American Educational Research Journal

http://aerj.aera.net

Is There an Expectations Gap? Educational Federalism and the Demographic Distribution of Proficiency Cut Scores
Douglas S. Reed
Am Educ Res J 2009 46: 718
DOI: 10.3102/0002831209340254

The online version of this article can be found at:
http://aer.sagepub.com/content/46/3/718

Published on behalf of
American Educational Research Association

and
SAGE

http://www.sagepublications.com

Additional services and information for American Educational Research Journal can be found at:

Email Alerts: http://aerj.aera.net/alerts

Subscriptions: http://aerj.aera.net/subscriptions

Reprints: http://www.aera.net/reprints

Permissions: http://www.aera.net/permissions
Is There an Expectations Gap? Educational Federalism and the Demographic Distribution of Proficiency Cut Scores

Douglas S. Reed
Georgetown University

Recent analyses of state-level proficiency standards under No Child Left Behind indicate that states’ benchmarks for determining whether a child is proficient in reading and math are widely divergent. This article explores whether the capacity of states to employ different proficiency standards imposes different performance standards across demographic groups. Using a newly devised metric that allows for interstate comparison of state-level proficiency cut scores, along with aggregated district-level demographic data, this article provides a descriptive analysis of the distribution of proficiency standards across demographic groups to determine whether some groups systematically face higher or lower proficiency cut scores. The findings indicate that while an “expectations gap” does exist, it does not operate consistently across categories of race, ethnicity, and poverty status. The article concludes by reviewing the trade-offs required under three possible policy responses to these findings.

KEYWORDS: educational policy, accountability, minorities, politics, assessment

Advocates of standards and robust accountability systems frequently insist that any program to close the achievement gap should ensure that all students are held to uniformly high performance standards. Indeed, the very framework of No Child Left Behind (NCLB) requires that all students within a state—including students who have limited English proficiency and students with special needs—take the same state-level, standards-based test in order to demonstrate a school’s adequate yearly progress toward the goal of ensuring that all students be proficient in math and reading by the year 2014. While there have been numerous calls to broaden the set of indicators
Is There an Expectations Gap?

schools might use to demonstrate a student’s proficiency (Kornhaber, 2008; Linn, 2008), NCLB as currently written requires students within a state to demonstrate a similar level of performance on the same test in order to achieve the designation of “proficient.”

Despite NCLB’s apparent commitment to uniformity among students, many analysts have pointed out that the system of educational federalism upon which NCLB rests allows each state to establish the performance standards that students must meet in order to be deemed “proficient” within a particular subject area, under the accountability provisions of No Child Left Behind. As a result, states exhibit large variation in the performance standards needed to earn the label proficient and in the percentages of students who in fact do earn the label (Cronin, Dahlin, Adkins, & Kingsbury, 2007). As one author noted, the definition of proficient achievement varies so “wildly from state to state” that the idea of proficiency “lacks any semblance of a common meaning across states” (Linn, 2008, p. 29). This variation in performance standards forces us to revisit the claim that NCLB promotes greater equality in educational opportunity, and it raises important questions about the distribution of students across varying definitions of proficiency. If particular demographic groups of students are clustered in states with low performance standards, state-level variation in performance standards may be producing a different kind of educational opportunity gap—a gap not of academic achievement but of academic expectation.

This article engages, then, in a descriptive analysis of the distribution of proficiency standards across major demographic groups in order to determine whether some groups systematically face higher or lower cut scores needed to be deemed proficient. Using a newly devised metric that allows for inter-state comparison of state-level proficiency cut scores, I seek to answer whether some groups, when viewed from a national perspective, are clustered in states with either higher or lower performance standards for proficiency. The findings raise important policy implications about educational federalism and the extent to which state-level accountability mechanisms are contributing to a systems-level gap in academic expectations for U.S. students.

To perform this demographic analysis, the article draws on a recently developed comparative ranking of state-level cut scores. This index enables us to make more meaningful comparisons among states than has been possible before, and by combining these comparable state-level measures of performance with district-level demographic data, we can explore with nuance how these proficiency standards are distributed across U.S. students. In sum, the aspirations for this article are modest: to map the range of proficiency cut scores against the demographic distribution of students in school districts across the country. Doing so enables us to see to what extent educational federalism in the United States systematically presents students of different racial, ethnic, and economic groups with different performance standards in order to be deemed “proficient” within math and reading.

The findings of this investigation indicate that while lower performance standards are expected of some demographic groups, the pattern is not
uniform: African American students and students who receive free and reduced price lunch are more concentrated in states that have lower proficiency cut scores for fourth-grade reading, but African American, Hispanic, and free and reduced price lunch recipients are more likely than White students to attend school in states that have higher proficiency cut scores for eighth-grade reading. The results for fourth- and eighth-grade math are more consistent: African Americans are overrepresented in states with lower performance expectations, but sizeable and nearly comparable percentages of Whites also attend school in states with lower proficiency cut scores. At the other end of the scale of expectations, slightly lower percentages of African American students attend school in states with high performance standards for eighth-grade math, compared to White students. For Hispanic and English Language Learner/Limited English Proficiency (ELL/LEP) students however, the picture is different: They are more likely to attend school in states that require higher cut scores to demonstrate proficiency.

These findings are significant for three reasons. First, they provide a new window on the question of whether poor or minority students are systematically held to lower performance standards than other students—a claim made frequently by supporters of standards-based reforms. Second, the findings raise questions about whether variation in systems-level definitions of proficiency across states has much to do with the Black-White achievement gap. Third, the article confirms that NCLB-style educational federalism produces irrationalities that significantly undermine the system of accountability envisioned by the authors of NCLB. As several scholars have pointed out, the variation in proficiency cut scores clearly indicates that not all students are held to the same standards, undermining a primary—if not the primary—aim of NCLB.

Standards-Based Accountability, Federalism, and State-Level Standards

The turn to standards and accountability is a now familiar history within the late 20th-century story of public education reform. Emerging on the heels of *A Nation at Risk*, which warned of a “rising tide of mediocrity” in U.S. schools, the standards movement was both an effort to increase national economic competitiveness and an almost moral crusade to improve education by simply demanding more of students. Conjointly with standards, the accountability movement also sought to improve performance by creating mechanisms by which students, schools, and/or teachers are held accountable for their inability to meet these newly articulated standards. Thus, accountability and standards together became the keystone reforms in many states in the 1980s and 1990s (Beneviste, 1985; Gordon, 2003). Typically these reforms established curricular content and performance standards requirements, which are generally assessed through standardized testing. In the most robust form of accountability, these tests come with a “high-stakes” consequence attached if a student or school fails to meet a state-level cut
score. Those consequences may fall either to students (in the form of promotion and/or graduation exam) or schools (in the form of increased supervision, state takeover, or other form of intervention).

