Proficiency Levels Vary Across States, Subjects, and Grades

Willard R. Daggett, Ed.D., International Center for Leadership in Education

International Center for
Leadership in Education
Rigor, Relevance, and Relationships for ALL Students

Proficiency Levels Vary Across States, Subjects, and Grades

by Willard R. Daggett, Ed.D.
International Center for Leadership in Education

In its effort to move public education towards a more rigorous and relevant standard of achievement for *all* students, the *No Child Left Behind Act of 2001* remains a work in progress. While the end goal is, and should remain, to have 100% of American students proficient in English, math, and science, the legislation does not define what "proficient" means. Hence, although the concept of proficiency itself – the minimum achievement a student must exhibit to be deemed proficient – is fairly constant among states, each state has complete autonomy in defining what that minimum achievement level is. The result is a great deal of variation among states in the level of achievement and learning necessary to be "proficient."

Comparing Proficiency Levels across States

A recent study by the Northwest Evaluation Association (NWEA), of Portland, Oregon, examined the levels of academic achievement used by 14 states to set proficiency levels for their high-stakes reading and math tests. NWEA then compared results – for example, how do the grade 8 reading proficiency target levels compare across states. The NWEA analysis confirmed that "proficient" performance is not consistent between states. This is not surprising in view of the independence given to states in setting goals.

NWEA also looked at proficiency levels by grade and by subject *within individual states*. The researchers found that proficiency levels are not consistent across grade levels or subject areas within a state. The achievement level at which students in a grade are considered proficient in math are not necessarily equivalent to the level required for the same students to be considered proficient in reading. Moreover, growth expectations between grade levels in the same subject are neither consistent nor what one might expect in terms of genuine academic progress.

The NWEA assessment instruments used in the study were designed to align with the content standards of each state, thus allowing the results from the different states to be placed on a common measurement scale, which NWEA refers to as a RIT scale. In each state, at least 1,000 students in each grade took the mandated state test and an NWEA test. The results of this testing were used to establish a common basis for comparing state proficiency test results and the relative proficiency levels set for all states in the study. From the data collected, NWEA was also able to use RIT scale calibrations to determine and compare the growth expectations within subjects tested and between states studied with

respect to the states' own definitions and relative levels of proficiency across grades. The complete study, including detailed methodology, is available from www.nwea.org.

Tables 1a and 1b show the results of the NWEA study relative to proficiency levels on state reading and mathematics assessments in South Carolina, California, Montana, and Wyoming. "Cut score" is the state's proficiency level equated to the RIT scale. The columns labeled "%ile" reflect the *percentage of the total tested student population at that grade in all 14 states studied who performed at or below the cut score.* From the tables, we can see that a 6th student with a cut score of 221 on the NWEA reading assessment instrument, for example, would exceed the proficiency levels set in California and Montana but not in South Carolina. To meet proficiency in South Carolina, that student would have had to score one point higher (222 on the RIT scale) and thus finish better than 63 percent of the tested student population in 6th grade reading.

Table 1a – Cut Scores Relative to Proficiency Levels for State Reading Assessments in CA, MT, SC and WY

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	205	67	**WY**	214	73	**SC**	220	73	**SC**	221	63	**SC**	227	70	**WY**	232	74	**MT**	224	43	**MT**	224	44
CA	200	51	**SC**	213	70	**CA**	214	54	**CA**	216	46	**CA**	221	50	**SC**	230	68	**SC**	-	-	**CA**	208	14
MT	193	35	**CA**	205	46	**MT**	206	35	**MT**	211	35	**MT**	218	43	**CA**	226	54	**CA**	-	-	**SC**	-	-
WY	-	-	**MT**	196	26	**WY**	-	-	**WY**	-	-	**WY**	-	-	**MT**	219	35	**WY**	-	-	**WY**	-	-

Table 1b – Cut Scores Relative to Proficiency Levels for State Math Assessments in CA, MT, SC and WY

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	208	75	**WY**	221	83	**SC**	227	76	**SC**	235	78	**SC**	242	78	**WY**	257	89	**MT**	242	47	**MT**	247	40
CA	204	60	**SC**	217	74	**CA**	225	70	**CA**	230	67	**CA**	238	70	**SC**	251	80	**SC**	-	-	**CA**	232	13
MT	197	39	**CA**	212	59	**MT**	212	38	**MT**	211	35	**MT**	224	42	**CA**	240	59	**CA**	-	-	**SC**	-	-
WY	-	-	**MT**	205	39	**WY**	-	-	**WY**	-	-	**WY**	-	-	**MT**	228	36	**WY**	-	-	**WY**	-	-

Source: Northwest Evaluation Association, The State of State Standards: Research investigating Proficiency Levels In Fourteen States.

