

Mapping the NAEP Performance Levels to the Iowa Standard Score Scale

Joshua Tudor, Shalini Kapoor, Stephen Dunbar, and Catherine Welch

The University of Iowa

Paper presented at the annual meeting of the National Council on Measurement in Education,
Vancouver, BC, April, 2012.

Abstract

This paper considers the value of translating National Assessment of Educational Progress (NAEP) achievement levels to the standard score scale of an annual statewide assessment, in this paper the *Iowa Assessments*. NAEP achievement levels are often revered as being the gold standard to which states should aspire. However, results from state NAEP administrations are only reported at the aggregate level, in turn providing no indication of how individual students perform relative to NAEP. Similarly, interpretations often focus on the rank order of states, which is only partially indicative of the differences between states. An equipercentile linking method is used to express NAEP achievement levels in terms of expected score ranges on the Iowa standard score scale. The utility of the mapping process is demonstrated by expressing differences between top performing states and select subgroups of students in terms of raw score units on the *Iowa Assessments*. Of interest is how the translation of NAEP scores and score differences into another metric might provide interpretive meaning to state NAEP results. Considerations about the validity of possible interpretive statements are discussed.

Introduction of the No Child Left Behind (NCLB) Act of 2001 marked an increased focus on standard setting, and arriving at achievement level indicators and their respective cut scores (Perie, 2008; Huff & Plake, 2010). Today, a decade after the introduction of NCLB, the discussion surrounding accountability and performance standards continues with initiatives such as the Race to the Top Assessment Program (U.S. Department of Education, April 2010), and the Common Core State Standards Initiative (2010). Most recently the Obama administration published the *ESEA Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act* (U.S. Department of Education, Office of Planning, Evaluation and Policy Development, 2010). With its focus on college- and career-readiness standards, the blueprint offered a glimpse of what the next version of NCLB might look like. The release of the blueprint also saw states discussing more rigorous standards in an attempt to align with the current administration's focus. For many states this has involved consideration of National Assessment of Educational Progress (NAEP) performance standards and college readiness benchmarks in redefining their achievement-level cut scores and descriptors (Wyatt, Kobrin, Wiley, Camara, & Proestler, 2011).

The consideration of external indicators as a way to inform achievement level descriptors is in keeping with the movement towards rigorous standards. NAEP performance standards are often cited because of the perception that the NAEP standards are rigorous or aspirational (Hambleton, Sireci, & Smith, 2009; Linn, McLaughlin, & Thissen, 2009). This perception stems from discrepancies between the percentage of students meeting proficiency standards on the NAEP and on state assessments. For instance, Iowa's annual yearly progress report for the 2002-2003 academic year indicated that in grade 4 Reading 75.9% of students met the state's proficiency standard whereas 35% of Iowa students met the NAEP proficiency standard.

Discrepancies like the one in the aforementioned example became more salient in 2003 when the National Center for Educational Statistics (NCES) mapped states' Reading and Mathematics proficiency levels onto the NAEP scale. The mapping work done by NCES continued for subsequent NAEP administrations and extended the comparisons of states by incorporating each state's achievement levels onto the NAEP scale. The comparisons afforded by the mapping study involved each state's aggregated student performance on the NAEP Reading and Mathematics assessments along with each participating state's proficiency standards (Bandeira de Mello, 2011).

A similar, yet grander comparison was made using international grade 8 data from the NAEP, the Trends in International Mathematics and Science Study (TIMSS), and the Program for International Student Assessment (PISA) Mathematics assessments (Hambleton, Sireci, & Smith, 2009). After mapping the NAEP achievement levels onto the TIMSS and PISA scales it was determined that top performing countries on the TIMSS and PISA had higher percentages of students meeting the NAEP Proficiency standard than the United States. In light of this evidence and given the demand for external frames of reference, it seems reasonable for states to map NAEP achievement levels to their own score scales. The purpose of the research herein is to translate the NAEP achievement levels to the standard score scale of the *Iowa Assessments* using an equipercentile linking method. The goal of this linking is to express NAEP achievement levels in terms of expected score ranges on the Iowa standard score scale.

The NAEP sampling framework is such that a representative sample of each state's grade 4 and grade 8 student populations is selected to take the main NAEP assessment every two years. Performance on the assessments is then interpreted at the state level. The *Iowa Assessments* are administered annually to all Iowa students in grades 3 through 8 and 11. Although the NAEP and

Iowa Mathematics and Reading tests have overlapping test specifications and domain definitions, they perhaps differ the most in terms of the measurement conditions and their uses (see Tables 1 and 2 for item-skill classifications). According to Kolen and Brennan (2004), the shared specifications between the tests warrant referring to the equating as a linking or a concordance, so the goal of this study concerns the development of comparable scores and achievement levels.

Methods

The distribution of scores of grade 4 and grade 8 populations of Iowa students on the Iowa Reading and Mathematics tests were used in establishing a link to the distributions of the sample of Iowa students who were selected to take the state NAEP in 2003. The state of Iowa results for the 2003 NAEP Mathematics and Reading tests were obtained from the NAEP Data Explorer and used to determine the percentage of examinees classified in each of the three NAEP achievement levels - Basic, Proficient, and Advanced (NCES, n.d.-a). The NAEP Data Explorer also provides the NAEP scale scores associated with the 10th, 25th, 50th, 75th and 90th percentiles in the state. After identifying the cut score that defines the lower bound of each NAEP achievement level, the corresponding 2003 Iowa percentile ranks were identified. This resulted in 8 percentile points that served as the means for aligning the two scales (see Figure 1 for plots of the 8 percentile points for Iowa).