The 1994 reauthorization of the Elementary and Secondary Education Act (ESEA; formally called the Improving America’s Schools Act) required states to adopt standards-based accountability mechanisms, but the law made virtually no provision for enforcement. Under Goals 2000 (another federal education law enacted just prior to the ESEA reauthorization in 1994), reformers had initially sought to create a mechanism by which a common set of educational standards could be adopted nationwide, but political resistance to those common standards doomed that effort (Superfine, 2005). With the enactment of NCLB in January 2002 however, the federal government became serious about requiring states to adopt standards-based accountability in order to receive federal funding for education. Nonetheless, this increasing concentration of educational authority in federal law has not generated political enthusiasm for national curricular standards. The commitment to educational federalism is still robust enough that uniform federal standards—either curricular or performance—are generally regarded as a political implausibility, if not impossibility. As a result, state-level curricular and performance standards have been the hallmark of the last generation of education reform and they form the fundamental benchmark mechanism for the accountability regime created by No Child Left Behind.

As befits educational federalism however, states have not adopted uniform performance standards against which to evaluate students. Not only does the substance of these state-level curricular standards-based reforms differ substantially, but the level of performance required to show mastery of these standards varies. Partly because of the varying contents of curricula and the varying tests linked to those standards, states establish widely varying cut scores to determine if a student is performing at a “basic” or “proficient” or “advanced” level. This cut score disparity across states provides the metric by which this article operationalizes the notion of an expectations gap within U.S. education. While the political project and policy implementation of educational standards are far more complex than the establishment of cut scores, the range of cut score variation is a particularly vivid (and useful) demonstration of the extent to which students in different states are expected to meet differing performance standards under NCLB’s accountability regime. The next section addresses both the nature of these disparities and employs a recently devised common metric to map the demographic distribution of students against disparities of proficiency standards.

The Problem of Test Score Comparisons Under NCLB

Not all tests are created alike. That principle is perhaps no better exemplified than in the testing regimes that exist within U.S. education. In large
Reed

part because of NCLB, every state has now promulgated performance standards against which its students are evaluated, annually, in Grades 3 through 8. The national commitment to educational federalism, however, ensures that different states employ different tests to assess student learning and, more importantly for the purposes of this article, rely upon different cut scores to establish whether students have successfully met the “proficient” or “basic” levels of achievement. This state-level variation could be sensible if a state with rigorous curricular standards or a particularly demanding test established a lower cut score than a state with a less demanding curriculum or test. From an assessment point of view, it could make sense to have a range of proficiency cut scores, depending on the state contexts. But without a common national test, or without a common benchmark of performance comparison, it is very difficult to evaluate systematically the degree to which some states expect more of their students than others, although it is clear that some states are expecting more of their students than others.

The primary response of educational researchers to this problem has been to compare percentages between the pass rates on state tests and the pass rates on the National Assessment of Educational Progress (NAEP) test within that state (Hall & Kennedy, 2006). These differences in percentages are firm evidence of the existence of divergent performance standards, but just knowing the percentages of proficient students within a particular state does not allow us to make systematic inferences across states about the populations of students studying and taking tests under those divergent standards.

Recently, however, the National Center for Educational Statistics (NCES) published a report that places the cut scores of state-level tests on a common benchmark, using the state-level NAEP test as the metric (National Center for Education Statistics, 2007). Using the distribution of student achievement in the 2005 state-level NAEP tests, NCES was able to locate a point within the state test distribution at which a similar percentage of students scored. For example, assume 30% of students in the state of Euphoria scored at or above the cut score deemed to demonstrate proficiency on Euphoria's own standards-based annual standardized test. We could then turn to the state-level NAEP test that was administered in the state of Euphoria and find the point within that test score distribution at which a comparable percentage of students demonstrated proficiency according to the NAEP's definition of proficiency. This would then give us a "NAEP-equivalent" proficiency cut score for Euphoria's test. Repeating this technique for math and reading tests at the fourth- and eighth-grade levels across the states with comparable data, the authors of the NCES study produced a common index of NAEP-equivalent cut scores of state proficiency standards for roughly 30 to 35 states (some states are excluded because of data incomparability). As an example, the distribution of these NAEP-equivalent proficiency cut scores for fourth-grade math is shown in Figure 1.

These ranked NAEP-equivalent cut scores show a significant degree of variation among states, with Massachusetts and Wyoming at the top end of
Is There an Expectations Gap?

Data and Method

Using these indices for fourth- and eighth-grade reading and math as a starting point, we can now explore the demographic distribution of students across the low, middle, and high standards states. Using the range of NAEP-equivalent cut scores for the state-level tests, I first calculated the mean and standard deviation of these ranges of scores, for fourth- and eighth-grade reading and math, shown in Table 1.
Then, using the 2005 district-level demographic data from the NCES Common Core of Data (National Center for Education Statistics, 2005) I assigned to each school district in the country the proficiency cut scores of their respective states for fourth- and eighth-grade math and reading.\textsuperscript{5} I then summed the numbers of students within particular demographic groups who attended school districts that fell within nine standard deviation bands (ranging from greater than 1.75 $SD$ below the mean to greater than 1.75 $SD$ above the mean, in 0.5 $SD$ increments). The tallies of these demographic groups (and the percentages) give us a clear picture of the demographic implications of states employing different proficiency standards. Before now, we have not had a concise description of the racial, ethnic, class, and linguistic distribution of students across differing standards of proficiency, using a common metric. Tables 2 through 5 provide those figures for fourth-grade and eighth-grade NAEP-equivalent cut scores. Note that these are not the percentages of students within each category who pass these proficiency standards. Instead, these are the percentages of students within each demographic category who attend a school district that is asked to meet that particular level of proficiency standard. Tables 2 through 5, in effect, show the demographic distribution of students who are expected to achieve at particular levels of proficiency, using a relative scale that relies upon a common metric.