Looking at the Wyoming data for grades 4 and 8, one could question the fairness of the state's expectations for proficiency. For example, nearly 75 percent of students tested in the 14 states would not meet minimum proficiency on Wyoming's reading assessments in t hose grades. Conversely, just looking at the data alone, one could question the level of rigor of Montana's proficiency tests. It appears that whether or not a school is doing a good job in getting students to proficiency could be as much a matter of state residency as actual performance. The data clearly shows a significant lack of consistency.

The NWEA report provides data that depicts vast differences across states in how NCLB proficiency levels are both set and achieved. Consider, for example, a student in Montana whose performance on the 8th grade math test equates to 256 on the NWEA RIT scale. According to how Montana categorizes proficiency for 8th grade math, that student would not just meet proficiency, but

exceed it by 27 cut score points. However, if that same student lived in Wyoming, he or she would not necessarily be deemed proficient, since Wyoming's cut score for minimum proficiency is 257 on the RIT scale.

Comparing Proficiency Levels across Grades within a State

The NWEA data also allows for comparison of proficiency levels within a state in a single subject across grades. A careful look at the data shows that some states are more consistent in how they set proficiency levels across grades than others. States that exhibit consistency can be expected to experience fewer problems in meeting adequate yearly progress (AYP) than those whose proficiency levels are less well aligned. NWEA uses the term "calibration" to describe how proficiency standards are set across grades, defining a calibrated standard as "one in which the minimum score required for proficiency at the exit grade is not substantively easier or more difficult than the standard at earlier grades."

Table 2 – Cut Scores Relative to Proficiency Levels for States' Reading Assessments in AZ, CA, and SC

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	205	67	SC	213	70	SC	220	73	SC	221	63	SC	227	70	SC	230	68	SC	-	-	SC	-	-
CA	200	51	CA	205	46	CA	214	54	CA	216	46	CA	221	50	CA	226	54	CA	-	-	CA	208	14
AZ	190	29	AZ	-	-	AZ	210	45	AZ	-	-	AZ	-	-	AZ	224	49	AZ	-	-	AZ	-	-

Source: Northwest Evaluation Association, The State of State Standards: Research Investigating Proficiency Levels in Fourteen States.

Table 2 shows the results of the NWEA study on cut scores relative to proficiency levels of performance on state reading assessments in South Carolina, California, and Arizona. Notice that South Carolina's cut scores increase somewhat linearly from grade 3 through grade 8 and the percentage of students scoring at or below the cut score is consistently around the 70th percentile mark. These numbers indicate that South Carolina has high expectation for student performance and their tests are indeed calibrated. California exhibits similar calibration from grade 3 through grade 8; however, the cut score of 208 at grade 10 is a substantial drop off from grade 8. This cut score falls between the 4th and 5th grade proficiency levels set in California, indicating that California's expectations for high school students – and teachers – are low in comparison to those in middle or elementary school. Moreover, a student in California taking the 10th grade reading assessment would have to only place in the 15th percentile or above relative to students in the 14 states to be deemed proficient.

Arizona is another example of a state whose calibration is questionable, as revealed byby percentiles. The minimum score required for proficiency at the grade 8 is far more difficult than the standard at grade 3 based on the percentile rankings. Take, for example, a student who has a cut score that

places him/her at the 48th percentile on the 3rd grade reading assessment in Arizona. That student would be considered proficient – and by a wide margin. If that student got a cut score that placed him/her in the same percentile in 8th grade, the student would not meet the minimum proficiency requirement. This would not necessarily indicate that the student's performance dropped off after 3rd grade; it simply means that he or she did not increase performance enough to meet the higher expectations of 8th grade. This situation could reflect poorly on middle school reading teachers, who might be viewed as not bringing enough students up to proficiency, even though the bar has been raised considerably. NWEA's report concluded that "applied to NWEA's norms, between 27 to 32 percent of the tested population [in the 14 states] may be identified as proficient in grade 3, but would not be performing at a level needed to achieve proficiency in grade 8."

A major provision of NCLB is that each school must exhibit adequate yearly progress by meeting increasingly higher annual targets for the percent of students who achieve proficiency in mathematics and reading. Arizona, and other states that set similarly inconsistent proficiency levels, will have difficulty in meeting the AYP goals because of the big jump between proficiency levels in grade 3 and grade 8, for example. Even though some states are not be consistent in establishing its proficiency levels across grades, schools identified as not passing muster may be subjected to criticism by the state, the media, and the community. These stakeholders would likely be unaware of the dynamics of the situation regarding proficiency, what it means, and how arbitrarily it may be set in some instances.