After identifying the Iowa standard score for each state percentile rank on NAEP, the standard score ranges associated with each NAEP range were identified. Figure 2 shows a representation of the Iowa NAEP distribution along with a plot of the unsmoothed equipercentile equivalents on the Iowa standard score scale. Cumulative distribution functions were then determined by fitting a curve to the 8 points using cubic spline smoothing. The smoothed curves depicted in Figure 3 show the cumulative distribution function of scores on the NAEP scale for

the state of Iowa along with a curve representing the cumulative distribution function of scores on the standard score scale of the *Iowa Assessments*. Note that differences between the locations of the cumulative distributions simply indicate metric differences between the Iowa and NAEP scales. Smoothing of the equipercentile equivalents provided an approximation of the scale scores and corresponding percentiles located between each of the observed data points used to conduct the linking.

Considerations for Interpretation

- NAEP and the *Iowa Assessments* rank order students similarly in Mathematics and Reading
- The NAEP and the *Iowa Assessments* have sufficiently comparable test specifications (cognitive levels and skills within domains) to support interpretations of linking results
- The NAEP state sample and the *Iowa Assessment* population represent equivalent groups needed to support linking

Results

In keeping with much of the research involving NAEP achievement levels, the relative location of NAEP achievement levels on another reporting metric is of interest. The equipercentile linking method identified for each grade and content area the particular standard scores on the *Iowa Assessments* that correspond with the NAEP achievement-level cut scores. Table 3 provides a summary of the NAEP achievement-level cut scores on the NAEP scale and the comparable Iowa standard scores obtained from the equipercentile linking procedure. One such result is that the cut score for proficient on the NAEP grade 4 Reading test of 238 has an equipercentile equivalent on the Reading test of the *Iowa Assessments* of a standard score of 222 (see Figure 4 for a graphical depiction of this linking relationship). Table 3 presents the Iowa

standard scores associated with each NAEP achievement-level cut score. The values in the table represent the lower bound of each NAEP achievement level range. As indicated in Table 3, it is estimated that an 8th grade student scoring at or above a standard score of 315 on the Mathematics test of the *Iowa Assessments* would score at a level associated with the lower bound of the Advanced range on NAEP. This interpretation, made possible by the fitted cumulative distribution function, can be made with respect to any Iowa standard score. Such an interpretation doesn't imply that if a student with a given Iowa standard score were to take NAEP s/he would score, say, at the Advanced level. It merely indicates that the student's relative performance on the Iowa scale is the same as a student's or group's relative performance at the NAEP-Advanced level.

In 2003 the state of Iowa began using a standard score with a national percentile rank of 41 when reporting out for adequate yearly progress purposes (Iowa Department of Education, 2003). Table 4 describes various achievement level cut scores (NAEP and state of Iowa) in terms of Iowa standard score equivalents. An indication of how different NAEP achievement levels are from state of Iowa achievement levels can be understood in terms of Iowa standard score units. For example, in grade 8 Reading, the cut score for "Proficient" on NAEP is 274 on the Iowa standard score scale. The corresponding cut score for the state of Iowa is 236, which is the standard score that corresponds to the 41st percentile on the *Iowa Assessments*. The difference of 38 points equals about 0.94 in Iowa standard deviation units. The metric conversion presented here represents a method of quantifying differences between NAEP achievement levels and those of a particular state (cf. Linn, McLaughlin, & Thissen, 2009). As another example, in grade 4 Reading, the cut score for "Proficient" on NAEP is 222 on the Iowa standard score scale, whereas the corresponding cut score for the state of Iowa is 190. The difference of 32 points

equals about 1.17 in Iowa standard deviation units (see Table 5 for differences between the NAEP cut score for proficient on the *Iowa Assessments* and the cut score for proficient on the state test expressed in terms of standard deviation units). Consistent with a finding first reported by Linn and Kiplinger (1995), in all of the grade/content pairs in Table 4, the state cut score for “Proficient” is closest to the NAEP cut score for “Basic.”

Iowa Raw Score Units as a Metric for Comparing Iowa to other States

One of the purposes of NAEP is to provide a common measure by which the student achievement of each state can be compared. One of the problems involved with using NAEP as the common measure to compare states is that attention is often directed towards the rank order of states when evaluating states relative to each other (Stoneberg, 2007). This is problematic from the standpoint that states’ scores are often misinterpreted as being free of error, and the difference between average scale scores can be within the margin of error. Similarly, using the difference between states’ average scale scores presents somewhat of an enigma as it is unclear how differences in scale score units translates to differences in raw score units. Expressing the differences between states in terms of the number of *Iowa Assessment* items provides a potentially useful metric for understanding the magnitude of the difference between states. Table 6 provides the average number of *Iowa Assessment* items for each NAEP scale score point.

