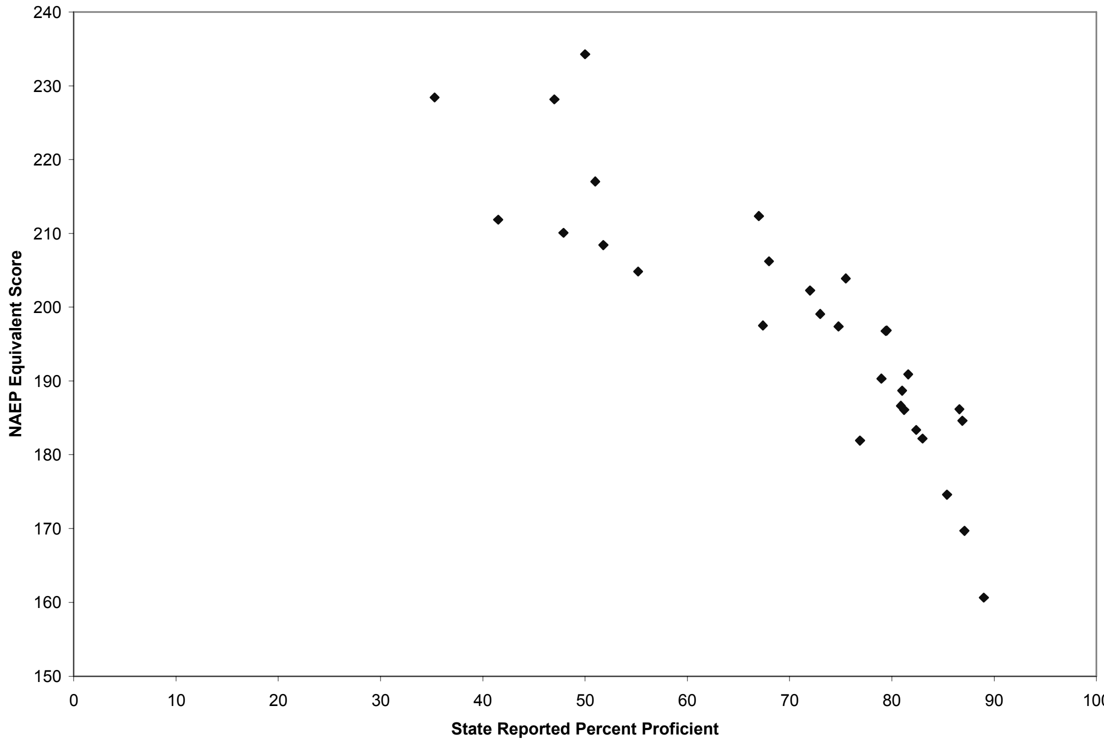
where $\left(1 + M^{-1} \right)$ is a finite adjustment factor (Rubin & Schenker, 1986). The estimation process mimics that of operational NAEP: The calculation of $v_J\left(y_{\text{WAM}} \right)$ is based on the first plausible value, and the estimation of *B* is based on all five plausible values. For further details, consult Allen et al. (2001).

Appendix C

Supplementary Plots

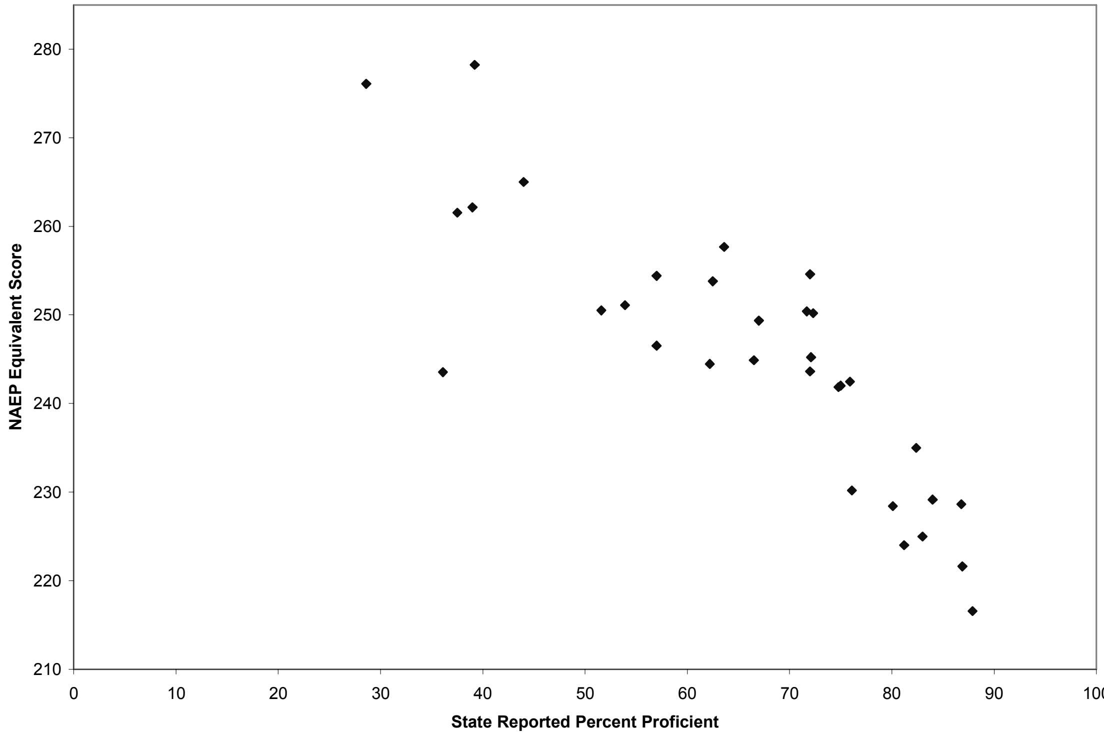
This appendix contains eight supplementary plots referred to in the main text. The first four plots display the relationship between states' NAEP score equivalents and states' reported percents proficient derived from the NLSLSAD for each subject and grade combination. In each plot, it is evident that there is a negative relationship between the two characteristics. The last four plots display the relationship between states' reported NAEP means and states' NAEP score equivalents for each subject and grade combination. In each plot, it is evident that there is a weak positive relationship between the two characteristics. Moreover, based on comparisons of the coefficients of variation, the heterogeneity among states' NAEP score equivalents is several times greater than that among states' reported NAEP means.