Now that NCLB is in its third year since enactment, states are beginning to understand the full implications and, in many cases, are being reactive rather than proactive in their proficiency-setting strategies. Several states have already redefined their proficiency levels, typically by lowering or softening their standards of proficiency so student performance "looks" better – at least on paper. The 2003 Princeton Review of state accountability systems uses Michigan as an example: "Michigan went from 1,513 failing schools last year to 216 this year not by having more successful students but by redefining a proficient school from one where 75 percent of students are proficient to one where 38 percent of students are."

Comparing Proficiency Levels across Subjects within a State

There's more. In a vast majority of the 14 states, the standards set for mathematics proficiency were higher, according to the percentiles, than those for reading (Oregon was the sole exception to this). NWEA theorizes about the underlying causes of this phenomenon in its report, but concedes that there is no clear answer, offering that it may have much to do with differences in culture between math and reading subject-area experts, who play a role in determining the standard-setting process. The report

deftly wonders, "… what happened when some states set math standards that cannot currently be attained by most college bound students and a reading standard that is below the 50th percentile."

Establishing Appropriate Levels of Proficiency

Regardless of whether proficiency levels are set too high or too low, the AYP provisions of NCLB are a step in the right direction. AYP raises the bar by giving educators incentives for helping students get to or maintain proficient status from one grade to the next. Proficiency is not indicative of appropriate and necessary academic growth, however. NCLB requirements do not ensure that an individual student is acquiring the skills necessary to be successful in the world beyond high school. In other words, proficiency can be set by state standards and measured by AYP, but a new and different set of standards is needed to measure and assure individual student proficiency. Consider the example of reading proficiency.

For the past several decades, and more recently in response to NCLB, schools have placed a great deal of importance on preK-6 reading initiatives. Little attention and few resources have been focused on reading for students in grades 7-12, however. Yet, these upper grades are exactly where, according to recent international studies, emphasis on reading is needed. The studies show that reading achievement of U.S. 4th graders ranks among the best in the world. However, by 8th grade, U.S. student performance declines to around the international average, and by 12th grade, our students rank even lower (Allington, 2001; National Center for Educational Statistics, 2001).

Why is this so? Reading demands increase dramatically for students around the 4th or 5th grade, when learning relies more on textbooks and other sources of information (Chall, 1983). The vocabulary encountered is less familiar because it contains more specialized or technical terms as content area subject matter increases. Syntax becomes more complex to handle ideas that are more complex. Greater reliance must be placed on inferential thinking and prior knowledge. More independent learning is expected than in lower grades, yet independent learning strategies are not directly taught at these levels.

When reading becomes the primary vehicle for learning, the demands on readers and the strategies they need to use in reading change. Unfortunately, just when the reading load increases and students shift from learning to read to reading to learn, most United States schools provide no corresponding instruction in reading to students. The scaffold of systematic and focused reading instruction typically diminishes or disappears altogether at the middle and high school levels in American schools. Student uses for reading begin to widen, and increasingly, schools begin to use more single-source instructional materials (textbooks, teacher handouts, etc.) for all students. (Baumann & Duffy, 1997)

Thus, a gap emerges between the overall reading ability levels of many students and the readability levels of the materials they are expected to read, a gap that widens with respect to reading demands beyond high school. The International Center has done extensive research on the reading comprehension ability of high school students versus the reading requirements for the real world (personal use reading, leisure reading, and occupational reading). Results show that a measurable gap exists. Three-quarters of all high school students will struggle to read or be unable to comprehend reading materials for entry-level jobs since the real world requires substantially higher levels of reading proficiency than most students possess.

States need to be sure that the reading proficiency levels set under *No Child Left Behind* reflect not just traditional academic measures of reading competence, but also the larger picture of what individuals will need for employability and success in life after graduation. We must also try harder to match students to instructional materials that will promote growth in reading proficiency without fostering frustration or limiting access to information needed for learning to occur. This broader view of reading competency is an example of the academic growth of individual students that must become part of program improvement under NCLB.

Studies like the one conducted by NWEA, along with benchmarking studies of the academic proficiency demands of the real world of adult roles such as those conducted by the International Center, will help us to fine-tune the standards of achievement mandated both by NCLB and by the necessity to help all students succeed in a rigorous and relevant curriculum.

References

Allington, Richard L. "You Can't Learn Much from Books You Can't Read," *Educational Leadership,* Vol. 60, No. 3, November 2002, pp. 16-19.

Chall, J.S. *Stages in Reading Development*. New York: McGraw-Hill, 1983.

Daggett, Willard R. (2003). "Achieving Reading for All." International Center for Leadership in Education. Rexford, NY.

Kingsbury, G. G., Olson, A., Cronin, J., Hauser, C., & Houser, R. (2003). *The State of State Standards: Research Investigating Proficiency Levels in Fourteen States*. Northwest Evaluation Association. Portland, OR.

The Princeton Review (2003). *Testing the Testers 2003: An Annual Ranking of State Accountability Systems.*