Findings and Results: Two Views of the Expectations Gap

This section will provide two windows on whether there is a demonstrable expectations gap among major demographic groups in the United States, based on common metric of estimated NAEP-equivalent cut scores for state-level tests. First, I will examine the relative percentages of groups attending school districts with low or high proficiency standards. Second I will explore the proficiency levels of the district attended by the median White student, the median African American student, the median Hispanic student, the median student receiving free and reduced priced lunch (FRPL; a measure of poverty), and the median English Language Learner/Limited English Proficiency student. In addition, the article graphically represents these median student NAEP-equivalent cut scores, thereby enabling us to

\begin{table}
\centering
\caption{Means and Standard Deviations of State National Assessment of Educational Progress (NAEP) Equivalent Proficiency Cut Scores}
\begin{tabular}{lcc}
\hline
 & Math & Reading \\
 & $M$ & $SD$ & $M$ & $SD$
\hline
Fourth grade & 224.12 & 14.19 & 197.72 & 16.59 \\
Eighth grade & 271.38 & 16.69 & 245.82 & 15.01 \\
\hline
\end{tabular}
\end{table}
Is There an Expectations Gap?

see the comparative distribution of demographic groups under both the state NAEP-equivalent proficiency and the actual NAEP basic and proficient cut scores.

Relative Concentration of Demographic Groups Within Proficiency Categories

If low systems-level expectations did not fall disproportionately within particular demographic groups, we would see roughly comparable percentages of racial, ethnic, and class groups across all categories of proficiency cut scores. There would be no clustering of students by demographic group within high or low cut score districts. Unfortunately, that “equitable” distribution of lower cut scores is not entirely borne out by the evidence. Tables 2 through 5 suggest some support for lower systems-level expectations in districts with higher percentages of minority students, but the picture is far from simple. There is significant variation among racial and ethnic and class groups and between reading and math and between fourth- and eighth-grade standards levels. Indeed, African Americans are proportionately more likely to attend school districts with higher cut scores in eighth-grade reading but lower cut scores in fourth-grade math and reading and eighth-grade math. Similarly, Whites are consistently overrepresented in the highest cut-score group.

Fourth-Grade Math

Fourth-grade math shows us that systems-level expectations do vary modestly by different racial groups. Greater percentages of African American students are found in districts with lower proficiency cut scores in fourth-grade math, but not by wide margins. Moreover, there is significant variation among minority groups. For example, 11.6% of African American students attend a district whose fourth-grade math standard is between 1.75 and 1.25 SD below the mean proficiency standard. At the same time, 7.7% of Whites attend school districts with the same standard, but only 4.0% of Hispanics do so.

The view from the top end of the fourth-grade math proficiency standard is somewhat different from the bottom end, but in general, Whites are still more likely to be in the top three categories than African Americans or Hispanics. In the top three categories, Whites are overrepresented 14.9% to 11.0% for African Americans, with Hispanics seriously underrepresented at 4.3%. There are some variations: We see greater percentages of African American students attending school districts at 1.25 to 1.75 SD above the mean category than Whites. (For African Americans, 7.1% attend a school district at that standard, while 5.7% of Whites do so.) But overall, the top end of fourth-grade math scale shows a higher concentration of Whites than other racial groups.
Table 2
PK–12 Students by State Proficiency Cut Score Group and Demographic, Fourth-Grade Math

<table>
<thead>
<tr>
<th>Standard Deviation Group: 9-Point Scale</th>
<th>All Students</th>
<th>White</th>
<th>African American</th>
<th>Hispanic</th>
<th>Free and Reduced Priced Lunch</th>
<th>English Language Learner/Limited English Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>&gt; 1.75 SD below mean</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.25 to 1.75 SD below mean</td>
<td>2,632,018</td>
<td>7.3</td>
<td>1,510,000</td>
<td>7.7</td>
<td>733,154</td>
<td>11.6</td>
</tr>
<tr>
<td>1.25 to 0.75 SD below mean</td>
<td>2,926,112</td>
<td>8.2</td>
<td>1,710,000</td>
<td>8.7</td>
<td>536,959</td>
<td>8.5</td>
</tr>
<tr>
<td>0.75 to 0.25 SD below mean</td>
<td>8,652,364</td>
<td>24.1</td>
<td>4,260,000</td>
<td>21.7</td>
<td>1,680,000</td>
<td>26.7</td>
</tr>
<tr>
<td>±0.25 SD from mean</td>
<td>6,387,780</td>
<td>17.8</td>
<td>4,410,000</td>
<td>22.5</td>
<td>1,210,000</td>
<td>19.2</td>
</tr>
<tr>
<td>0.75 to 0.25 SD above mean</td>
<td>10,968,538</td>
<td>30.6</td>
<td>4,790,000</td>
<td>24.4</td>
<td>1,440,000</td>
<td>22.9</td>
</tr>
<tr>
<td>0.75 to 1.25 SD above mean</td>
<td>1,481,860</td>
<td>4.1</td>
<td>1,040,000</td>
<td>5.3</td>
<td>164,456</td>
<td>2.6</td>
</tr>
<tr>
<td>1.25 to 1.75 SD above mean</td>
<td>1,799,115</td>
<td>5.0</td>
<td>1,110,000</td>
<td>5.7</td>
<td>449,972</td>
<td>7.1</td>
</tr>
<tr>
<td>&gt;1.75 SD above mean</td>
<td>1,014,191</td>
<td>2.8</td>
<td>765,000</td>
<td>3.9</td>
<td>80,512</td>
<td>1.3</td>
</tr>
<tr>
<td>Total</td>
<td>35,861,979</td>
<td>100.0</td>
<td>19,595,001</td>
<td>100.0</td>
<td>6,295,054</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note. NV and TN omitted because of missing race and ethnicity data from Common Core of Data (National Center for Education Statistics, 2005).
### Table 3
PK–12 Students by State Proficiency Cut Score Group and Demographic, Fourth-Grade Reading