Figure C-1. NAEP score equivalent vs. state-reported percent proficient for reading, grade 4: 2005



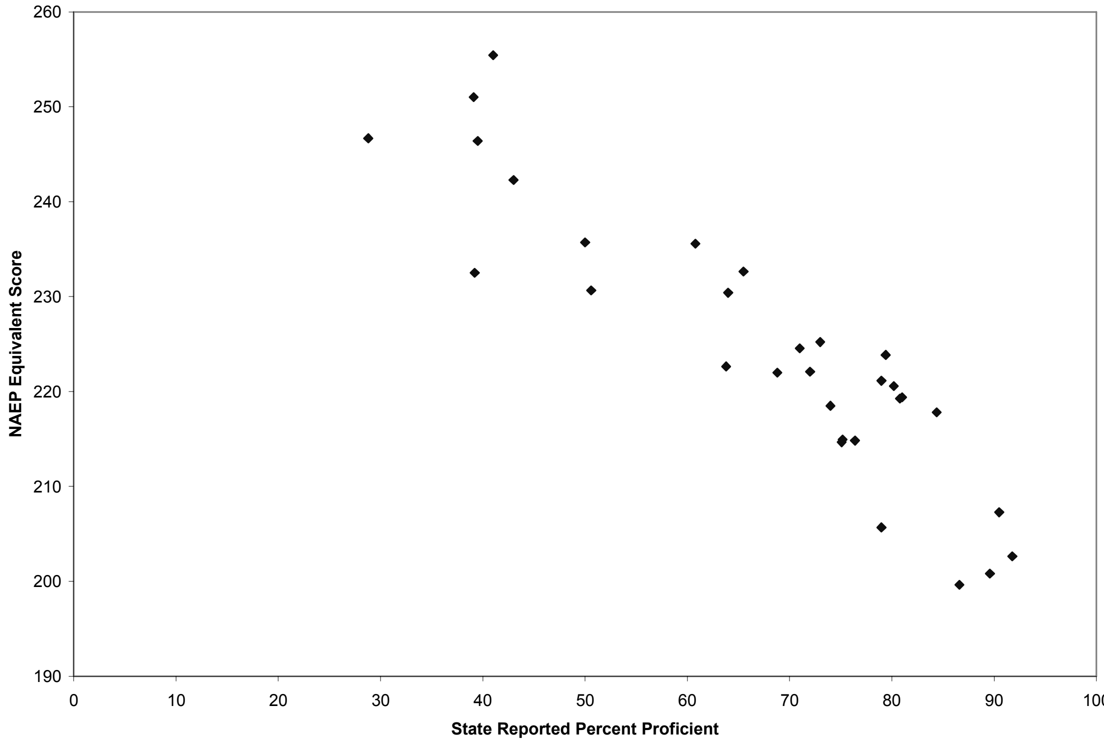
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-2. NAEP score equivalent vs. state-reported percent proficient for reading, grade 8: 2005



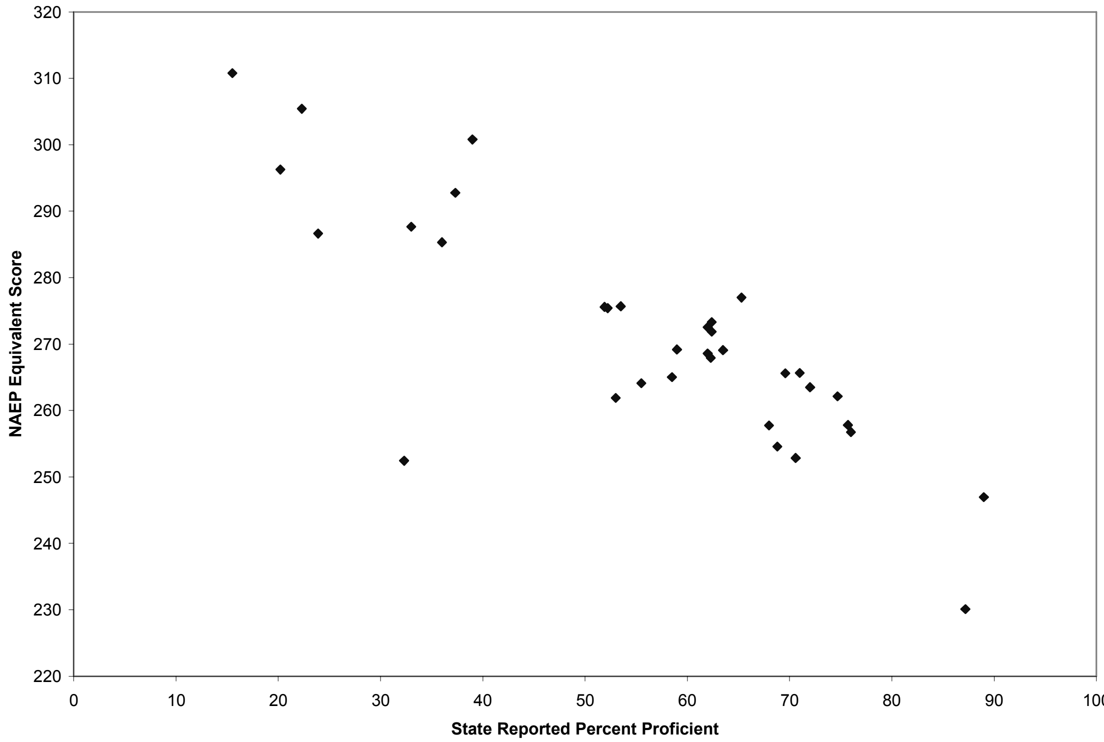
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-3. NAEP score equivalent vs. state-reported percent proficient for mathematics, grade 4: 2005