Additional data from the 2003 and 2011 NAEP was obtained from the NAEP State Comparisons Tool (NCES, n.d.-b). Mathematics and Reading data for both grades 4 and 8 were generated separately. After generating each data set, Iowa was chosen as the reference state and the data were sorted by the scale score. Selecting Iowa as the reference state populates information regarding the number of jurisdictions that are significantly higher, not different, and lower than Iowa with respect to the sorted column (i.e., 2003 NAEP Scale Score). Table 7

describes for each grade and content area, differences between Iowa and the state with the score closest to Iowa that is identified by NCES as being significantly higher than Iowa. As can be seen in Table 7, the magnitude of the difference that constitutes statistical significance relative to Iowa's NAEP score corresponds to a difference of 1 raw score unit in grade 4 Reading and Mathematics, while in grade 8 Reading and Mathematics the smallest score difference on NAEP that constitutes statistical significance relative to Iowa corresponds to 1 and 2 Iowa items, respectively. The three top performing states on NAEP grade 4 Reading were Connecticut, New Hampshire, and Massachusetts. The average Reading NAEP scale score for 4th grade public school students in Connecticut, New Hampshire, and Massachusetts was 228 while the average scale score for 4th grade public school students in Iowa was 223. When expressed in terms of raw score units on the *Iowa Assessment*, this rounded difference of 5 scale score units corresponds to a rounded difference of 1 raw score unit. It is also worth noting that the gain made by Massachusetts in grade 4 Reading between 2003 and 2011 is equivalent to 2 items correct on the *Iowa Assessment*, while the corresponding change made by Iowa is equivalent to 1 raw score unit less in 2011 than in 2003. As alluded to earlier, focusing interpretations on the rank alone doesn't help policymakers understand the magnitudes of score differences. In 2003 Massachusetts had a rank order of 3 while Iowa had a rank order of 11. In this case the rank order is a function of the relative standing of a state in terms of its unrounded scale score and the associated standard error. Interpretations can be better understood in terms of magnitude when an easily understood metric can be associated with results, and focusing on cross-state significant differences in this research helps quantify the magnitude of the differences in terms of raw scores on an assessment familiar to state educators and policymakers.

Iowa Raw Score Units as a Metric for Comparing Student Subgroups

One of the limiting aspects of NAEP in terms of reporting is that results are only provided at an aggregate level and individuals who are selected to participate in the assessment do not receive an indication of how they performed. State results are disaggregated by particular subgroups of interest (i.e., race/ethnicity, free and reduced-price lunch, etc.), and similar to the aforementioned problem in comparing states, it is difficult to attach a meaningful interpretation to differences between scale scores when evaluating the differences in achievement of student subgroups.

For instance, the 2003 achievement gap in NAEP grade 8 Reading in Iowa between Hispanic and White students is equivalent to 10 items on the *Iowa Assessment* while the 2011 achievement gap corresponds to 7 items on the state test. These raw score differences correspond to NAEP scale score differences of 25.27 and 16.11 for 2003 and 2011, respectively. The 2003 achievement gap in NAEP grade 4 Reading in Iowa between Hispanic and White students is equivalent to 6 items correct on the *Iowa Assessment* while the 2011 achievement gap corresponds to 6 items on the state test. These raw score differences correspond to NAEP scale score differences of 20.86 and 23.95 for 2003 and 2011, respectively. According to this, the 2011 achievement gap in grade 8 Reading is less than it was in 2003 by 3 items, while the 2003 achievement gap in grade 4 Reading is equivalent to the same number of rounded raw score units in 2011.

In terms of the 2003 achievement gap between Hispanic and White students in grade 8 Mathematics in Iowa, a NAEP scale score difference of 31.62 is equivalent to a rounded difference of 17 items on the *Iowa Assessment* while the 2011 achievement gap of 19.52 NAEP scale score points corresponds to a rounded difference of 11 items on the state test. Thus, the Hispanic-White achievement gap on NAEP grade 8 Mathematics decreased in magnitude by 6

items in the Iowa raw-score metric. The 2003 achievement gap between Hispanic and White students in grade 4 Mathematics in Iowa corresponds to a NAEP scale score difference of 18.47, which is equivalent to a rounded difference of 10 items on the *Iowa Assessment*. The 2011 achievement gap of 16.69 NAEP scale score points corresponds to a rounded difference of 9 items on the state test. According to these results the achievement gap in grade 8 Mathematics between White and Hispanics in Iowa has decreased from 2003 to 2011 while it has remained relatively stable for grade 4 Mathematics. These results indicate the reduction of the Hispanic-White achievement gap on NAEP grade 8 Mathematics is bigger in magnitude than the corresponding reduction in grade 4. Similar comparisons can be made with respect to the achievement of other groups of interest.

Limitations of the Study

As with any research that seeks to link scores from one test to another, the extent to which the assumptions are satisfied influences the extent to which results from the linking are useful. Although the *Iowa Assessments* and NAEP serve different purposes and have different test specifications, the classification of items in Tables 1 and 2 seem sufficiently comparable to warrant the linking. Furthermore, it seems reasonable that the randomly equivalent groups assumption required to make assertions regarding the relative difficulty of one test form compared to another is satisfied. In the case of the Iowa NAEP, sample-based results representative of the population of Iowa students exists, while in the case of the *Iowa Assessments* a census of Iowa students that is representative of the achievement of Iowa students exists. On the subject of linking non-parallel tests Lindquist (1964) stated, “We can, in a certain sense, establish ‘comparable scales’ for such tests, but we cannot equate the scores to one another” (p. 9). That goes to say that while the statistical mechanisms used to conduct the linking

are the same as in the context of equating parallel test forms, referring to the process and product as mapping and comparable scores is intended to be more than a difference in verbiage from the equating context. Rather, referring to the process and product as mapping and comparable scores is indicative of the type of interpretations supported by linking non-parallel tests, which is consistent with the interpretations made in this paper.