<table>
<thead>
<tr>
<th>Standard Deviation Group: 9-Point Scale</th>
<th>All Students</th>
<th>White</th>
<th>African American</th>
<th>Hispanic</th>
<th>Free and Reduced Priced Lunch</th>
<th>English Language Learner/Limited English Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1.75 SD below mean</td>
<td>481,780</td>
<td>1.4</td>
<td>233,000</td>
<td>1.3</td>
<td>250,922</td>
<td>4.3</td>
</tr>
<tr>
<td>1.25 to 1.75 SD below mean</td>
<td>1,552,758</td>
<td>4.6</td>
<td>766,000</td>
<td>4.3</td>
<td>590,038</td>
<td>10.1</td>
</tr>
<tr>
<td>1.25 to 0.75 SD below mean</td>
<td>2,397,974</td>
<td>7.1</td>
<td>1,470,000</td>
<td>8.2</td>
<td>513,980</td>
<td>8.8</td>
</tr>
<tr>
<td>0.75 to 0.25 SD below mean</td>
<td>8,477,441</td>
<td>25.2</td>
<td>4,290,000</td>
<td>23.9</td>
<td>1,330,000</td>
<td>22.8</td>
</tr>
<tr>
<td>±0.25 SD from mean</td>
<td>5,218,077</td>
<td>15.5</td>
<td>3,850,000</td>
<td>21.4</td>
<td>854,955</td>
<td>14.7</td>
</tr>
<tr>
<td>0.75 to 0.25 SD above mean</td>
<td>12,753,338</td>
<td>37.9</td>
<td>5,540,000</td>
<td>30.8</td>
<td>1,740,000</td>
<td>29.9</td>
</tr>
<tr>
<td>0.75 to 1.25 SD above mean</td>
<td>1,017,139</td>
<td>3.0</td>
<td>697,000</td>
<td>3.9</td>
<td>181,486</td>
<td>3.1</td>
</tr>
<tr>
<td>1.25 to 1.75 SD above mean</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>&gt;1.75 SD above mean</td>
<td>1,715,367</td>
<td>5.1</td>
<td>1,140,000</td>
<td>6.3</td>
<td>365,995</td>
<td>6.3</td>
</tr>
<tr>
<td>Total</td>
<td>33,613,875</td>
<td>100.0</td>
<td>17,986,001</td>
<td>100.0</td>
<td>5,827,377</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note. NV and TN omitted because of missing race and ethnicity data from Common Core of Data (National Center for Education Statistics, 2005).
Reed

In the category with the largest percentage of students (0.25–0.75 SD above the mean), there are significant ethnic and class differences. Nearly half of all Hispanics fall into this group, as do more than 35% of all students who receive free and reduced price lunch. At the same time however, only about 23% of African Americans are in this most populous group, while 24.4% of White students are in the category.

Fourth-Grade Reading

For fourth-grade reading however, the Black-White racial contrast at the low end of the cut score distribution is starker: More than 10% of African American students attend a district in which the proficiency standard is 1.75 to 1.25 SD below the mean, but less than half that percentage of Whites (4.3%) attend a school district at the same proficiency standard. Interestingly however, comparatively few Hispanic students are clustered in districts with low cut scores for fourth-grade reading. Indeed, only 1.6% of Hispanic students attend a school district at 1.75 to 1.25 SD below the mean, while nearly 56% of all Hispanics attend a school district at the 0.25 to 0.75 SD above the mean proficiency standard. Comparatively, only 30.8% and 29.9% of Whites and African Americans are in that same category. This, presumably, is due to the concentration of Hispanic students in relatively few states, with higher cut scores.

At the top end of the fourth-grade reading proficiency standards however, there is little difference between Whites and African Americans: Equal percentages of Whites and African Americans attend a school district at the top end of the fourth-grade reading scale (6.3%), and at the 0.75 to 1.25 SD category, less than 1 percentage point separates the two groups (3.9% vs. 3.1%). Thus, in fourth-grade reading, the bottom end of the scale shows an expectations gap that is distributed unevenly across racial groups, but the top end of the scale does not. While disparate percentages of Whites and African American fourth graders face relatively low proficiency cut scores in reading, comparable percentages of them are confront relatively high proficiency standards.

Eighth-Grade Math

At the level of eighth-grade math, we see only modest evidence of a racially disparate gap in systems-level expectations, at least at the low end of the scale. While 32.6% of all African American students are located in the bottom three categories (from 0.75 SD below the mean to greater than 1.75 SD below the mean), a nearly comparable percentage of 30.4% of all Whites are at similarly low proficiency cut scores. In other words, only a 2 percentage point difference exists between African American and White students in the size of the student populations at the bottom end of the eighth-grade math proficiency standard. Somewhat surprisingly, Hispanics comprise the smallest group in the bottom third, with 20.2% of Hispanics at the bottom end of
Table 4
PK–12 Students by State Proficiency Cut Score Group and Demographic, Eighth-Grade Math

<table>
<thead>
<tr>
<th>Standard Deviation Group: 9-Point Scale</th>
<th>All Students</th>
<th>White</th>
<th>African American</th>
<th>Hispanic</th>
<th>Free and Reduced Priced Lunch</th>
<th>English Language Learner/Limited English Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 1.75 SD below mean</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.25 to 1.75 SD below mean</td>
<td>1,384,850</td>
<td>4.0</td>
<td>795,000</td>
<td>3.9</td>
<td>437,200</td>
<td>6.6</td>
</tr>
<tr>
<td>1.25 to 0.75 SD below mean</td>
<td>3,847,875</td>
<td>11.2</td>
<td>2,280,000</td>
<td>11.1</td>
<td>855,078</td>
<td>12.8</td>
</tr>
<tr>
<td>0.75 to 0.25 SD below mean</td>
<td>4,810,645</td>
<td>13.9</td>
<td>3,190,000</td>
<td>15.5</td>
<td>880,315</td>
<td>13.2</td>
</tr>
<tr>
<td>±0.25 SD from mean</td>
<td>17,151,857</td>
<td>49.7</td>
<td>9,820,000</td>
<td>47.6</td>
<td>3,020,000</td>
<td>45.4</td>
</tr>
<tr>
<td>0.75 to 0.25 SD above mean</td>
<td>3,035,996</td>
<td>8.8</td>
<td>1,690,000</td>
<td>8.2</td>
<td>753,285</td>
<td>11.3</td>
</tr>
<tr>
<td>0.75 to 1.25 SD above mean</td>
<td>1,457,348</td>
<td>4.2</td>
<td>975,000</td>
<td>4.7</td>
<td>181,054</td>
<td>2.7</td>
</tr>
<tr>
<td>1.25 to 1.75 SD above mean</td>
<td>267,113</td>
<td>0.8</td>
<td>109,000</td>
<td>0.5</td>
<td>5,595</td>
<td>0.1</td>
</tr>
<tr>
<td>&gt;1.75 SD above mean</td>
<td>2,546,193</td>
<td>7.4</td>
<td>1,770,000</td>
<td>8.6</td>
<td>524,889</td>
<td>7.9</td>
</tr>
<tr>
<td>Total</td>
<td>34,501,878</td>
<td>100.0</td>
<td>20,629,001</td>
<td>100.0</td>
<td>6,657,417</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note. NV and TN omitted because of missing race and ethnicity data from Common Core of Data (National Center for Education Statistics, 2005).
the eighth-grade math proficiency cut score distribution. Focusing on the largest category within the bottom third (1.25–0.75 SD below the mean), there is a slight difference between Blacks and Whites: 12.8% of all African American students are concentrated in this category, while 11.1% of Whites are at this expected proficiency cut score range.