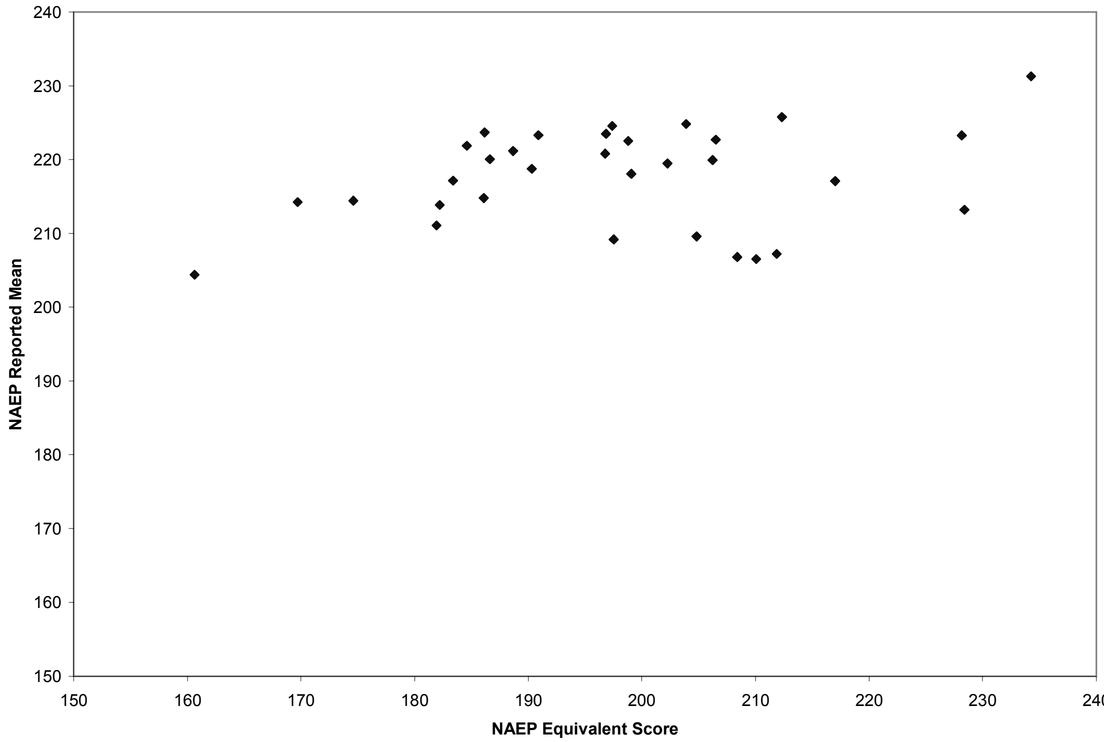
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-4. NAEP score equivalent vs. state-reported percent proficient for mathematics, grade 8: 2005



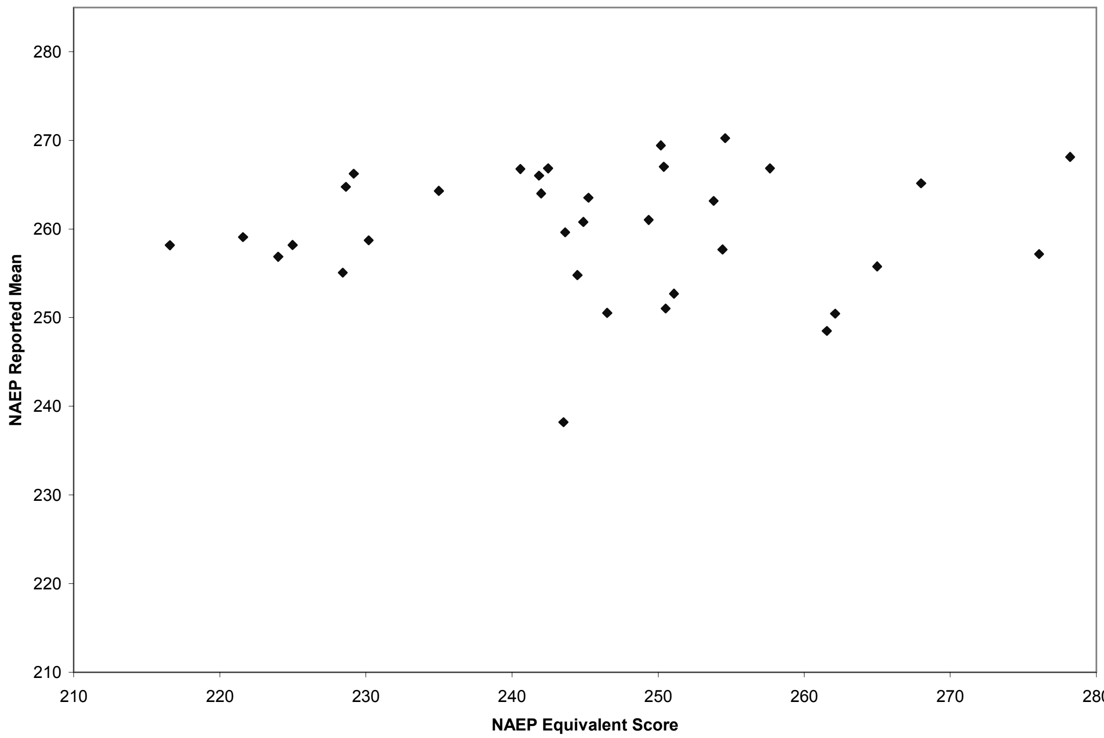
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-5. NAEP reported means vs. NAEP score equivalent for reading, grade 4: 2005