Summary and Concluding Discussion

Many measurement professionals have suggested that mapping the NAEP achievement levels onto the scale of another test provides a frame of reference that can serve to inform policymakers, educators, and the public in evaluating student performance and performance standards (Beaton, Linn, & Bohrnstedt, 2012; Hambleton, Sireci, & Smith, 2009). In September of 2011, the Obama administration responded to the lack of progress from congress in reauthorizing ESEA and announced a plan that allows states flexibility from certain NCLB mandates should they demonstrate a transition to college- and career-ready standards. Some of the activities the U.S. Department of Education (September 2011) cited as acceptable evidence for demonstrating the transition includes conducting NAEP mapping studies and using an advanced achievement level on State assessments instead of the proficient achievement level as the standard for students to meet. To that end, finding standard scores on the *Iowa Assessments* that are comparable to NAEP scale scores provides an alternative way to understand differences between any two NAEP scale score points, and mapping the NAEP achievement levels onto the Iowa standard score scale provides an external indicator of proficient that is comparable to NAEP “Proficient” in terms of difficulty.

References

- Bandeira de Mello, V. (2011), Mapping state proficiency standards onto the NAEP scales: Variation and change in state standards for Reading and Mathematics, 2005–2009 (NCES 2011-458). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC: Government Printing Office.
- Beaton, A. E., Linn, R. L., & Bohrnstedt, G. W. (2012). *Alternative approaches to setting performance standards for the National Assessment of Educational Progress*. Washington, DC: American Institutes for Research.
- Common Core State Standards Initiative. (2010). *Common core state standards*. Retrieved from <http://corestandards.org>
- Hambleton, R. K., Sireci, S. G., & Smith, Z. R. (2009). How do other countries measure up to the Mathematics performance levels on the national assessment of educational progress? *Applied Measurement in Education*, 22(4), 376-393.
- Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2003). *The Iowa Tests interpretive guide for school administrators*. Chicago, IL: Riverside.
- Huff, K., & Plake, B. S. (2010). Innovations in setting performance standards for K–12 test-based accountability. *Measurement: Interdisciplinary Research & Perspective*, 8, 130-144.
- Iowa Department of Education. (2003). *The State Report Card for No Child Left Behind*. Des Moines, IA: Department of Education. Retrieved from http://educateiowa.gov/index.php?option=com_content&view=article&id=346&Itemid=4439#StateReportCards
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and*

practices. New York: Springer.

Lindquist, E. F. (1964). Equating scores on non-parallel tests. *Journal of Educational Measurement, 1*, 5-9.

Linn, R. L., & Kiplinger, V. L. (1995). Linking statewide tests to the National Assessment of Educational Progress: Stability of results. *Applied Measurement in Education, 8* (2), 135-155.

Linn, R. L., McLaughlin, D. H., & Thissen, D. (2009). *Utility and validity of NAEP linking efforts*. Washington, DC: American Institutes for Research.

National Center for Education Statistics. (2003). *2003 Mathematics assessment*. Retrieved from http://nces.ed.gov/nationsreportcard/tdw/instruments/2002_2003/cog_dev_math_number2003.asp.

National Center for Education Statistics. (2003). *2003 Reading assessment*. Retrieved from http://nces.ed.gov/nationsreportcard/tdw/instruments/2002_2003/cog_dev_read_number2003.asp.

National Center for Education Statistics. (n.d.-a). *NAEP data explorer* [Data file]. Washington, DC: U. S. Department of Education. Retrieved from <http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx>.

National Center for Education Statistics. (n.d.-b). *NAEP state comparisons tool* [Data file]. Washington, DC: U. S. Department of Education. Retrieved from <http://nces.ed.gov/nationsreportcard/statecomparisons/>.

No Child Left Behind Act of 2001. Public Law 107-110.

Perie, M. (2008). A guide to understanding and developing performance-level descriptors. *Educational Measurement: Issues and Practice, 27*(4), 15–29.

Stoneberg, B. D. (2007). Using NAEP to confirm state test results in the No Child Left Behind Act. *Practical Assessment Research & Evaluation*, 12(5). Retrieved from

<http://pareonline.net/pdf/v12n5.pdf>

U.S. Department of Education. (2010, April). *Race to the top assessment program executive summary*. Washington, DC: U.S. Department of Education. Retrieved from

<http://www2.ed.gov/programs/racetothetop-assessment/executive-summary-042010.pdf>.

U.S. Department of Education. (2011, September). *ESEA flexibility review guidance*.

Washington, DC: U.S. Department of Education. Retrieved from

<http://www.ed.gov/esea/flexibility>.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2010).

ESEA blueprint for reform. Washington, DC: U.S. Department of Education. Retrieved from <http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf>.