At the midrange of eighth-grade math proficiency cut scores, significantly greater percentages of Hispanics are found clustered around the mean. Fully 65% of all Hispanics are in this category, while 47.6% of Whites are found at this level (0.25 SD plus or minus around the mean). A comparable percentage of African Americans (45.4%) are similarly clustered around the mean, a figure identical to the percentage of students in poverty at that level.

The top third of the distribution shows us a similar picture: Slightly lower percentages of African Americans face higher proficiency cut scores than Whites, but by a relatively small gap. Taking the top three categories as a whole, 13.8% of Whites are enrolled in school districts at the top end of the eighth-grade math proficiency cut score range, while 10.7% of African American students are enrolled in districts with similar definitions of proficient. Again, Hispanics are underrepresented in the top three categories at 6.9%.

In sum, eighth-grade math proficiency cut scores show us that African Americans are overrepresented in districts that have low proficiency cut scores and Hispanics are overrepresented in districts that cluster around the mean, compared to Whites and all students. At the top end, the disparities continue, with modestly lower percentages of African Americans and significantly lower percentages of Hispanics enrolled in districts with the highest eighth-grade math proficiency cut scores.

**Eighth-Grade Reading**

In eighth-grade reading, the picture is somewhat mixed, at least for African American students. The low end of the distribution tells us a now-familiar story of modest overrepresentation of African American students: Roughly 24% of Whites and 27.5% of African American students are in the bottom three categories. We see a slightly higher percentage of Hispanics, 29.2%, enrolled in these low proficiency standard districts. Thus, in contrast to eighth-grade math, eighth-grade reading sees overrepresentation of Hispanic students within the bottom tier of proficiency standards.

At the top end of the distribution of eighth-grade reading proficiency standards however, we see little demonstration of an expectations gap, as revealed by proficiency cut scores. More than 50% of all Hispanic students are clustered in the top three categories of proficiency standards, as are 34.3% of African American students. A similar percentage of White students, 33.5%, are in districts whose proficiency cut scores lie within the top three categories. In the top two proficiency cut score groups in eighth-grade reading, we see higher percentages of African Americans than any
Table 5

PK–12 Students by State Proficiency Cut Score Group and Demographic, Eighth-Grade Reading

<table>
<thead>
<tr>
<th>Standard Deviation Group: 9-Point Scale</th>
<th>All Students</th>
<th>White</th>
<th>African American</th>
<th>Hispanic</th>
<th>Free and Reduced Priced Lunch</th>
<th>English Language Learner/Limited Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 1.75 SD below mean</td>
<td>1,384,850</td>
<td>3.8</td>
<td>794,978</td>
<td>4.0</td>
<td>437,200</td>
<td>6.7</td>
</tr>
<tr>
<td>1.25 to 1.75 SD below mean</td>
<td>5,952,795</td>
<td>16.1</td>
<td>2,430,000</td>
<td>12.3</td>
<td>1,210,000</td>
<td>18.5</td>
</tr>
<tr>
<td>1.25 to 0.75 SD below mean</td>
<td>2,036,701</td>
<td>5.5</td>
<td>1,500,000</td>
<td>7.6</td>
<td>150,945</td>
<td>2.3</td>
</tr>
<tr>
<td>0.75 to 0.25 SD below mean</td>
<td>3,216,443</td>
<td>8.7</td>
<td>2,420,000</td>
<td>12.3</td>
<td>459,594</td>
<td>7.0</td>
</tr>
<tr>
<td>±0.25 SD from mean</td>
<td>6,130,060</td>
<td>16.6</td>
<td>3,530,000</td>
<td>17.9</td>
<td>1,310,000</td>
<td>20.0</td>
</tr>
<tr>
<td>0.75 to 0.25 SD above mean</td>
<td>3,944,060</td>
<td>10.7</td>
<td>2,450,000</td>
<td>12.4</td>
<td>736,271</td>
<td>11.2</td>
</tr>
<tr>
<td>0.75 to 1.25 SD above mean</td>
<td>8,140,425</td>
<td>22.1</td>
<td>3,330,000</td>
<td>16.9</td>
<td>784,746</td>
<td>12.0</td>
</tr>
<tr>
<td>1.25 to 1.75 SD above mean</td>
<td>5,315,388</td>
<td>14.4</td>
<td>2,830,000</td>
<td>14.3</td>
<td>1,170,000</td>
<td>17.9</td>
</tr>
<tr>
<td>&gt; 1.75 SD above mean</td>
<td>785,522</td>
<td>2.1</td>
<td>450,024</td>
<td>2.3</td>
<td>286,661</td>
<td>4.4</td>
</tr>
<tr>
<td>Total</td>
<td>36,906,045</td>
<td>100.0</td>
<td>19,735,003</td>
<td>100.0</td>
<td>6,545,418</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note. NV and TN omitted because of missing race and ethnicity data from Common Core of Data (National Center for Education Statistics, 2005).
other demographic. For LEP/ELL students, the contrast is even more striking: 53.9% of all LEP/ELL students attend school districts in which the eighth-grade reading proficiency cut scores are at least 0.75 standard deviations above the mean proficiency cut score. Eighth-grade reading, then, shows us that the argument that minority groups are systematically held to a lower proficiency standard is not accurate.

Based on these percentages, it is clear that an expectations gap, while it does exist, is not uniform, at least as measured by proficiency cut scores, across fourth- and eighth-grade reading and math. African Americans are overrepresented at the low end in fourth-grade reading but also overrepresented at the high end in eighth-grade reading. Hispanics tend to be clustered more toward the middle of the distribution and Whites are, typically, overrepresented within the top ranges for fourth-grade math and reading and eighth-grade math, but not by overwhelming percentages. In eighth-grade reading in fact, the overrepresentation of African Americans at the top end of the proficiency scale is larger than overrepresentation of African Americans at the bottom end of the scale for fourth-grade math, when compared to Whites.