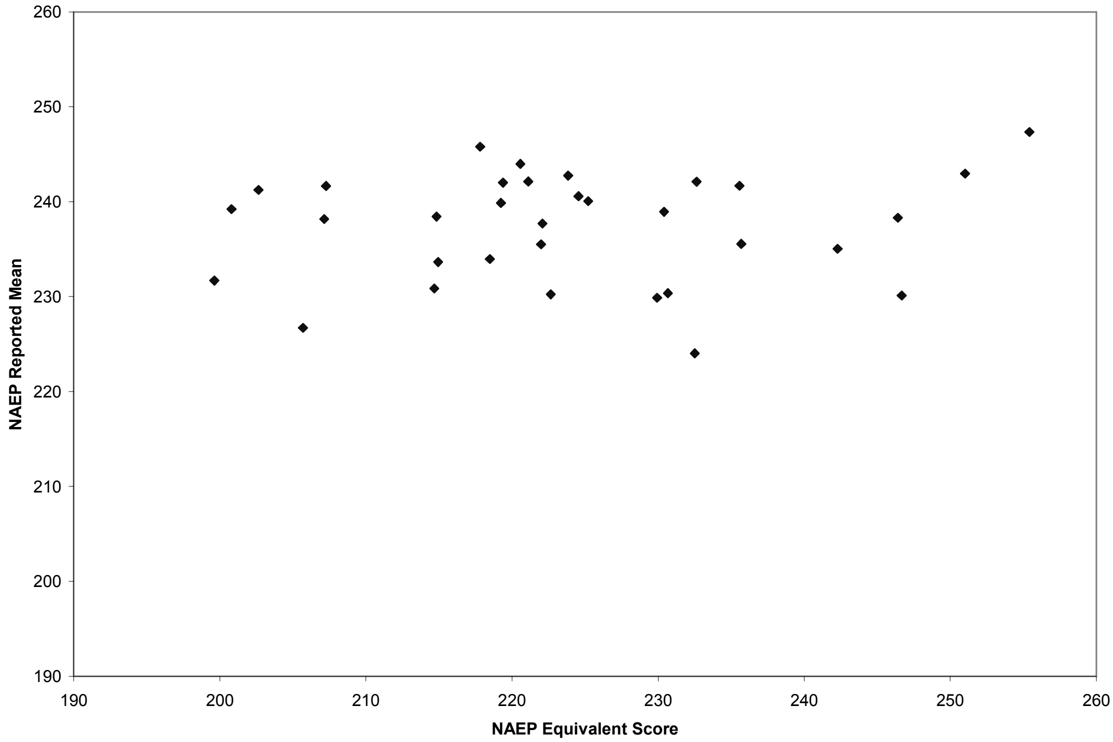
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-6. NAEP reported means vs. NAEP score equivalent for reading, grade 8: 2005



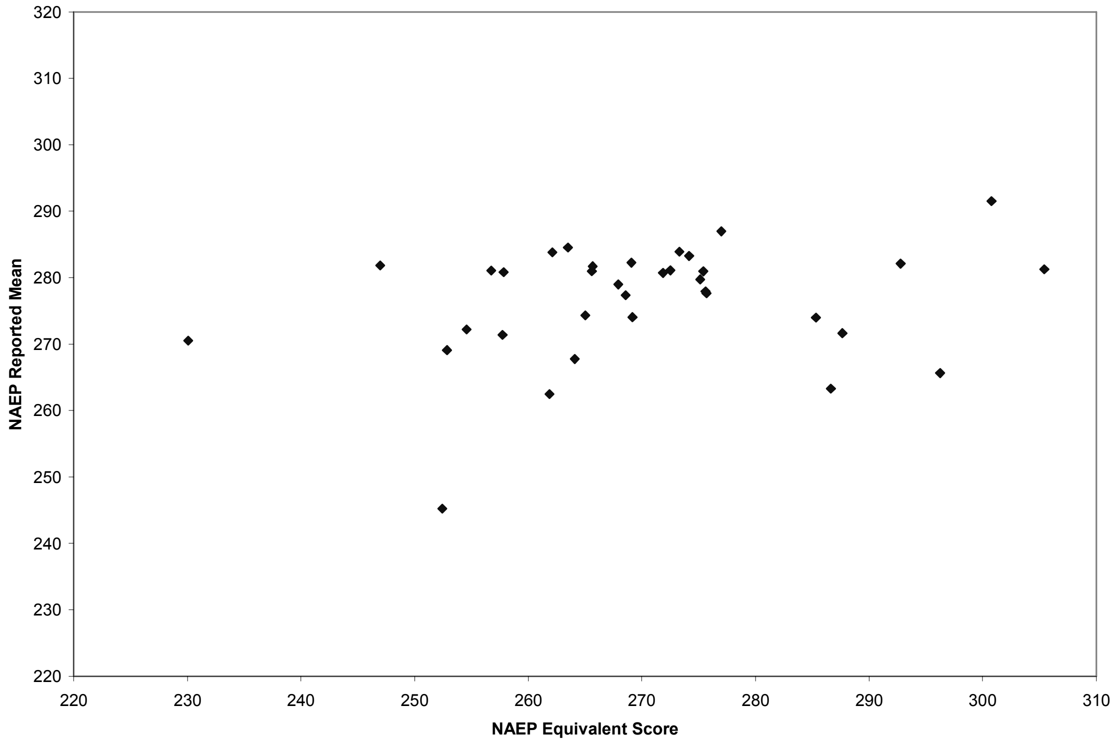
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-7. NAEP reported means vs. NAEP score equivalent for mathematics, grade 4: 2005



SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure C-8. NAEP reported means vs. NAEP score equivalent for mathematics, grade 8: 2005



SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

THIS PAGE INTENTIONALLY LEFT BLANK.

Appendix D

Results of 2003 Data Analysis

This appendix contains the results of carrying out the analyses described in the main text for state test data from the academic year 2002–03 and the NAEP 2003 assessment. Tables D-1 to D-4 parallel tables 1–4 in the main text, and figures D-1 to D-4 parallel figures 2–5 in the main text. Because of changes in relevant state policies that may have occurred between 2002–03 and 2004–05, the results of the two sets of analyses may not be comparable.

In the future it could prove useful, however, to identify those states where both policies and assessments have remained essentially unchanged over the academic years from 2001–02 to 2004–05. For these states, comparisons between the two sets of NAEP equivalents could be instructive. Such comparisons would provide evidence with respect to the stability of the linkages. (See, for example, a comprehensive document that provides a profile of NAEP and state assessment standards at http://nces.ed.gov/nationsreportcard/researchcenter/profile_standards.asp.)

Table D-1. Results of mapping state standards to the grade 4 NAEP reading scale: 2003

State	State name	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska	103	0.73	193	2.6
AR	Arkansas	115	0.62	206	1.7
CA	California	216	0.38	219	1.3
CO	Colorado	115	0.87	184	2.1
CT	Connecticut	108	0.68	215	1.8
DC	District of Columbia	103	0.47	192	0.8
FL	Florida	104	0.58	212	1.3
GA	Georgia	147	0.80	183	1.6
IA	Iowa	129	0.77	201	1.8
ID	Idaho	114	0.75	197	1.6
KY	Kentucky	121	0.62	211	1.6
LA	Louisiana	109	0.59	198	2.0
MA	Massachusetts	161	0.54	226	1.4
ME	Maine	145	0.50	226	1.1
MI	Michigan	133	0.74	197	1.8
MS	Mississippi	107	0.87	165	1.7
MT	Montana	141	0.77	199	2.0
NC	North Carolina	147	0.81	191	1.4
ND	North Dakota	176	0.75	201	1.0
NJ	New Jersey	109	0.78	198	1.2
NV	Nevada	106	0.49	211	1.5
NY	New York	145	0.64	211	1.6
OH	Ohio	163	0.69	207	2.2
SC	South Carolina	101	0.31	234	1.7
TX	Texas	194	0.85	177	1.7
WA	Washington	95	0.65	210	1.3
WI	Wisconsin	127	0.82	190	1.2
WY	Wyoming	145	0.44	230	1.0

NOTE: NAEP reading cut scores at grade 4 are 208 for *Basic* and 238 for *Proficient.* Median (SE of the NAEP equivalent) = 1.6. The following states' grade 4 reading test data were not used in the analysis or received special treatment: CT, LA, and OH—standard used for mapping is different from the standard specified as the AYP level; CA and LA—reading data not available for the state assessment, so English Language Arts (ELA) data used; AL, AZ, DE, HI, IL, IN, KS, MD, MN, MO, NH, NM, OK, OR, PA, RI, SD, TN, UT, VA, VT, and WV—state assessment data not available; NE—results deleted because state results are based on assessments developed by each local education agency.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Table D-2. Results of mapping state standards to the grade 8 NAEP reading scale: 2003

State	State name	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska	51	0.71	241	1.7
AR	Arkansas	99	0.44	267	1.8
AZ	Arizona	105	0.54	256	1.5
CA	California	180	0.32	271	1.2
CO	Colorado	104	0.88	229	1.9
CT	Connecticut	102	0.79	239	2.2
DC	District of Columbia	26	0.45	244	1.0
DE	Delaware	32	0.70	249	0.9
FL	Florida	96	0.47	263	1.6
GA	Georgia	113	0.81	230	2.1
HI	Hawaii	53	0.39	264	1.0
IA	Iowa	115	0.70	253	0.8
ID	Idaho	85	0.73	247	1.5
IL	Illinois	169	0.65	256	1.3
IN	Indiana	99	0.63	257	1.1
KS	Kansas	118	0.69	253	1.3
LA	Louisiana	94	0.52	253	1.5
MD	Maryland	95	0.62	252	1.7
ME	Maine	106	0.45	274	1.3
MS	Mississippi	102	0.55	250	1.3
MT	Montana	100	0.72	253	1.1
NC	North Carolina	129	0.86	226	1.6
ND	North Dakota	31	0.69	255	1.2
NJ	New Jersey	107	0.74	249	1.6
NY	New York	141	0.47	272	1.3
OK	Oklahoma	123	0.78	238	1.8
OR	Oregon	105	0.59	258	1.0
PA	Pennsylvania	100	0.63	256	1.5
SC	South Carolina	92	0.21	285	1.5
TX	Texas	142	0.88	221	1.7
WI	Wisconsin	103	0.84	232	2.9
WY	Wyoming	74	0.39	277	0.9

