

Looking Worse Before We Look Better: Student Performance in a Common Core World

by MetaMetrics President and Co-founder Malbert Smith III, Ph.D.,
and Director of Professional Development Jason Turner

While political leaders and pundits wrangle over the various impending fiscal cliffs, it is important to call attention to a lesser known, but just as urgent, impending cliff in education. First, it is widely hoped that in 2013 we will finally see the reauthorization of the Elementary and Secondary Education Act (ESEA). The failure of Congress to reauthorize ESEA has led the Department of Education to circumvent issues of compliance by granting state waivers. However, these waivers are largely temporary, allowing states some breathing room until the reauthorization of ESEA can be addressed. The second critical event regards the final development and release of the common core assessments by two testing consortia, the Partnership for Assessment of Readiness for College and Careers (PARCC) and Smarter Balanced Assessment Consortium (SBAC). Thoughtful and careful deliberation of both of these factors will be required if we are to avoid a storm of uncertainty and confusion.

It has been almost a decade since Congress last reauthorized the Elementary and Secondary Education Act (ESEA) with the passage of the No Child Left Behind Act (NCLB). President Bush signed it into law on January 8, 2002, and at that time it represented a sweeping overhaul of federal efforts to support elementary and secondary education in the United States. Undertaken as a major effort to enhance the academic performance of *all* children, NCLB had many important and promising features that Congress hopefully retains when addressing the reauthorization—features such as disaggregation of data, annual assessments and expanded parental options. However, there are at least three flaws in the current bill that need to be corrected.

The first flaw in NCLB is its allowance for each state to define and set their own internal standards for proficiency. Fifty states determining their own “proficiency” standards, combined with the “proficiency” standards associated with the National Assessment of Educational Progress (NAEP), has resulted in an academic Tower of Babel when it comes to measuring progress in reading and math in our country.

The failure to establish clear, national standards of proficiency yielded highly disparate definitions of “proficiency” and, in practice, led to radically different ideas on what it meant to be proficient in reading and math. The unintended consequence of allowing states to set their own proficiency standards was that a “proficient 4th grade student in reading in Mississippi” was not the same as a “proficient 4th grade student in reading in Connecticut.” Even worse, the proficiency level set by the majority of states was not consistent with the rigor of NAEP. This lack of a common scale led to a general confusion among the public about whether our academic performance actually improved or not. In a 2004 white paper, *The Need for Objective Measurement Under the No Child Left Behind Act*, Malbert Smith wrote:

The serendipitous benefit of the high-stakes consequences of NCLB is that it will expose one of the most profound limitations of measurement in the social sciences: The lack of unification of metrics (universal and standard scales). Without universal, exchangeable scales in the social sciences, our assessment systems across states may employ the same labels (advanced, proficient, basic and below basic), but may vary dramatically in the achievement implied by these labels.

The second inherent flaw in NCLB is that it permitted states to back load their growth into the outer years. In other words, as Chester Finn of the Thomas Fordham Institute noted, this feature essentially allowed states to assume a balloon mortgage-type model when it came to ensuring that every student was proficient by school year 2013-14. Instead of requiring an equal amount of annual growth (like a fixed mortgage), much of the growth was deferred until the later years. An unintended consequence of this approach is that it became increasingly difficult for states to hit annual yearly progress (AYP) targets with each passing year.

A final and fundamental flaw of NCLB is that it led states to build accountability models that focused on

status measures of student performance as opposed to growth measures. Within the 670 pages of NCLB, terms like “achievement,” “progress,” “learning growth,” and “development” are used over 1,660 times. Unfortunately, these terms are used interchangeably, as if there is no distinction between growth and status. The bill could have been more accurately titled No “Cohort” Left Behind. The real and substantive concern of educators and parents involves growth, not status. Obviously, the reauthorization of a new bill presents an opportunity for correcting these previous flaws.

Around the same time that Congress will address these issues, we will be approaching the release of the common core assessments by PARCC and SBAC. While there is a healthy sense of optimism about these assessments, there are a number of nagging concerns that need addressing. The first and most obvious issue concerns cost. In today’s economic environment, all states are seeking to reduce the amount they spend on their annual assessments. Though most educators are optimistic about the promise of these long-awaited assessments, many privately worry that ballooning costs—costs far greater than many current state assessment programs—will render their state unable to afford the high price tag. While those fears are completely reasonable, a thoughtful analysis by Linda Darling-Hammond of Stanford University indicates that while the two assessments are likely to be more expensive than current assessments, there is good reason to believe that a number of creative solutions, including cost-sharing among states, may offer ways to significantly lower cost.

Assuming that cost can be brought in line with current state assessment budgets and that the assessments pass basic psychometric assumptions of reliability and validity, a major remaining issue will be how to communicate the assessment results to the public. If one subscribes to the notion that the “proficiency” levels on PARCC and SBAC assessments will likely be more rigorous than what most states have previously set, then great care and attention must be paid to how we communicate these results. To paraphrase Warren Buffett, only when the tide goes out can you tell who has been swimming without a bathing suit. The release of two “national assessments” will take the tide out! For example, assume a state reports that 75 percent of their third graders are proficient in reading in 2013. Then, in subsequent years, the state reports that (under the new assessment program) only 50 percent are proficient. To the uninitiated, it appears that the state has suffered a precipitous decline in reading growth, when, in fact,

it is the benchmark for what constitutes “proficiency,” along with the introduction of a common scale, that has been raised and is responsible for a perceived drop in student performance. It is easy to imagine how, divorced from context and history, these reports will provide headline fodder and can fuel an inaccurate picture of state educational progress.