Wyatt, J., Kobrin, J., Wiley, A., Camara, W. J., & Proestler, N. (2011). SAT benchmarks:

Development of a college readiness benchmark and its relationship to secondary and postsecondary school performance. (College Board Research Report 2011-5). New York: The College Board.

Appendix

Table 1

Percentage of NAEP and Iowa Assessment Mathematics items classified by skill

Mathematics	Grade 4		Grade 8	
	ITBS	NAEP	ITBS	NAEP
Number Sense, Properties, and Operations	18	42	19	26
Measurement	05	18	05	15
Geometry and spatial sense	08	15	07	18
Data Analysis, Statistics, and Probability	05	11	05	15
Algebra and functions	10	15	10	25
Estimation	13	-	15	-
Problem Solving	23	-	23	-
Data Interpretation	17	-	16	-

Note. Item percentages obtained from Hoover, Dunbar, and Frisbie (2003), and NCES (2003). Totals may not sum to 100 due to rounding.

Table 2

Percentage of NAEP and Iowa Assessment Reading items classified by skill

Reading	Grade 4		Grade 8	
	ITBS	NAEP	ITBS	NAEP
Reading for literary experience (Factual Understanding)	41	50	35	33
Reading to gain information (Inference and Interpretation)	34	50	35	38
Reading to perform a task (Analysis and Generalization)	24	-	31	28

Note. Item percentages obtained from Hoover, Dunbar, and Frisbie (2003), and NCES (2003). Totals may not sum to 100 due to rounding.

Table 3

NAEP achievement levels and comparable cut scores on the Iowa standard score scale

		NAEP PR	NAEP SS	Iowa SS
Grade 4 Reading	Basic	30	208	197
	Proficient	65	238	222
	Advanced	93	268	255
Grade 8 Reading	Basic	21	243	226
	Proficient	64	281	274
	Advanced	97	323	323
Grade 4 Math	Basic	17	214	186
	Proficient	64	249	216
	Advanced	97	282	251
Grade 8 Math	Basic	24	262	234
	Proficient	67	299	281
	Advanced	95	333	315

Table 4

Iowa and NAEP achievement levels on the Iowa standard score scale

	Achievement Levels	Iowa NSS
Grade 4 Reading	Not Proficient _{Iowa}	130
	Proficient _{Iowa}	190
	Basic	197
	Proficient	222
	Advanced _{Iowa}	236
	Advanced	255
Grade 8 Reading	Not Proficient _{Iowa}	145
	Basic	226
	Proficient _{Iowa}	236
	Proficient	274
	Advanced _{Iowa}	299
	Advanced	323
Grade 4 Math	Not Proficient _{Iowa}	130
	Basic	186
	Proficient _{Iowa}	190
	Proficient	216
	Advanced _{Iowa}	228
	Advanced	251
Grade 8 Math	Not Proficient _{Iowa}	145
	Basic	234
	Proficient _{Iowa}	237
	Proficient	281
	Advanced _{Iowa}	295
	Advanced	315

Note. Each number represents the lower bound of the achievement level. The achievement levels that contain the _{Iowa} subscript correspond to the achievement levels used for the state test.

Table 5

Cut scores for proficient expressed as Iowa standard scores and Iowa standard deviation units

	NAEP Proficient	Iowa Proficient	Difference	Effect Size
Grade 4 Reading	222	190	32	1.17
Grade 8 Reading	274	236	38	0.94
Grade 4 Mathematics	216	190	26	1.14
Grade 8 Mathematics	281	237	44	1.25

Table 6

Average number of Iowa items for each NAEP scale score point

Content Area	Reading	4	8
		Mathematics	.23
		.42	.55

Table 7

Select 2003 NAEP comparisons expressed as raw score differences on the Iowa Assessments

	Rank	Higher	Equal	Lower	State	NAEP SS	Iowa SS	Iowa RS
Grade 4 Reading	3	0	4	47	MA	227.60	210.70	28.90
	11	3	21	27	IA	223.27	207.00	27.67
					<i>Difference</i>	4.33	3.70	1.23
Grade 8 Reading	5	1	12	38	SD	269.97	259.83	34.61
	13	5	19	27	IA	267.50	256.48	33.49
					<i>Difference</i>	3.35	3.35	1.12
Grade 4 Math	7	0	9	42	WY	241.09	208.88	42.54
	11	7	15	29	IA	238.48	206.71	41.22
					<i>Difference</i>	2.61	2.17	1.32
Grade 8 Math	3	1	9	41	MA	286.52	263.88	50.62
	10	3	18	30	IA	283.95	260.23	48.97
					<i>Difference</i>	2.57	3.65	1.65

Note. Higher, Equal, and Lower refer to statistical significance as defined by NAEP, based on the unrounded average NAEP scale scores relative to those of the other jurisdictions.