NOTE: NAEP reading cut scores at grade 8 are 243 for *Basic* and 281 for *Proficient*. Median (SE of the NAEP equivalent) = 1.5. The following states' grade 8 reading test data were not used in the analysis or received special treatment: VA—results deleted due to discrepancies between state assessment data and the state document; CT and LA—standard used for mapping is different from the standard specified as the AYP level; CA and LA—reading data not available for the state assessment, so English Language Arts (ELA) data used; AL, KY, MA, MI, MN, MO, NV, NH, NM, OH, RI, SD, TN, UT, VT, WA, and WV—state assessment data not available; NE—results deleted because state results are based on assessments developed by each local education agency.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Table D-3. Results of mapping state standards to the grade 4 NAEP mathematics scale: 2003

State	State name	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska	110	0.67	223	1.3
AR	Arkansas	115	0.60	223	0.9
CA	California	216	0.45	231	1.1
CT	Connecticut	108	0.80	217	1.1
DC	District of Columbia	103	0.54	201	0.7
FL	Florida	103	0.56	231	1.3
GA	Georgia	147	0.74	212	1.1
IA	Iowa	130	0.77	220	1.1
ID	Idaho	114	0.77	217	0.9
KS	Kansas	130	0.74	226	1.1
LA	Louisiana	109	0.58	221	1.1
MA	Massachusetts	161	0.38	251	1.1
ME	Maine	145	0.29	252	0.8
MI	Michigan	133	0.64	226	1.2
MO	Missouri	126	0.37	244	1.0
MS	Mississippi	107	0.74	205	1.3
MT	Montana	142	0.75	220	0.9
NC	North Carolina	151	0.92	203	1.0
ND	North Dakota	176	0.59	234	0.7
NJ	New Jersey	109	0.68	227	1.4
NV	Nevada	106	0.51	228	1.0
NY	New York	145	0.79	213	1.1
OH	Ohio	163	0.59	232	1.0
SC	South Carolina	101	0.33	248	0.9
TX	Texas	194	0.88	207	1.5
WA	Washington	96	0.54	236	1.2
WI	Wisconsin	127	0.70	223	1.1
WY	Wyoming	145	0.36	250	0.6

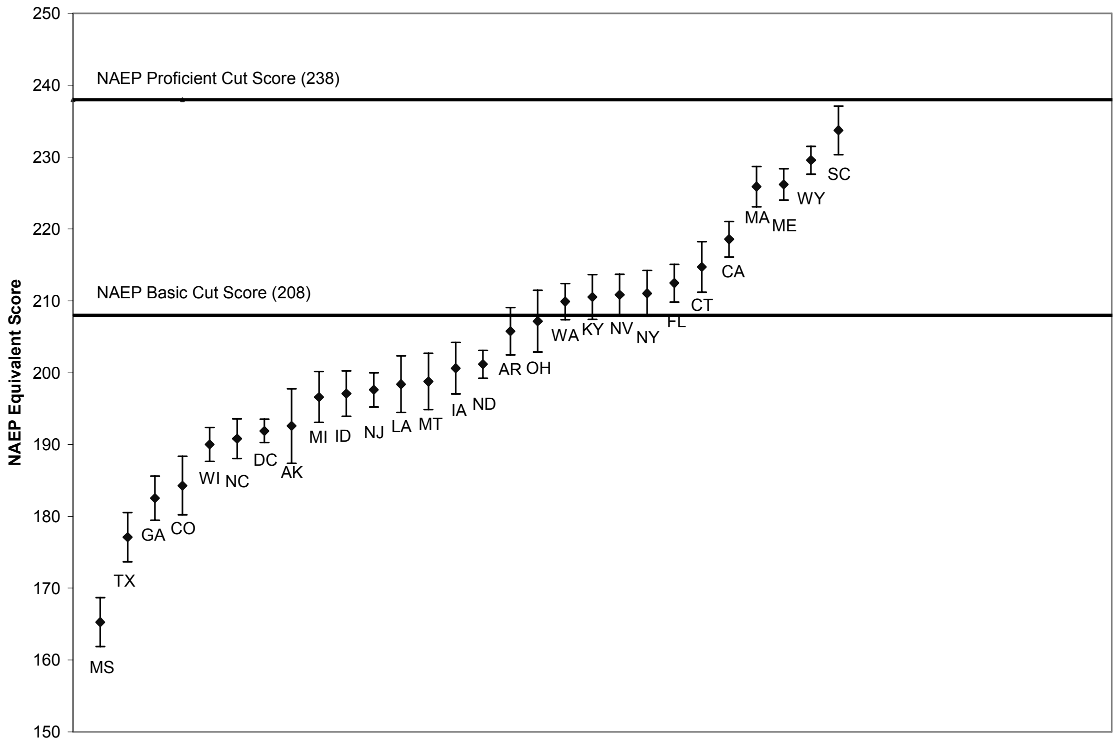
NOTE: NAEP mathematics cut scores at grade 4 are 214 for *Basic* and 249 for *Proficient.* Median (SE of the NAEP equivalent) = 1.1. The following states' grade 4 mathematics test data were not used in the analysis or received special treatment: CT, LA, and OH: standard used for mapping is different from the standard specified as the AYP level; AL, AZ, CO, DE, HI, IL, IN, KY, MD, MN, NE, NH, NM, OK, OR, PA, RI, SD, TN, UT, VA, VT, and WV—state assessment data not available.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Table D-4. Results of mapping state standards to the grade 8 NAEP mathematics scale: 2003