When it comes to the implementation of the new assessments, it is helpful to take a page from the book of Wall Street. Publicly traded companies know the importance of managing expectations, and it is critical that education departments proactively shape expectations. Education departments would do well to take the lead in informing the public about what to expect from the reauthorization of ESEA and what the adoption of our new ‘national assessments’ will mean for the measurement of growth. Central to that effort will be the need to communicate the reason why state-specific multiple scales simply will not suffice as tools to measure our students’ progress in reading and mathematics. As Smith wrote in 2004:

Consequently, the real reason that the multiple measures requirement is on such a slippery slope is that our instruments do not have exchangeable scales. Without standard objective scales, like those employed in the hard sciences, educators will be left with less-than-satisfactory methods and very confusing, complicated schemes for reporting such data.

In fact, given the lack of exchangeable, highly disparate scales, the public should *expect* to see dips in student performance as we transition to a common metric and a unified way of measuring student growth.

While most of the nation’s attention has been focused on the financial cliffs (debt ceiling, sequestration, etc.), educators should pay special attention to the fog of uncertainty surrounding the reauthorization of NCLB and the release of the PARCC and SBAC assessments. A recent Gallup Poll indicates that the public’s confidence in our schools is at an all-time low. We analyzed over 40 years of Gallup data on this issue, along with empirical assessment data, to see if such a dismal public view was justified. We concluded our analyses with the steps that we think are needed to restore public confidence in education. In this transition into what could be a transformative moment, we are concerned that no child be left behind and that each student graduates college and career ready. During this time of transition—as we move toward a higher level of academic rigor and

proficiency standards—we may look worse before we look better. Getting this message right and helping the public understand this important transition will be necessary if we are to avoid further misperceptions and misunderstandings about the great work our nation's educators are doing.

REFERENCES

Darling-Hammond, L. & Adamson, F. (2010). "Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning." Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.

Finn, Jr., C. (2009). "Can we get to national standards, considering the pitfalls?" *The Education Gadfly*, The Thomas B. Fordham Institute.

Smith, M. (2004). "The Need for Objective Measurement Under the No Child Left Behind Act." MetaMetrics, Inc.

Smith, M., Turner, J., & Lattanzio, S. (2012). "Public Schools: Glass Half Full or Half Empty?" *Education Week*.

ABOUT THE AUTHORS:



Malbert Smith III, Ph.D., is president of MetaMetrics®, an educational measurement and research organization. Together with co-founder and CEO A. Jackson Stenner, Ph.D., Dr. Smith created The Lexile® Framework for Reading; El Sistema Lexile para Leer; The Lexile Framework for Writing; and The Quantile® Framework for Mathematics. Dr. Smith strives to make educational measurement actionable in the classroom and at home. His vision of common metrics for reading, writing and mathematics opens the way for differentiated instruction. In each state—and increasingly abroad—educators use Lexile and Quantile measures to blend instruction and assessment in whole-class and intervention settings. Concerned with the relationship between early literacy and college- and career-readiness, Dr. Smith led research to build a continuum of text complexity that places academic and life goals on the Lexile scale. He and Dr. Stenner were members of the team that contributed to the Common Core State Standards. They are also senior investigators on a National Center for Education Statistics research study to examine NAEP benchmark scores in relationship to college- and career- readiness. Dr. Smith serves on the UNC School of Education Foundation Board, the advisory board of Capstone Digital, and is a member of the advisory board for EdSteps, a joint project of the Council of Chief State School Officers and The Bill and Melinda Gates Foundation. He and Dr. Stenner are leading a three-year grant from The Gates Foundation on the efficacy of personalized learning platforms. Dr. Smith is a member of The American Association for the Advancement of Science, The American Educational Research Association and The National Council on Measurement in Education. He has taught graduate seminars in educational research and test development and design at Duke University and the University of North Carolina at Chapel Hill, from which he received the Distinguished Alumni Award. Dr. Smith frequently speaks at various events on educational research and measurement.



Jason Turner is the Professional Development Director at MetaMetrics, where he has played a variety of roles. Mr. Turner has managed implementations of and professional development for The Lexile Framework for Reading and The Quantile Framework for Mathematics. Additionally, he has overseen various state and district outreach efforts and developed content and training modules for the Lexile and Quantile Frameworks. Mr. Turner has also worked extensively in the training and implementation of the Lexile and Quantile Frameworks in classrooms across North America.

METAMETRICS® POLICY BRIEFS: MetaMetrics is focused on improving education for learners of all ages. For over twenty years, our work has been increasingly recognized for its distinct value in differentiating instruction and personalizing learning. Our research on postsecondary reading demands, for example, informed the Common Core State Standards for college- and career- readiness. In addition to the white papers and position papers we publish throughout the year, our policy briefs will encompass our research on a variety of educational issues, such as closing the achievement gap, next-generation assessments, and college and career readiness. The policy briefs will explore potential ways to address these critical issues by focusing on education as the foundation of student success and the stepping stone to social and economic growth in our country.