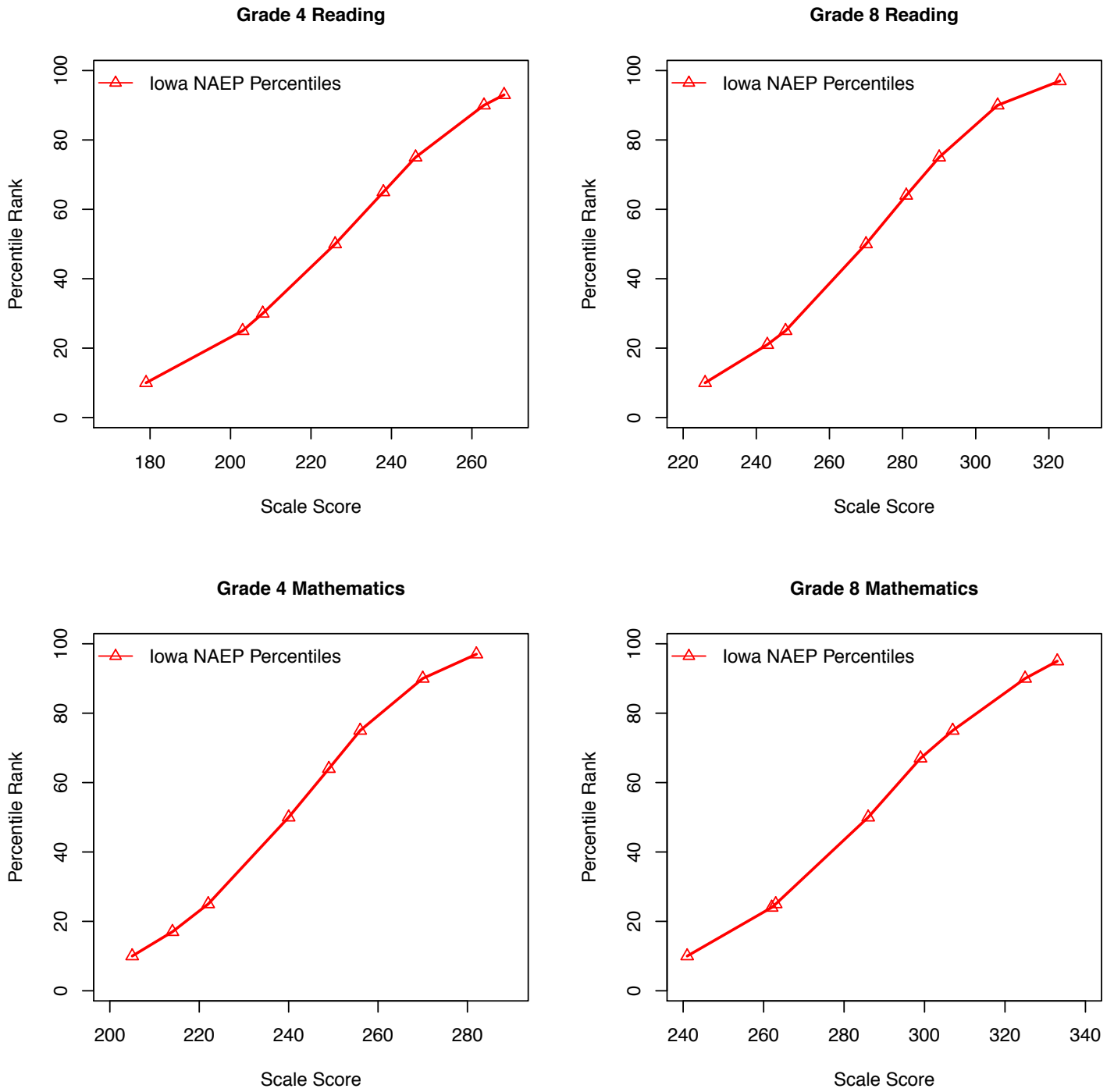


Figure 1. Iowa NAEP percentiles on the NAEP scale.