State	State name	Number of schools used in mapping	Estimate of proportion meeting the state proficiency standard	Estimated NAEP score equivalent to the state standard	Estimated standard error of the NAEP score equivalent
AK	Alaska	57	0.65	268	1.5
AR	Arkansas	99	0.22	296	1.5
AZ	Arizona	105	0.21	300	1.3
CO	Colorado	104	0.68	268	1.5
CT	Connecticut	102	0.77	258	1.6
DC	District of Columbia	27	0.43	250	0.9
DE	Delaware	32	0.48	278	1.0
FL	Florida	96	0.54	269	1.7
GA	Georgia	113	0.66	255	1.3
HI	Hawaii	54	0.17	299	1.9
IA	Iowa	115	0.72	266	1.3
ID	Idaho	86	0.52	280	0.9
IL	Illinois	169	0.54	276	1.4
IN	Indiana	99	0.66	269	1.5
KY	Kentucky	112	0.32	291	1.2
LA	Louisiana	94	0.52	265	1.4
MA	Massachusetts	128	0.38	299	0.8
MD	Maryland	95	0.43	286	1.2
ME	Maine	105	0.17	311	1.0
MI	Michigan	105	0.51	278	1.4
MO	Missouri	113	0.13	314	1.0
MS	Mississippi	102	0.46	261	1.0
MT	Montana	101	0.70	271	1.0
NC	North Carolina	129	0.82	247	2.1
ND	North Dakota	31	0.44	293	1.1
NJ	New Jersey	107	0.56	278	1.3
NY	New York	141	0.54	279	1.4
OK	Oklahoma	123	0.71	256	1.5
OR	Oregon	103	0.58	275	1.6
PA	Pennsylvania	100	0.52	279	1.4
SC	South Carolina	92	0.20	306	1.5
TX	Texas	142	0.71	260	1.2
WI	Wisconsin	103	0.76	261	1.6
WY	Wyoming	74	0.35	297	1.1

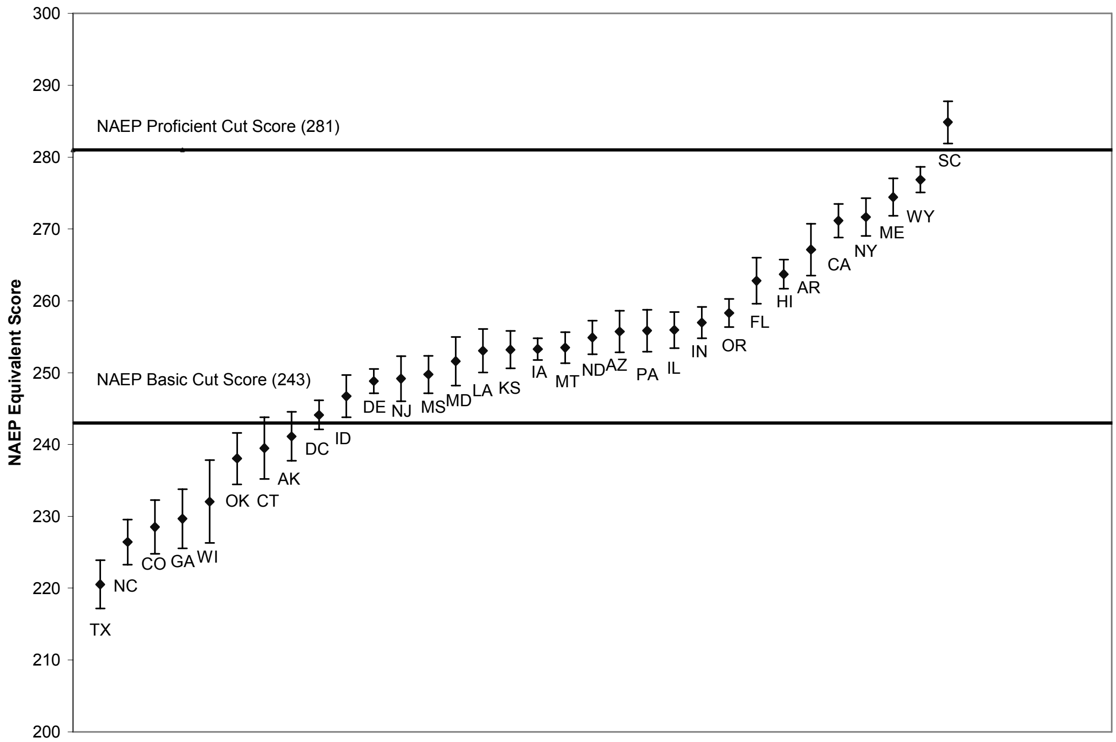
NOTE: NAEP mathematics cut scores at grade 8 are 262 for *Basic* and 299 for *Proficient.* Median (SE of the NAEP equivalent) = 1.3. The following states' grade 8 mathematics test data were not used in the analysis or received special treatment: VA—results deleted due to discrepancies between state assessment data and the state document; CT and LA—standard used for mapping is different from the standard specified as the AYP level; AL, CA, KS, MN, NE, NH, NM, NV, OH, RI, SD, TN, UT, VT, WA, and WV—state assessment data not available.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure D-1. NAEP score equivalents of states' proficiency standards for reading, grade 4: 2003