**Where Do the Median Students Lie?**

A second way of examining a possible expectations gap is to determine the proficiency cut score that a hypothetical median student confronts and to determine whether meaningful differences exist in the location of the median student of various demographic groups. For our purposes, the median student is simply the student located at the midpoint of students along the range of proficiency scores. For each demographic group, the median student lies at the point where half the students of that demographic group face proficiency standards lower than the median student and half the students face proficiency standards that are higher. This metric enables us to compare directly the average experience of White students or Hispanic or African American students, rather than examine their clustering within particular proficiency cut score groupings.

The results for six demographic groups (White, African American, Hispanic, ELL/LEP, FRPL recipients, and all students) are presented in Table 6. The same results are produced graphically in Figures 2 and 3, with the actual NAEP cut scores for basic and proficient also represented. Again, these are *not* the median scores produced by students within these demographic groups; rather, these figures show us the proficiency cut score at which half the students within each demographic group face a higher cut score and half face a lower cut score.

If we look at the median fourth-grade reading and math cut scores for each group, we see some interesting relationships. First, the median proficiency cut scores for Whites in fourth grade are the second lowest cut scores, with only the median African American student facing lower proficiency cut scores. The median Hispanic student in poverty and median LEP/ELL student
Table 6
National Assessment of Educational Progress (NAEP) Equivalent Proficiency Cut Scores of Median Student, by Demographic Group

<table>
<thead>
<tr>
<th></th>
<th>All Students</th>
<th>White</th>
<th>African American</th>
<th>Hispanic</th>
<th>Free and Reduced Priced Lunch</th>
<th>English Language Learner/ Limited English Proficiency</th>
<th>NAEP Basic</th>
<th>NAEP Proficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourth-grade reading</td>
<td>206.0</td>
<td>203.2</td>
<td>198.0</td>
<td>204.6</td>
<td>206.2</td>
<td>205.2</td>
<td>208</td>
<td>238</td>
</tr>
<tr>
<td>Fourth-grade math</td>
<td>228.8</td>
<td>228.0</td>
<td>222.0</td>
<td>228.4</td>
<td>231.7</td>
<td>230.8</td>
<td>214</td>
<td>249</td>
</tr>
<tr>
<td>Eighth-grade reading</td>
<td>256.4</td>
<td>253.6</td>
<td>253.7</td>
<td>257.2</td>
<td>255.0</td>
<td>268.3</td>
<td>243</td>
<td>281</td>
</tr>
<tr>
<td>Eighth-grade math</td>
<td>274.2</td>
<td>274.1</td>
<td>273.6</td>
<td>274.8</td>
<td>272.9</td>
<td>272.4</td>
<td>262</td>
<td>299</td>
</tr>
</tbody>
</table>

Note. NV and TN omitted because of missing race and ethnicity data from Common Core of Data (National Center for Education Statistics, 2005).

Figure 2. National Assessment of Educational Progress equivalent proficiency cut scores of median student, by race or ethnicity.
all confront fourth-grade reading and math proficiency cut scores that are higher than those faced by the median White student. While the differences may not be large (particularly for math), the relative placement of Whites among these groups strongly challenges the notion that all minority groups uniformly face lower systems-level expectations.

Nonetheless, it is clear that the median African American student faces a lower proficiency cut score than all other groups. It is in fourth grade that we see the clearest expression of an expectations gap. The median African American fourth grader faces a lower proficiency cut score in both math and reading than the median student in poverty, the median ELL/LEP student, the median Hispanic student, and the median White student. If 10 NAEP points translate into a full grade-level year of learning, the median African American fourth grader who achieves proficiency under state definitions in math or reading is one half year behind the median White fourth grader who achieves proficiency in those same subjects.

Figure 3. NAEP equivalent proficiency cut scores of median student, by poverty and language.
Note. FRPL = free and reduced priced lunch; LEP/ELL = Limited English Proficiency/English Language Learner.
In the eighth grade however, those gaps are largely gone. The median African American eighth grader faces an NAEP-equivalent proficiency cut score in reading that is essentially the same as the White median student (253.7 for African Americans, 253.6 for Whites) and the gap in math between the two groups in the eighth grade is negligible (273.6 vs. 274.1). In fact, in terms of race and ethnicity, the median cut scores for both reading and math converge dramatically in eighth grade, with the greatest spread between African Americans and Hispanics, with the median Hispanic facing an eighth-grade reading proficiency cut score 3.5 points higher than the median African American.

One surprising result in the eighth grade is that ELL/LEP students face, on average, a higher proficiency reading cut score than any other demographic group. In fact, the median ELL/LEP student confronts a cut score that is nearly 15 points higher than the cut score facing the median White student. Again, if an increase of 10 NAEP points equals 1 year of schooling, our system of educational federalism defines proficiency for the median ELL/LEP student at a level that represents 1½ years beyond the level of proficiency it asks the median White student to meet. While this result is most likely due to the concentration of large numbers of ELL/LEP students in relatively high proficiency standard states, it also highlights the inequities that educational federalism is capable of creating.

Overall, these median student calculations show a clearer expectations gap for the median African American student in the fourth grade, in both reading and math, when compared to the median students of other demographic groups, but little racial or ethnic difference in the formal proficiency expectations for eighth-grade students. In addition, we hold the median ELL/LEP student to a higher proficiency standard in reading than we do for the median student in any other group.

**Discussion and Policy Implications**

Several important issues emerge from this analysis. First, something akin to an expectations gap does exist, but it is not monolithic. The average African American student does face proficiency cut scores that are lower than the proficiency cut scores facing the average White student, particularly in elementary school. That disparity, however, does not uniformly apply to all racial and/or ethnic minorities or across all grade spans. Hispanics are more likely to attend school districts in states that have relatively high standards for proficiency. Also, the eighth-grade results for both math and reading show that the expectation gap in eighth grade is modest, at best.

While the determinants of (and changes in) the Black-White achievement gap have been extensively studied (Hall & Kennedy, 2006; Jencks & Phillips, 1998; Lee, 1998, 2004), there is much less agreement on the relationship between higher standards and closing the achievement gap. To date, the policy debate surrounding high standards has been characterized by two
distinctive understandings of the consequences of setting high performance expectations. On one side, scholars, analysts, and policy advocates have contended that high performance expectations (either in the form of high-stakes testing and increased graduation requirements or in the form of higher standards for proficiency) will induce greater achievement and more rigorous course taking and improve educational opportunity for previously disadvantaged groups (Alexander, 2002; Carnoy & Loeb, 2002; Harris & Herrington, 2006; Haycock, 2001). Indeed, Peterson and Hess (2006) contended that setting high standards is a “crucial first step” in realizing the promise of NCLB (p. 29).