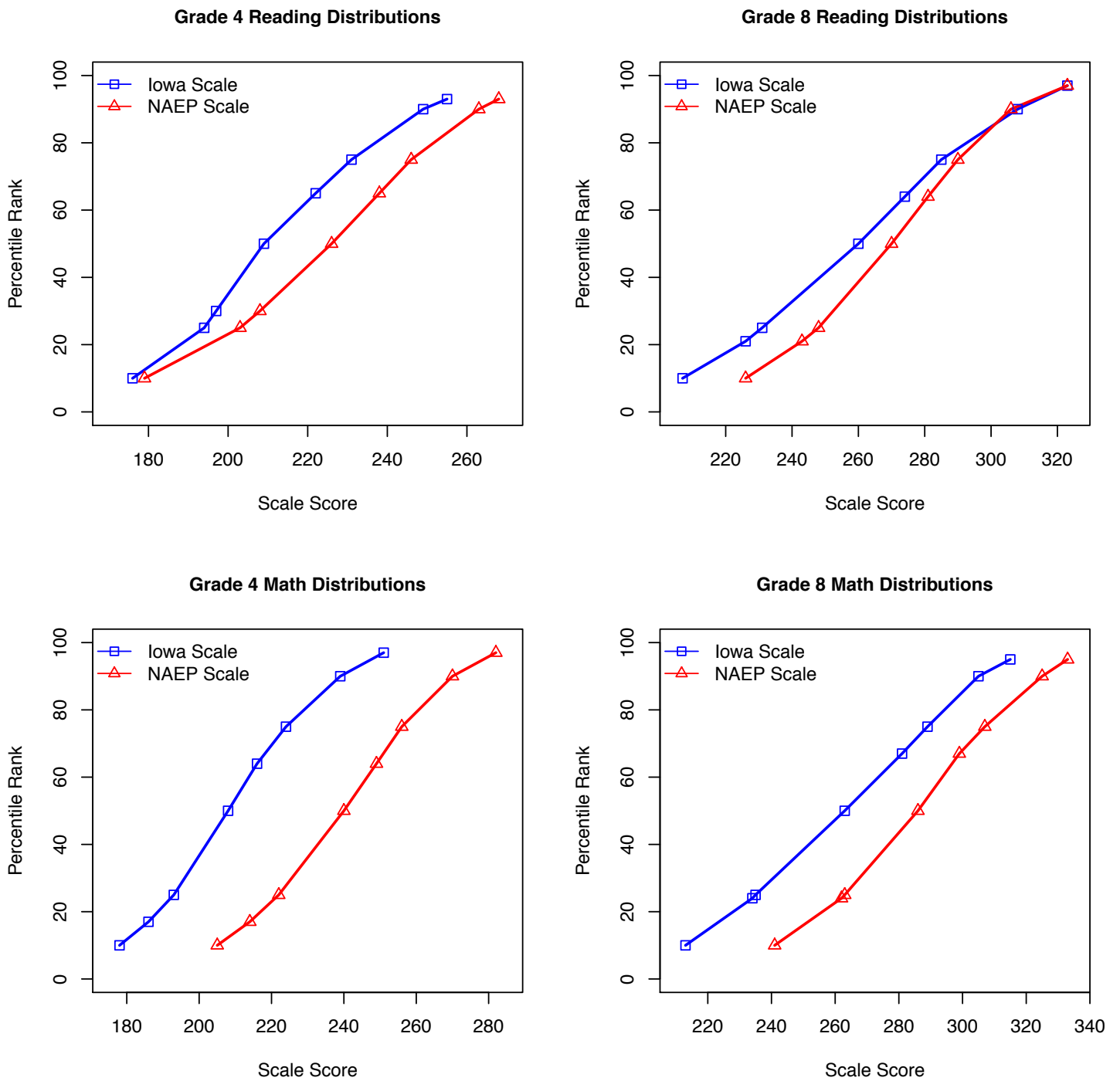


Figure 2. Iowa NAEP percentiles on the NAEP scale along with the corresponding equivalents on the Iowa standard score scale.