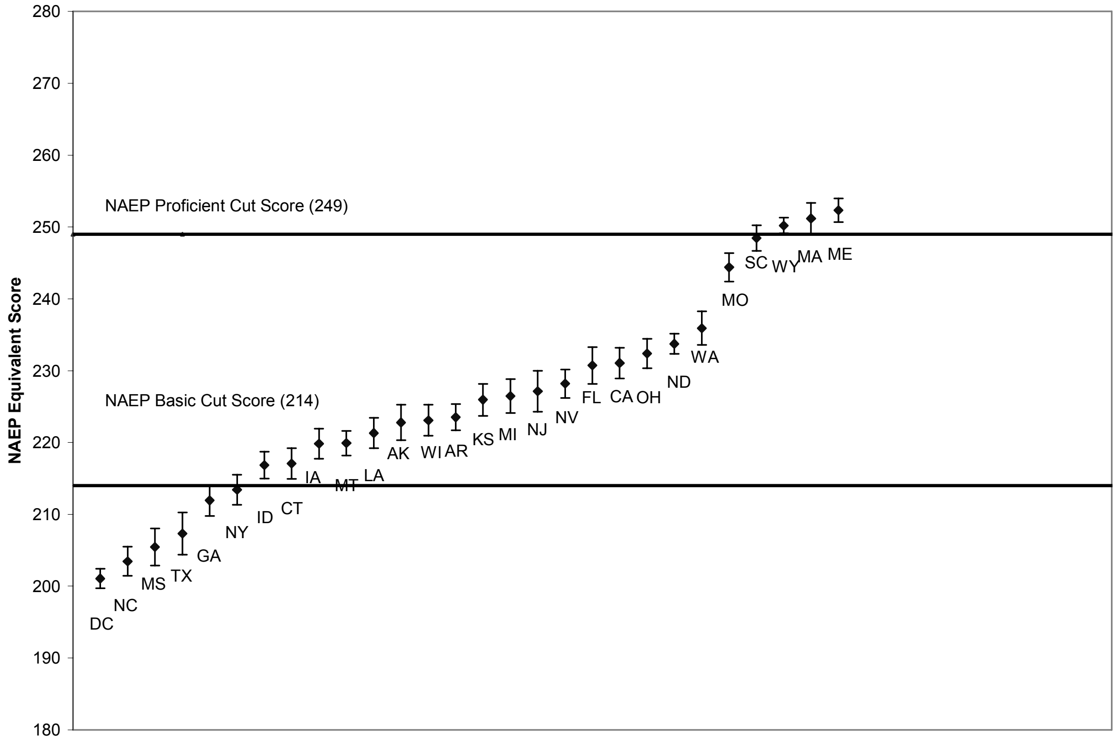
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure D-2. NAEP score equivalents of states' proficiency standards for reading, grade 8: 2003



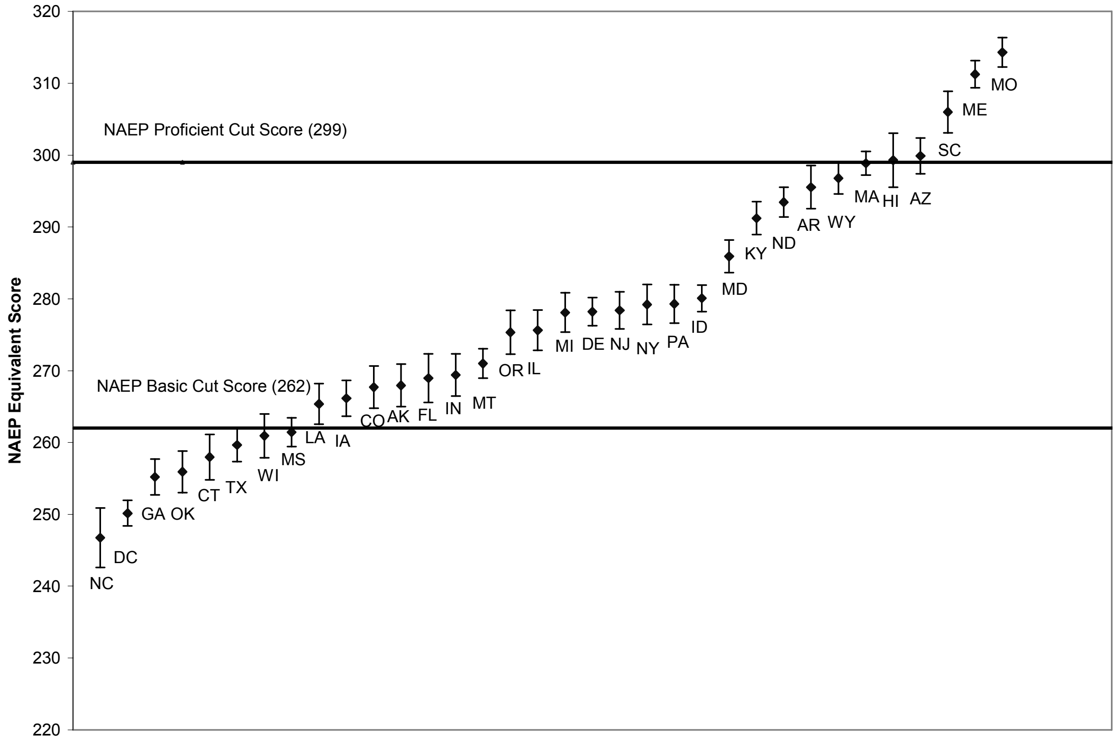
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Reading Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure D-3. NAEP score equivalents of states' proficiency standards for mathematics, grade 4: 2003



SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

Figure D-4. NAEP score equivalents of states' proficiency standards for mathematics, grade 8: 2003



SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2003 Mathematics Assessment, and National Longitudinal School-Level State Assessment Score Database (NLSLSASD).

THIS PAGE INTENTIONALLY LEFT BLANK.