Another side of the debate has focused more on the undesired or unanticipated consequences of heightened standards, high-stakes testing, and high-threshold cut scores—reductions in equity within U.S. schools (Kim & Sunderman, 2005), increasing drop-out and force-out rates (Jacob, 2001), higher rates of grade-level retentions (Haney, 2008), the narrowing of the school districts curricula (Tyree, 1993), and in general, a myopic view of what counts as learning (Darling-Hammond, 1994).

The primary finding presented here—that the expectations gap, while real, is relatively small—recasts this debate over ratcheting up of systems-level proficiency expectations and academic achievement. Compared to the achievement gap, the expectations gap is both much smaller and much less consistent. There is only a negligible eighth-grade Black-White expectations gap. The median Hispanic student faces a higher proficiency cut score than the median White student. Yet both Black and Hispanic students achieve at much lower levels than the median White student. Even where there is something of a Black-White expectations gap, it does not approach the size of the Black-White achievement gap: Comparing the size of the Black-White fourth-grade reading expectations gap (5 NAEP points, at most) to the size of the Black-White fourth-grade reading achievement gap (27 points in 2004, 24 points in 2008; National Center for Education Statistics, 2009), it becomes clear that other factors besides divergent proficiency standards are driving the continued existence of the achievement gap.

Those who argue that reducing the Black-White achievement gap requires raising systems-level proficiency expectations for African American students can only point to modest evidence that African American students systematically confront definitions of proficiency lower than the definitions of proficiency confronting White students. Aside from fourth-grade reading proficiency standards, there is not much evidence of this. Taking proficiency standards out of the policy discussion of the Black-White test score gap enables us to devote greater attention to other issues that may be more important determinants of the gap. Indeed, the system-level definitions of what constitutes “proficient” most likely have little to do with the daily routines, resources, and informal expectations within a student’s life that will enable him or her to meet those standards (Elmore & Fuhrman, 1995; Greenwald, Hedges, & Laine, 1996; Lee & Wong, 2004). Establishing higher
proficiency cut scores for greater numbers of African Americans in fourth grade will quite possibly accomplish little if the daily expectations (and abilities) of teachers, principals, staff, and parents are not correspondingly high. And as recent studies have shown, the ratcheting up of systems expectations most likely will only harm students at risk if they are imposed in a punitive, high-stakes fashion (Heilig & Darling-Hammond, 2008).

But beyond the relationship of a possible Black-White expectations gap to the Black-White achievement gap, other issues also merit the attention of policy makers. In particular, these findings suggest that the present accountability regime’s reliance on state-level standards helps produce perverse or irrational policy outcomes, outcomes that are desired by no one and may exacerbate the least desirable consequences of NCLB. In particular, the system’s exposure of the median LEP/ELL student to an eighth-grade reading proficiency cut score that is 15 points higher than the cut score confronting the median White student is simply irrational. Setting a higher bar for precisely those students who understandably will have difficulty meeting even the median student cut score is virtually impossible to justify on policy grounds. This consequence of educational federalism sustains criticisms that NCLB is not promoting equality of educational opportunity for LEP/ELL students but represents an unfair double standard.

So, what is to be done? What ought policy makers do in light of this evidence? Does the expectations gap merit significant direct attention from policy makers? While the existing system of accountability and proficiency standards may, because of its commitment to educational federalism, produce irrational performance expectations, or define proficiency differently for a median White student than for a median African American student, we need to consider the trade-offs inherent in removing these elements from the existing system. Let us consider three obvious options.

**Option 1: Create a system of uniform national proficiency standards.** In a technical sense, the simplest policy response would be to create a common national test with uniform cut scores to assign students to such performance categories as basic, proficient, and advanced. This response would effectively nationalize the accountability regime created by NCLB, and without changes in the sanctions that follow a school’s consistent failure to meet adequate yearly progress (AYP), these federally determined proficiency cut scores would, presumably, determine whether some local schools stay open. This option is politically unpalatable for a variety of reasons, but it clearly strikes at the heart of a constitutional order of public education in the United States that imposes an affirmative obligation on states to operate schools, under state constitutional provisions (Reed, 2001). While there may be defensible, even sensible, rationales for a nationalized accountability system, the politics of its enactment would most likely produce a system of proficiency standards that would be relatively low, cut scores that a clear majority of U.S. students would readily pass. Thus, the politics of a nationalized accountability system would undermine one of the very rationales used for
its enactment (the use of high proficiency standards to induce higher achievement among students). A uniform system of low proficiency standards would do little to change the nature of the achievement gap and could very well induce even greater constriction of school curricula. Absent changes in the sanctions imposed on schools that fail to meet AYP, a national system of proficiency standards, employing a common test, could very well create more problems than it solves.6

Option 2: Scrap NCLB-style standards-based accountability entirely. An alternative would be to eliminate proficiency standards as the basis of state-level accountability regimes. While the absence of a federal requirement that states use proficiency cut scores on standardized tests to demonstrate adequate yearly progress could induce some states to abandon standards-based reform, evidence exists that there is broad political and public support for the existence of standards (Hochschild & Scovronick, 2003). Parents and politicians very much like the idea of using test scores to demonstrate the caliber of school, and there is no certainty that the removal of a federal requirement would lead to the demise of proficiency cut scores. Moreover, the withdrawal of the federal government from a standards-based accountability regime could lead to greater, not less, variation in proficiency cut scores among states, as states’ educational leaders respond differentially to the social and political contexts of public education within their states, without consistent federal pressure. That increased variation could in turn induce more of a Black-White expectations gap.

It is not entirely clear that the presence of a federal accountability regime has created the expectations gap (it may have preexisted NCLB), and it is equally clear that a federal abandonment of standards-based accountability would not necessarily erase it. To be sure, the federally required sanctions that fall upon schools and school districts as a result of the federal accountability regime are undoubtedly prompting all manner of cut score “gamesmanship” at the state and district levels in an effort to evade those sanctions, but allowing states free rein would most likely remove a key positive benefit of NCLB: consistent policy attention from state and local school officials to the Black-White test score gap. Absent federal pressure, there may not be sufficient internal state-level commitment to address what has been called “the most significant educational challenge facing American society in the 21st century” (Kim & Sunderman, 2005, p. 3).