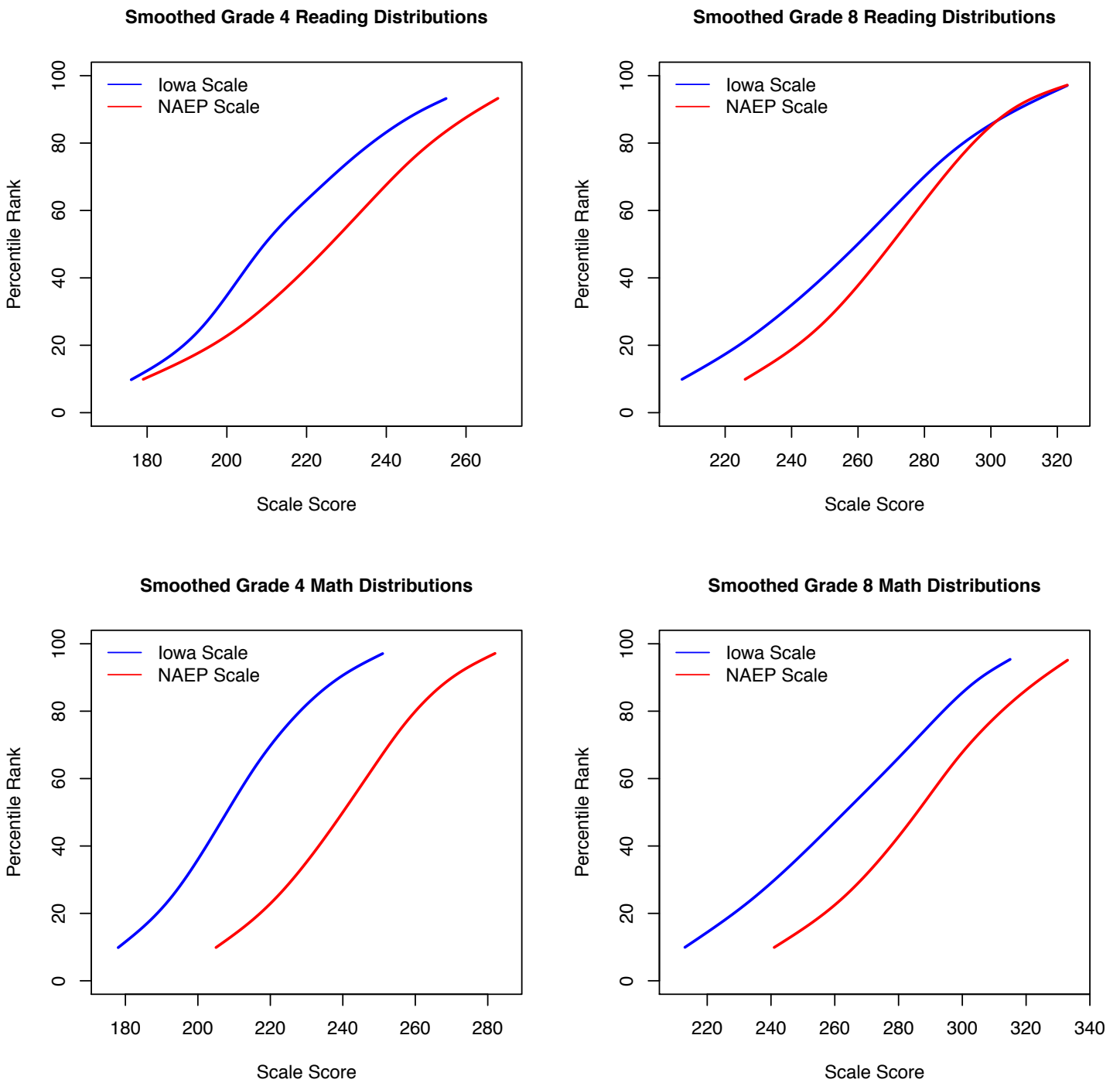


Figure 3. Smoothed NAEP and Iowa distributions. Both the Iowa NAEP distribution on the NAEP scale and the corresponding linked distribution on the Iowa scale were smoothed using cubic splines.

Smoothed Grade 4 Reading Distributions

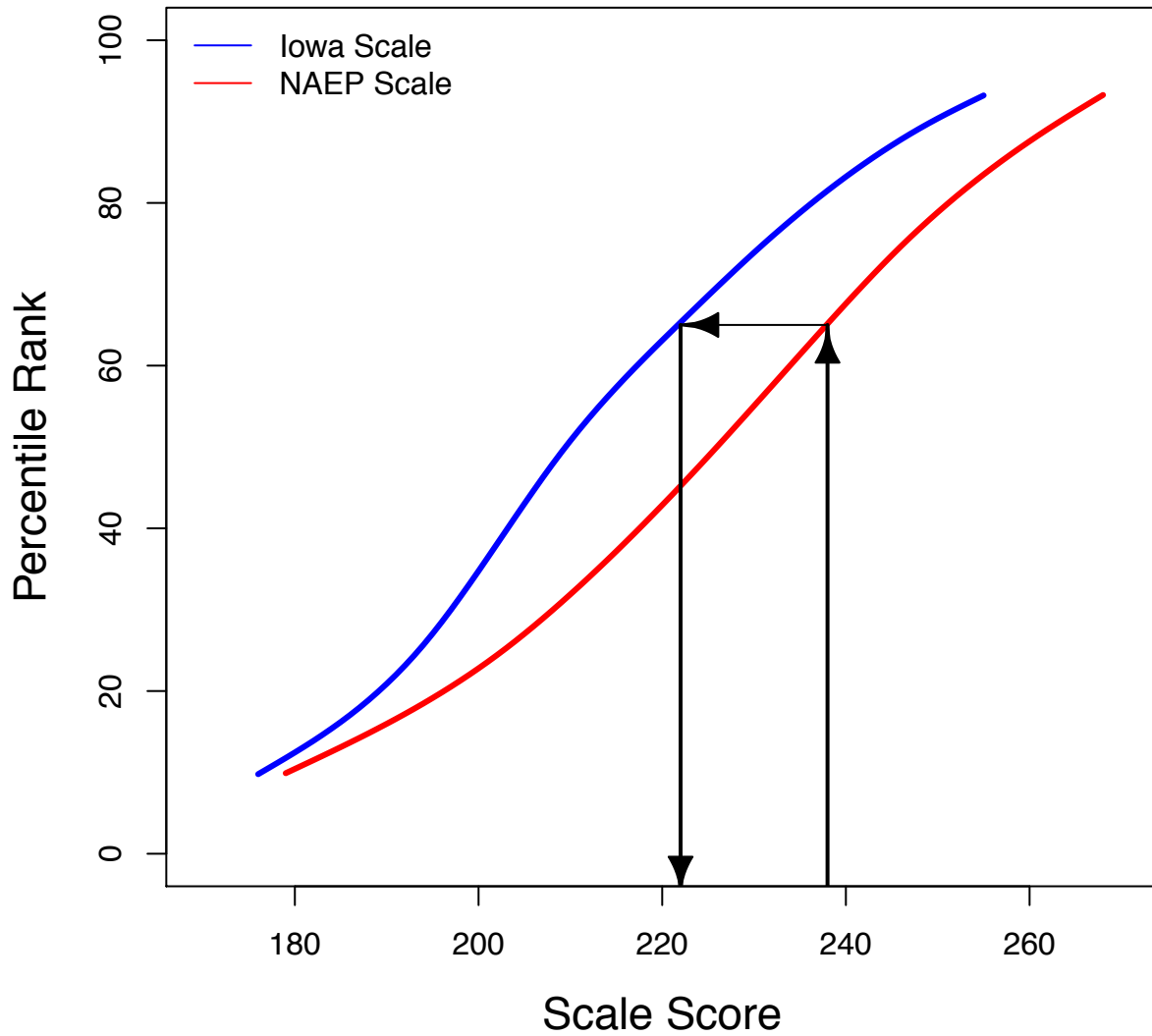


Figure 4. Smoothed grade 4 Reading distributions. The blue line reflects the comparable NAEP distribution on the Iowa scale while the red line reflects the Iowa NAEP distribution on the NAEP scale.