Option 3: Pay less attention to “proficiency” and more attention to capacity. Assuming, then, that some form of federal oversight of the existing accountability mechanism has some positive benefit and is most likely politically durable, we might best focus our policy energies on those states whose proficiency cut scores are accounting for the greatest irrationality within the present system. In other words, we could alleviate a Black-White expectation gap by increasing the reading proficiency cut scores in those low cut score states that enroll a significantly higher percentage of African American
Is There an Expectations Gap?

students. Similarly, we could lower the proficiency cut scores in those high proficiency standard states with larger percentages of ELL/LEP students.

Two issues, however, complicate this scenario. First, raising cut scores in low proficiency standards states that enroll comparatively higher percentages of African Americans makes no sense if those states and their local districts lack the capacity and resources to ensure that students will have the educational opportunity to meet the raised standard. Simply raising cut scores in order to reduce a fourth-grade reading Black-White expectation gap without paying any attention to the contexts of learning within those states and districts would only increase the specific, negative consequences of a more robust accountability regime, including higher numbers of dropouts, force-outs, and grade-level retentions (Heilig & Darling-Hammond, 2008). Policy makers need to be mindful of the capacity of states and local districts to provide opportunities to learn before rushing to adopt higher proficiency standards.

Second, we need to be aware that the attention paid to the racial dimension of a proficiency cut score expectations gap would not, in itself, eliminate differential expectations among students. If this “triage” were successful, some students would still be held to higher proficiency standards than others; there simply would not be an overrepresentation of one racial or demographic group at either the high or low end of the distribution of proficiency cut scores. The “triage” policy intervention does nothing to eliminate the spectrum of proficiency expectation—which is a product of educational federalism—but seeks only to ensure that there is no racial or class bias to the distribution of students along that spectrum. Even if the policy achieves its objective, there still would be students who would aim at—and achieve—a proficiency standard that is lower than the proficiency other students would aim at and quite likely achieve. As a result, we would still confront the normative objection that differential standards exist, and despite our removal of a racial or ethnic cast to those differential standards, one could nonetheless argue that lower proficiency standards for some students is inherently unfair. The unfairness of a lower proficiency standard would be even more acute if a state established lower levels of proficiency because of the state’s inability to deliver a high-quality education to its students, which would in turn enable them to meet a higher proficiency standard.

Educational federalism, by its very nature, allows for variation in the establishment of proficiency definitions. An NCLB-style accountability system, operating within educational federalism, will always have to confront those charges of unfairness, whether racially tinged or not. The questions for policy makers is whether the benefits of educational federalism outweigh claims about its inherent unfairness.

Conclusion

The evidence presented here shows that educational federalism’s commitment to state-level definitions of proficiency standards is not entirely neutral among major demographic groups. When viewed from a national
level, particular groups are held to modestly differing systems-level expectations about what constitutes proficient performance in math and reading. Determining the best policy response to this condition may require analysts to more fully understand what determines the range of proficiency cut scores that states establish in the first place. This article provides an account of what the distribution of proficiency cut scores across major demographic groups looks like; we now need to explain how states establish proficiency cut scores and what the determinants of those cut score levels are. Those kinds of explanations are important because they could provide insight into whether there are rational and important relationships between state educational contexts and state-level proficiency standards that need to be preserved. That is, if states are establishing differing proficiency expectations because of sensible differences in local political culture or an incapacity of local systems to meet educational need, those factors would help us determine the desirability of a single, uniform set of proficiency standards. Research suggests that there are important relationships between state-level policy contexts and student outcomes (Elmore & Fuhrman, 1995; Greenwald et al., 1996; Lee, 1998). We now need to determine the relationships between those state-level characteristics and contexts and the proficiency expectations that states establish for children. If no consistent, compelling, or predictable reasons explain why states at the bottom end of the distribution cluster there, then a single, or uniform, national level response is more justifiable. If, however, we find consistent and sufficiently compelling reasons or justifications for the existing range of proficiency scores among states, then supporters of educational federalism could legitimately argue that the present system is achieving political or policy goals that a common national system could not. Until we can better account for the variations that exist at the state level, we have no basis upon which to make those arguments.

Notes

The author would like to thank the editors of the American Educational Research Journal and the anonymous reviewers for their very helpful comments and suggestions. This research was supported by the Carnegie Corporation of New York. The author is solely responsible for the views expressed in this article.

1Educational federalism is defined here as a system of educational governance in which states, not the U.S. government, bear the primary legal and constitutional obligation to operate public schools, but they must operate those schools within parameters established by the U.S. government in order to obtain federal financial assistance for education.

2The widely cited phrase “soft bigotry of low expectations” is indicative of the kind of claim made by those who contend that schools harm minority students when they do not impose uniformly high performance standards on all students. Because this article seeks to determine if a differential exposure of minority or poor students to lower proficiency cut scores exists, it is, in some way, a test of this “soft bigotry” thesis, but I do not attribute any bias or ill motive on the part of officials in charge of establishing these cut scores. Indeed, the political or educational rationales for establishing proficiency cut scores at particular points is a topic that very much needs further investigation.
Is There an Expectations Gap?

However, Goodwin Liu (2006) recently sought to make both a constitutional and historical argument for a much greater federal role in the provision of an educational minimum, based on the citizenship clause.

The number of states for which National Assessment of Educational Progress (NAEP) equivalents could be calculated for 2005 test scores ranged from 32 to 36: 32 states for fourth-grade reading, 34 states for eighth-grade reading, 33 states for fourth-grade reading, and 36 states for eighth-grade math.

I could have also used state-level rather than district-level demographic data to perform these calculations. I chose district-level data, however, because I also want to explore, in another paper, the extent to which there are correlations between interdistrict school expenditure levels within the United States and the exposure of students to different proficiency cut scores. That paper asks (in a fashion similar to this article) whether students exposed to lower levels of educational expenditures are also exposed to lower proficiency cut scores. Relying on state data would not enable me to capture interdistrict variation in expenditure levels, and I wanted to ensure that the two sets of data were comparable. Thus, this article uses summed district-level data rather than just state-level data.

A variation on this option would be to create a common national test with uniform cut scores but eliminate federally imposed sanctions if a school fails to make adequate yearly progress. This "de-fanged" accountability system may provide useful information about the performance of schools and school districts, but according to the supporters of school accountability would do little to induce changes in school practices and expectations necessary to increase student achievement.

References


Manuscript received February 4, 2009
Revision received May 21, 2009
Accepted May 21, 2009